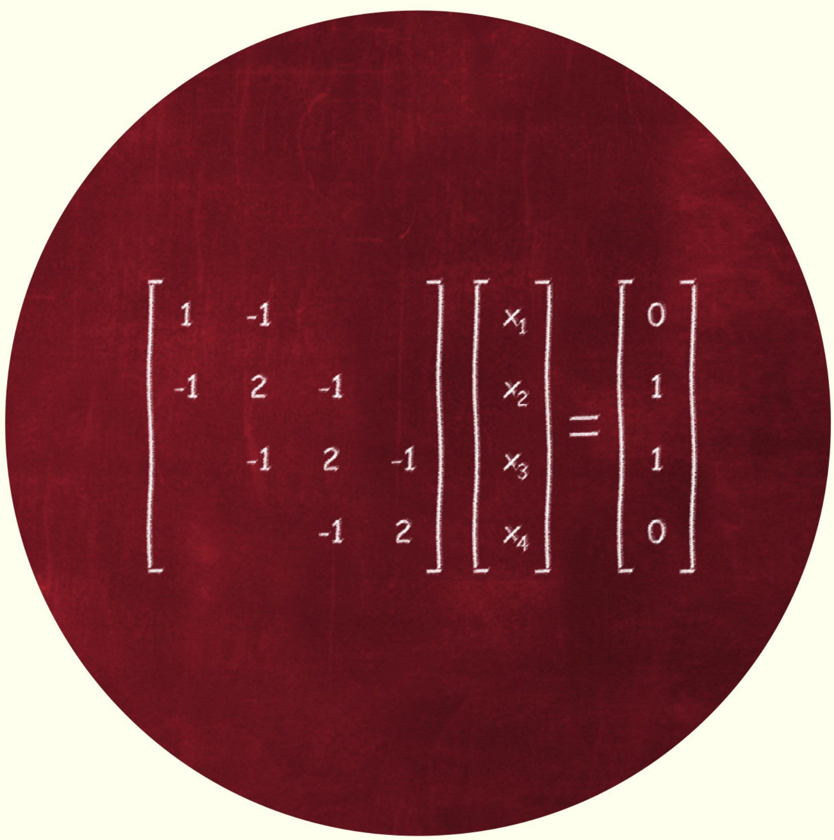
MATEMATISK ANALYS & LINJÄR ALGEBRA (III)

Linjär algebra och system av linjära ekvationer



STIG LARSSON   ANDERS LOGG   AXEL MÅLQVIST

Kompendiumutgåva tryckt på Chalmers tekniska högskola 3 mars 2021.

Innehåll

Förord och läsanvisningar

Denna bokserie i fyra delar ger en introduktion till matematisk analys och linjär algebra för teknisk högskola. Texten bygger på vår undervisning på Chalmers maskinteknikprogram och vår ambition har varit att skriva en bokserie som ger en koncis och lättillgänglig men samtidigt rigorös beskrivning av det matematiska teoribygget, med tydliga kopplingar till modellering, beräkning, algoritmer och programmering.

Vår ambition har också varit att ge studenten en bok som kan läsas från pärm till pärm, istället för en tegelsten sprängfylld med exempel och utvikningar; därav det kompakta formatet och det relativt sparsamma utrymme som ges åt lösta exempel.

Bokserien är strukturerad i fyra delar, med avsikt att varje del kan läsas som en kurs. De olika delarnas teman motiveras utifrån *ekvationslösning* som är ett centralt verktyg inom modellering och problemlösning i naturvetenskaperna och ingenjörskonsten. Upplägget kan därför sägas vara *problemmotiverat*. Varje del motiveras utifrån vår vilja att kunna lösa en viss klass av ekvationer: $f(x) = 0$ (skalära, ickelinjära ekvationer; del I), $u' = f$ (integraler och ordinära differentialekvationer; del II), $Ax = b$ (system av linjära ekvationer; del III) och $A(u) = f$ (system av partiella differentialekvationer; del IV). Dessa teman bildar en röd tråd som löper genom texten, men även om upplägget är problemmotiverat så är det *inte* problembaserat. Texten följer istället en traditionell matematisk beskrivning med definition, sats och bevis. Den som så önskar kan därför välja att istället tänka på bokseriens delar som differentialkalkyl (del I), integralkalkyl (del II), linjär algebra (del III) och flervariabelanalys (del IV).

Övningsuppgifterna är indelade i ”Övningar”, ”Problem” och ”Datorövningar”. Övningarna är standarduppgifter på specifika teman och avsedda för kunskapskontroll och mängdträning, medan problemen kombinerar olika teman och kan kräva mer eftertanke. Datorövningarna innebär matematisk problemlösning med hjälp av dator (programmering). Dessa övningar kan utföras i valfritt programmeringsspråk, exempelvis Python eller MATLAB.

Bevis och uppgifter märkta med ★ är av lite mer utmanande karaktär och ingår normalt inte i en standardkurs för godkänt betyg. I övrigt är rekommendationen att läsa (och förstå!) texten i sin helhet och att arbeta flitigt med övningar, problem och datorövningar. En lämplig omfattning kan vara att göra hälften av alla uppgifter; gör alla övningar märkta (a) och (b) samt alla udda problem och datorövningar. Resterande uppgifter kan fungera som extra träning eller repetitionsmaterial inför tentamen.

Lycka till!

Författarna
Göteborg, 23 september 2020

Tack!

Stort tack till de många studenter, övningsledare och kollegor som har läst, kommenterat och korrigerat tidiga utkast av bokserien: Thomas Bäckdahl, Rickard Cullman, Niklas Ericsson, Axel Flinth, Robert Forslund, Tomas Forssmark, Hussein Hamoodi, Felix Held, Andreas Jonsson, Per Ljung, Carl Lundholm, Douglas Molin, Vincent Molin, Aladdin Persson, Raad Salman och Joel Sjögren. Speciellt tack till Joar Axås och Christoffer Hansson som bidragit med många av bokseriens övningsuppgifter samt till Mikael Enelund för stöd och uppmuntran.

I arbetet med vår bokserie har vi tagit inspiration från ett flertal författare som på olika sätt har satt avtryck på både innehåll och stil, däribland Jan Petersson, Kenneth Eriksson, Donald Estep, Claes Johnson och Terrence Tao.

Ekvationen $Ax = b$

Läsperioden kan sägas handla om att lösa linjära ekvationssystem på formen $Ax = b$.

Introduktion

Linjära ekvationssystem kan skrivas på formen $Ax = b$. Här är A en matris och x och b vektorer av tal.

Linjär algebra handlar om att räkna med matriser och vektorer av allmän dimension. Innan vi börjar med detta ska vi studera vektorer i det vanliga 3-dimensionella rummet. Eftersom dessa definieras som riktade sträckor kallar vi dem *geometriska vektorer*. Med hjälp av dem kan vi studera geometri i rummet, till exempel, avstånd, vinkel, ortogonal projektion, area, volym, rät linje och plan. De spelar även en viktig roll i mekanik och fysik: kraftvektor, hastighetsvektor, rotationsvektor, elektrisk fältvektor, och så vidare.

Sedan presenterar vi en systematisk metod, Gauss eliminationsmetod, som transformerar ett linjärt ekvationssystem till en standardform där vi kan läsa av lösningsmängden. Därefter studerar vi hur vi räknar med matriser. Gauss-elimination kommer in här också. ...

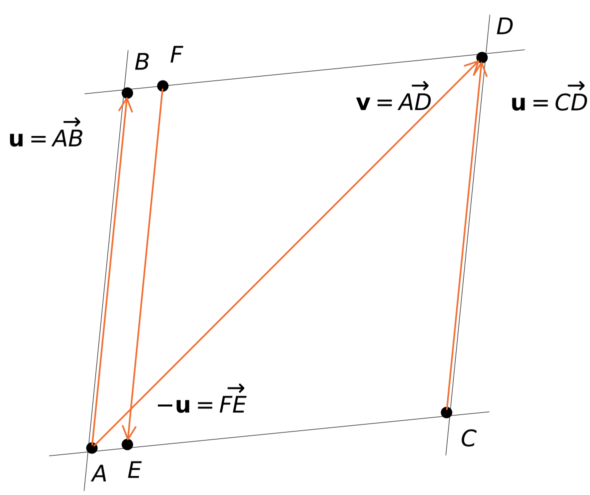
1. Vektorer i rummet

Vi studerar vektorer i rummet. Dessa kan vara geometriska vektorer, dvs riktade sträckor mellan punkter i rummet, men också fysikaliska vektorer såsom hastighet, kraft, elektriskt eller magnetiskt fält. Det gemensamma för dessa är att de karakteriseras endast av sin riktning och sin magnitud (längd). En annan gemensam sak är hur vi räknar med vektorer: vi kan addera vektorer och vi kan multiplicera vektorer med tal och på så sätt bilda nya vektorer. Vi ska också införa skalärprodukt och kryssprodukt av vektorer. Genom att införa ett Cartesiskt koordinatsystem i rummet kan vi identifiera geometriska vektorer med vektorer i $\mathbb{R}^3$*. Slutligen ska vi använda geometriska vektorer för att härleda ekvationer för den räta linjen och för planet. I senare kapitel ska vi generalisera till vektorer i* n *dimensioner, dvs vektorer i* $\mathbb{R}^n$*, men även abstrakta vektorer. Då kommer analogin med geometriska vektorer att hjälpa oss att visualisera den abstrakta teorin.*

1.1 Geometriska vektorer

Vi börjar med att definiera geometriska vektorer som riktade sträckor i rummet.

Definition 1.1 (Geometrisk vektor) Med en *geometrisk vektor* menas en riktad sträcka, dvs en sträcka med en bestämd genomloppsriktning. Den riktade sträckan från punkten A till punkten B tecknas $\overrightarrow{AB}$.
Längd. Avståndet mellan A och B kallas längden av vektorn $\overrightarrow{AB}$ och tecknas $|\overrightarrow{AB}|$.
Likhet. Två vektorer $\overrightarrow{AB}$ och $\overrightarrow{CD}$ sägs vara lika, $\overrightarrow{AB} = \overrightarrow{CD}$, om $\overrightarrow{AB}$ kan parallellförflyttas så att den sammanfaller med $\overrightarrow{CD}$.
Nollvektor. En vektor med samma start- och slutpunkt, $\overrightarrow{AA}$, kallas nollvektor. Alla nollvektorer är lika, $\overrightarrow{AA} = \overrightarrow{BB}$.
Motsatt vektor. Den motsatta vektorn till $\overrightarrow{AB}$ är vektorn $\overrightarrow{BA}$, vilken tecknas $-\overrightarrow{AB}$.



Figur 1.1: Geometriska vektorer. (De tunna linjerna är parallella.)

Två geometriska vektorer är lika om de har samma längd och samma riktning. En geometrisk vektor bestäms alltså entydigt av sin längd och sin riktning, men inte av sitt läge i rummet. Två vektorer som har samma riktning eller motsatt riktning (utan att nödvändigtvis ha samma längd), dvs som är parallella eller antiparallella, kallas *kolinjära*.

Vi skriver ofta vektorer med fetstil, till exempel, $\boldsymbol{u} = \overrightarrow{AB}$. Nollvektorn tecknas $\mathbf{0}$ och den motsatta vektorn till $\boldsymbol{u}$ tecknas $-\boldsymbol{u}$. Notera att $|-\boldsymbol{u}| = |\boldsymbol{u}|$ och att nollvektorn är den enda vektor vars längd är 0, dvs $|\boldsymbol{u}| = 0$ om och endast om $\boldsymbol{u} = \mathbf{0}$. Ibland skriver vi vektorns längd med samma bokstav i vanlig stil, $u = |\boldsymbol{u}|$.

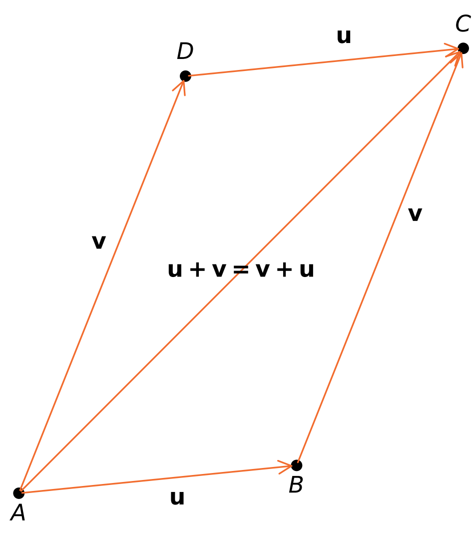
Exempel 1.1 (Fysiska vektorer) En fysisk vektor är en kvantitet som bestäms av en *riktning* och en *magnitud* som är en fysikalisk storhet. En geometrisk vektor $\boldsymbol{u} = \overrightarrow{AB}$ är en riktad sträcka i rummet och dess magnitud är därför *längd* med SI-enhet meter, $u = |\boldsymbol{u}|$ [m]. *Hastighet* $\boldsymbol{v}$ är en vektorkvantitet vars magnitud mäts i meter per sekund och kallas *fart*, $v = |\boldsymbol{v}|$ $[\mathrm{m\,s^{-1}}]$. *Kraft* $\boldsymbol{F}$ är en vektor vars magnitud mäts i newton, $F = |\boldsymbol{F}|$ [N]. *Elektrisk fältstyrka* $\boldsymbol{E}$ är vektor vars magnitud mäts i volt per meter, $E = |\boldsymbol{E}|$ $[\mathrm{V\,m^{-1}}]$.

1.2 Vektoralgebra

Vektorer adderas genom att vi parallellförflyttar den andra vektorn till spetsen av den första och tar den riktade sträckan som bildas.

Definition 1.2 (Addition av geometriska vektorer) Låt $\boldsymbol{u}$ och $\boldsymbol{v}$ vara vektorer och välj punkter A, B, C så att $\boldsymbol{u} = \overrightarrow{AB}$ och $\boldsymbol{v} = \overrightarrow{BC}$. Vektorn $\overrightarrow{AC}$ kallas summan av $\boldsymbol{u}$ och $\boldsymbol{v}$ och tecknas $\boldsymbol{u} + \boldsymbol{v}$ (se figur 1.2).

Med hjälp av den motsatta vektorn kan vi definiera *subtraktion* av vektorer: $\boldsymbol{u} - \boldsymbol{v} = \boldsymbol{u} + (-\boldsymbol{v})$.



Figur 1.2: Addition av vektorer.

Sats 1.1 (Räkneregler för addition)

$$\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u} \quad \text{(kommutativa lagen för addition)} \tag{1.1}$$
$$(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w}) \quad \text{(associativa lagen för addition)} \tag{1.2}$$
$$\boldsymbol{u}+\boldsymbol{0}=\boldsymbol{u} \quad \text{(addition av nollvektorn)} \tag{1.3}$$
$$\boldsymbol{u}+(-\boldsymbol{u})=\boldsymbol{0} \quad \text{(addition av motsatta vektorn)} \tag{1.4}$$

Bevis. Bevis av (1.1): Välj punkter så att $\boldsymbol{u}=\overrightarrow{AB}=\overrightarrow{DC}$, $\boldsymbol{v}=\overrightarrow{BC}=\overrightarrow{AD}$ (se figur 1.2). Då bildar punkterna en parallellogram $ABCD$ som spänns upp av vektorerna $\boldsymbol{u}$ och $\boldsymbol{v}$. Vi får

$$\boldsymbol{u}+\boldsymbol{v}=\overrightarrow{AB}+\overrightarrow{BC}=\overrightarrow{AC} \tag{1.5}$$
$$\boldsymbol{v}+\boldsymbol{u}=\overrightarrow{AD}+\overrightarrow{DC}=\overrightarrow{AC} \tag{1.6}$$

Bevis av (1.2): Välj nu punkter så att $\boldsymbol{u}=\overrightarrow{AB}$, $\boldsymbol{v}=\overrightarrow{BC}$, $\boldsymbol{w}=\overrightarrow{CD}$. Då fås

$$(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=(\overrightarrow{AB}+\overrightarrow{BC})+\overrightarrow{CD}=\overrightarrow{AC}+\overrightarrow{CD}=\overrightarrow{AD} \tag{1.7}$$
$$\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})=\overrightarrow{AB}+(\overrightarrow{BC}+\overrightarrow{CD})=\overrightarrow{AB}+\overrightarrow{BD}=\overrightarrow{AD} \tag{1.8}$$

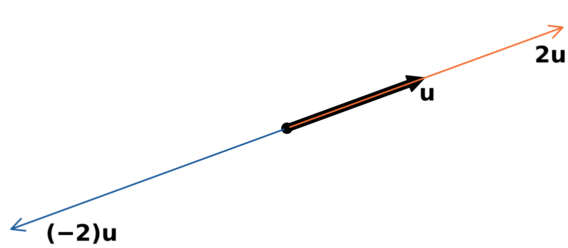
Bevis av (1.3):

$$\boldsymbol{u}+\boldsymbol{0}=\overrightarrow{AB}+\overrightarrow{BB}=\overrightarrow{AB}=\boldsymbol{u} \tag{1.9}$$

Bevis av (1.4):

$$\boldsymbol{u}+(-\boldsymbol{u})=\overrightarrow{AB}+\overrightarrow{BA}=\overrightarrow{AA}=\boldsymbol{0} \tag{1.10}$$

□



Figur 1.3: Multiplikation av vektor med skalär.

Definition 1.3 (Multiplikation med skalär) Låt $\boldsymbol{u}$ vara en vektor och t en skalär, dvs ett reellt tal. Då är produkten $t\boldsymbol{u}$ den unika vektor som har längden $|t\boldsymbol{u}| = |t||\boldsymbol{u}|$ och som är parallell med $\boldsymbol{u}$ om $t > 0$ och antiparallell med $\boldsymbol{u}$ om $t < 0$ (se figur 1.3).

Notera att i uttrycket $|t||\boldsymbol{u}|$ är $|t|$ absolutbeloppet av ett tal medan $|\boldsymbol{u}|$ är längden av en vektor. Definitionen innebär att två vektorer $\boldsymbol{u}, \boldsymbol{v}$ är kolinjära om och endast om det finns $t \in \mathbb{R}, t \neq 0$, så att

$$\boldsymbol{u} = t\boldsymbol{v} \tag{1.11}$$

Sats 1.2 (Räkneregler för multiplikation med skalär)

$$t(\boldsymbol{u} + \boldsymbol{v}) = t\boldsymbol{u} + t\boldsymbol{v} \qquad \text{(distributiv lag)} \tag{1.12}$$
$$(t + s)\boldsymbol{u} = t\boldsymbol{u} + s\boldsymbol{u} \qquad \text{(distributiv lag)} \tag{1.13}$$
$$t(s\boldsymbol{u}) = (ts)\boldsymbol{u} \qquad \text{(associativ lag)} \tag{1.14}$$
$$1\boldsymbol{u} = \boldsymbol{u} \tag{1.15}$$
$$(-1)\boldsymbol{u} = -\boldsymbol{u} \tag{1.16}$$
$$0\boldsymbol{u} = \boldsymbol{0} \tag{1.17}$$

Bevis. Vi lämnar beviset till problem 1.2. □

En vektor som inte är nollvektorn, $\boldsymbol{u} \neq \boldsymbol{0}$, kan *normeras*. Det innebär att vi bestämmer en parallell vektor $\hat{\boldsymbol{u}}$ som har längden 1, dvs $\hat{\boldsymbol{u}} = t\boldsymbol{u}$ med $t > 0$ och $1 = |\hat{\boldsymbol{u}}| = |t\boldsymbol{u}| = t|\boldsymbol{u}|$. Vi får $t = 1/|\boldsymbol{u}|$ och

$$\hat{\boldsymbol{u}} = \frac{1}{|\boldsymbol{u}|}\boldsymbol{u} = \frac{\boldsymbol{u}}{|\boldsymbol{u}|} \tag{1.18}$$

Den normerade vektorn $\hat{\boldsymbol{u}}$ har samma riktning som $\boldsymbol{u}$ men längden 1, $|\hat{\boldsymbol{u}}| = 1$. En sådan vektor kallas *enhetsvektor*.

Vektorn

$$\boldsymbol{w} = s\boldsymbol{u} + t\boldsymbol{v} \tag{1.19}$$

kallas *linjär kombination* av vektorerna $\boldsymbol{u}$ och $\boldsymbol{v}$ med koefficienterna s och t. Linjär kombination är ett viktigt sätt att bilda nya vektorer. Mer allmänt bildar vi linjär kombination av flera vektorer.

Definition 1.4 (Linjär kombination) Låt $\{\boldsymbol{v}_j\}_{j=1}^n$ vara geometriska vektorer och $\{c_j\}_{j=1}^n$ vara skalärer. Vektorn

$$v = \sum_{j=1}^{n} c_j \boldsymbol{v}_j = c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n \tag{1.20}$$

kallas linjär kombination av $\{\boldsymbol{v}_j\}_{j=1}^n$ med koefficienter $\{c_j\}_{j=1}^n$.

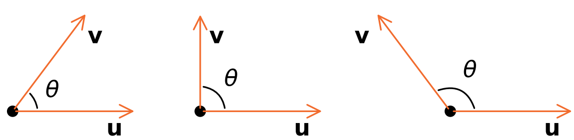
1.3 Skalärprodukt, ortogonalitet och projektion

Definition 1.5 (Skalärprodukt) Den skalära produkten $\boldsymbol{u} \cdot \boldsymbol{v}$ definieras av

$$\boldsymbol{u} \cdot \boldsymbol{v} = |\boldsymbol{u}||\boldsymbol{v}| \cos(\theta) \tag{1.21}$$

där θ är vinkeln mellan $\boldsymbol{u}$ och $\boldsymbol{v}$ (se figur 1.4).

Notera att skalärprodukten skrivs med en prick, $\boldsymbol{u} \cdot \boldsymbol{v}$. Uttrycket $\boldsymbol{u}\boldsymbol{v}$ har ingen mening. Vinkeln θ varierar mellan 0 och π. Om en av vektorerna är nollvektorn, till exempel $\boldsymbol{u} = \boldsymbol{0}$, så är vinkeln θ odefinierad men skalärprodukten är 0 eftersom $|\boldsymbol{u}| = 0$.



Figur 1.4: Vänster: $\cos(\theta) > 0$, mitten: $\cos(\theta) = 0$, höger: $\cos(\theta) < 0$.

Sats 1.3 (Räkneregler för skalärprodukt)

$$\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u} \quad \text{(kommutativ lag)} \tag{1.22}$$

$$\boldsymbol{u} \cdot (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{u} \cdot \boldsymbol{w} \quad \text{(distributiv lag)} \tag{1.23}$$

$$(t\boldsymbol{u}) \cdot \boldsymbol{v} = t(\boldsymbol{u} \cdot \boldsymbol{v}) = \boldsymbol{u} \cdot (t\boldsymbol{v}) \quad \text{(distributiv lag)} \tag{1.24}$$

$$\boldsymbol{u} \cdot \boldsymbol{u} = |\boldsymbol{u}|^2 \tag{1.25}$$

Bevis. Bevisen av (1.22), (1.24), (1.25) följer omedelbart av definitionen i (1.21). För (1.24) noterar vi då att vinkeln mellan $t\boldsymbol{u}$ och $\boldsymbol{v}$ är $\pi - \theta$ då $t < 0$ och att $\cos(\pi - \theta) = -\cos(\theta)$. Distributiva lagen (1.23) bevisar vi senare med hjälp av ortogonal projektion som vi strax ska introducera. □

Skalärproduktens tecken indikerar vinkelns typ. Antag att $\boldsymbol{u} \neq \boldsymbol{0}$, $\boldsymbol{v} \neq \boldsymbol{0}$. Vi har tre fall (se figur 1.4):

1. $\boldsymbol{u} \cdot \boldsymbol{v} > 0$. Då är $\cos(\theta) > 0$, dvs $0 \leqslant \theta < \pi/2$, och vektorerna bildar en *spetsig vinkel*.

2. $\boldsymbol{u} \cdot \boldsymbol{v} = 0$. Då är $\cos(\theta) = 0$, dvs $\theta = \pi/2$, och vektorerna bildar en *rät vinkel*.

3. $\boldsymbol{u} \cdot \boldsymbol{v} < 0$. Då är $\cos(\theta) < 0$, dvs $\pi/2 < \theta \leqslant \pi$, och vektorerna bildar en *trubbig vinkel*.

Fallet med rät vinkel är så viktigt att vi gör en definition.

Definition 1.6 (Ortogonala vektorer) Vektorerna $\boldsymbol{u}$ och $\boldsymbol{v}$ kallas *ortogonala* om

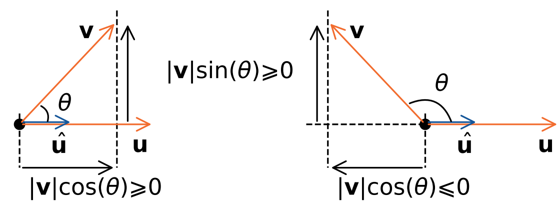
$$\boldsymbol{u} \cdot \boldsymbol{v} = 0 \tag{1.26}$$

Vi noterar att nollvektorn per definition är ortogonal mot alla vektorer.

Vi kan nu introducera begreppet *ortogonal projektion*. Antag $\boldsymbol{u} \neq \mathbf{0}$ och bilda enhetsvektorn $\hat{\boldsymbol{u}} = \boldsymbol{u}/|\boldsymbol{u}|$. I figur 1.5 ser vi att talet $|\boldsymbol{v}|\cos(\theta)$ fås genom att projicera vektorn $\boldsymbol{v}$ ortogonalt på $\boldsymbol{u}$. Men vi har ju $\boldsymbol{u} \cdot \boldsymbol{v} = |\boldsymbol{u}||\boldsymbol{v}|\cos(\theta)$ och projektionen ges alltså av

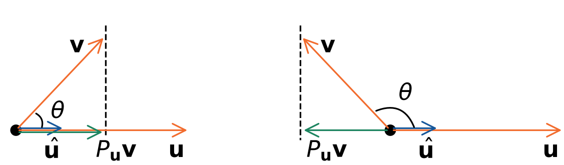
$$|\boldsymbol{v}|\cos(\theta) = \frac{\boldsymbol{v} \cdot \boldsymbol{u}}{|\boldsymbol{u}|} = \boldsymbol{v} \cdot \frac{\boldsymbol{u}}{|\boldsymbol{u}|} = \boldsymbol{v} \cdot \hat{\boldsymbol{u}} \tag{1.27}$$

Detta är en skalär som kan vara positiv, negativ eller 0 beroende på vinkeln. Den kallas *kom-*



Figur 1.5: Komponenten $|\boldsymbol{v}|\cos(\theta)$ kan ha olika tecken, men höjden $|\boldsymbol{v}|\sin(\theta)$ är alltid icke-negativ eftersom $\theta \in [0, \pi]$.

ponenten av $\boldsymbol{v}$ längs riktningen $\hat{\boldsymbol{u}}$ eller den skalära projektionen. Om vi multiplicerar den med enhetsvektorn $\hat{\boldsymbol{u}}$ får vi en vektor $(\boldsymbol{v} \cdot \hat{\boldsymbol{u}})\hat{\boldsymbol{u}}$ som är kolinjär med $\boldsymbol{u}$ och vars längd är lika med absolutbeloppet av den skalära projektionen. Denna kallas vektorprojektionen eller helt enkelt ortogonala projektionen av $\boldsymbol{v}$ på $\boldsymbol{u}$ (se figur 1.6).



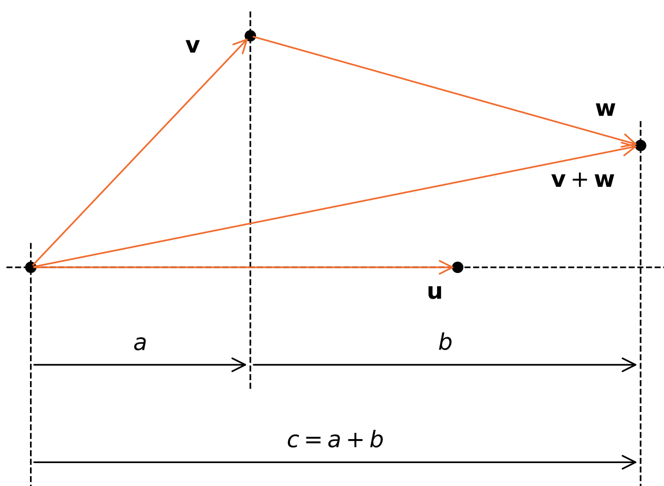
Figur 1.6: Ortogonal projektion.

Definition 1.7 (Ortogonal projektion) Antag $\boldsymbol{u} \neq \mathbf{0}$ och bilda enhetsvektorn $\hat{\boldsymbol{u}} = \boldsymbol{u}/|\boldsymbol{u}|$. *Komponenten av $\boldsymbol{v}$ längs riktningen $\hat{\boldsymbol{u}}$* (skalära projektionen) är

$$\boldsymbol{v} \cdot \hat{\boldsymbol{u}} = \frac{\boldsymbol{v} \cdot \boldsymbol{u}}{|\boldsymbol{u}|} \tag{1.28}$$

Den *ortogonala projektionen* av $\boldsymbol{v}$ på $\boldsymbol{u}$ är

$$P_{\boldsymbol{u}}(\boldsymbol{v}) = (\boldsymbol{v} \cdot \hat{\boldsymbol{u}})\hat{\boldsymbol{u}} = \frac{\boldsymbol{v} \cdot \boldsymbol{u}}{|\boldsymbol{u}|^2}\boldsymbol{u} \tag{1.29}$$



Figur 1.7: Bevis av distributiva lagen.

Vi kan nu bevisa distributiva lagen (1.23) för skalär produkt:

$$\boldsymbol{u} \cdot (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{u} \cdot \boldsymbol{w} \tag{1.30}$$

Om $\boldsymbol{u} = \mathbf{0}$ så är båda sidor lika med noll. Annars ser vi från (1.28) att alla tre termerna är lika med $|\boldsymbol{u}|$ gånger respektive skalära projektion på $\boldsymbol{u}$:

$$\boldsymbol{u} \cdot \boldsymbol{v} = |\boldsymbol{u}|\, \boldsymbol{v} \cdot \hat{\boldsymbol{u}} \tag{1.31}$$
$$\boldsymbol{u} \cdot \boldsymbol{w} = |\boldsymbol{u}|\, \boldsymbol{w} \cdot \hat{\boldsymbol{u}} \tag{1.32}$$
$$\boldsymbol{u} \cdot (\boldsymbol{v} + \boldsymbol{w}) = |\boldsymbol{u}|\, (\boldsymbol{v} + \boldsymbol{w}) \cdot \hat{\boldsymbol{u}} \tag{1.33}$$

Men projektionerna av $\boldsymbol{v}$ och $\boldsymbol{w}$ adderas längs en rät linje som är parallell med $\boldsymbol{u}$ och summan är projektionen av $\boldsymbol{v} + \boldsymbol{w}$:

$$(\boldsymbol{v} + \boldsymbol{w}) \cdot \hat{\boldsymbol{u}} = \boldsymbol{v} \cdot \hat{\boldsymbol{u}} + \boldsymbol{w} \cdot \hat{\boldsymbol{u}} \tag{1.34}$$

dvs $c = a + b$ i figur 1.7, varav (1.30) följer. Alternativt kan vi skriva med formler:

$$P_{\boldsymbol{u}}\boldsymbol{v} + P_{\boldsymbol{u}}\boldsymbol{w} = (\boldsymbol{v}\cdot\hat{\boldsymbol{u}})\hat{\boldsymbol{u}} + (\boldsymbol{w}\cdot\hat{\boldsymbol{u}})\hat{\boldsymbol{u}} = \Big((\boldsymbol{v}+\boldsymbol{w})\cdot\hat{\boldsymbol{u}}\Big)\hat{\boldsymbol{u}} = P_{\boldsymbol{u}}(\boldsymbol{v}+\boldsymbol{w}) \tag{1.35}$$

dvs

$$\Big(\boldsymbol{v}\cdot\hat{\boldsymbol{u}} + \boldsymbol{w}\cdot\hat{\boldsymbol{u}} - (\boldsymbol{v}+\boldsymbol{w})\cdot\hat{\boldsymbol{u}}\Big)\hat{\boldsymbol{u}} = \boldsymbol{0} \tag{1.36}$$

Vi multiplicerar skalärt med $\hat{\boldsymbol{u}}$:

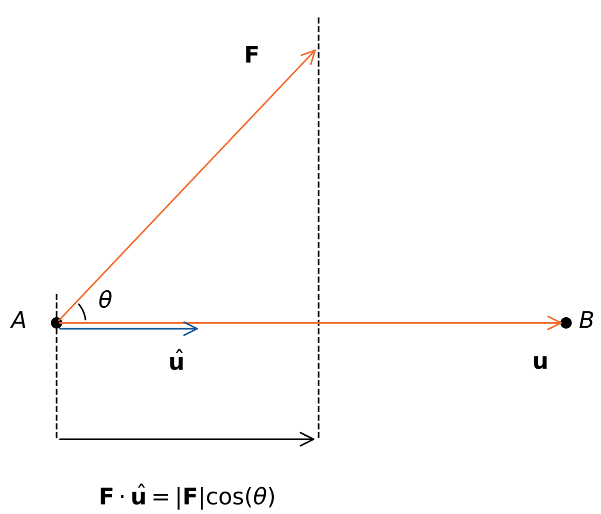
$$\Big(\boldsymbol{v}\cdot\hat{\boldsymbol{u}} + \boldsymbol{w}\cdot\hat{\boldsymbol{u}} - (\boldsymbol{v}+\boldsymbol{w})\cdot\hat{\boldsymbol{u}}\Big)\underbrace{\hat{\boldsymbol{u}}\cdot\hat{\boldsymbol{u}}}_{=1} = \underbrace{\boldsymbol{0}\cdot\hat{\boldsymbol{u}}}_{=0} \tag{1.37}$$

vilket leder till (1.34).

Exempel 1.2 (Arbete) En kropp förflyttas längs den riktade sträckan $\boldsymbol{u} = \overrightarrow{AB}$ [m] under inverkan av en kraft $\boldsymbol{F}$ [N]. Arbetet som uträttas blir då lika med komponenten av kraften längs $\boldsymbol{u}$, dvs $\boldsymbol{F}\cdot\hat{\boldsymbol{u}} = |\boldsymbol{F}|\cos(\theta)$, gånger den tillryggalagda sträckan $|\boldsymbol{u}|$, se figur 1.8:

$$W = (\boldsymbol{F}\cdot\hat{\boldsymbol{u}})|\boldsymbol{u}| = |\boldsymbol{u}||\boldsymbol{F}|\cos(\theta) = \boldsymbol{F}\cdot\boldsymbol{u} \quad [\mathrm{J}] \tag{1.38}$$

Arbetet ges helt enkelt som skalärprodukten av kraften och den riktade sträckan. Arbetet kan vara positivt, negativt eller noll beroende på tecknet på $\cos(\theta)$, dvs beroende på hur kraften är riktad i förhållande till rörelsen. (SI-enheten för arbete är joule = newtonmeter, J =N m.)



Figur 1.8: Arbete.

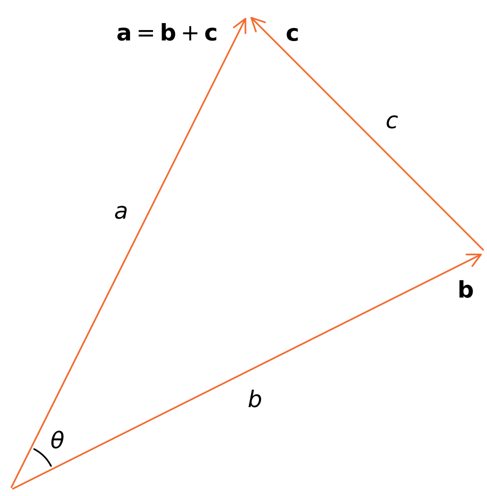
Sambandet (1.25), $|\boldsymbol{u}|^2 = \boldsymbol{u}\cdot\boldsymbol{u}$, är mycket användbart, vilket följande två exempel visar.

Exempel 1.3 (Beräkna längd) Vektorerna $\boldsymbol{u}$ och $\boldsymbol{v}$ har längderna 2 respektive 3 och bildar vinkeln $\pi/3$. Vi beräknar längden av vektorn $\boldsymbol{u} - 5\boldsymbol{v}$.

Vi använder (1.25) och utvecklar kvadraten med hjälp av räknereglerna:

$$\begin{aligned}
|\boldsymbol{u} - 5\boldsymbol{v}|^2 &= (\boldsymbol{u} - 5\boldsymbol{v}) \cdot (\boldsymbol{u} - 5\boldsymbol{v}) && (1.39)\\
&= \boldsymbol{u} \cdot \boldsymbol{u} + \boldsymbol{u} \cdot (-5\boldsymbol{v}) + (-5\boldsymbol{v}) \cdot \boldsymbol{u} + (-5\boldsymbol{v}) \cdot (-5\boldsymbol{v}) && (1.40)\\
&= \boldsymbol{u} \cdot \boldsymbol{u} + 2\boldsymbol{u} \cdot (-5\boldsymbol{v}) + (-5\boldsymbol{v}) \cdot (-5\boldsymbol{v}) && (1.41)\\
&= |\boldsymbol{u}|^2 + 2(-5)\boldsymbol{u} \cdot \boldsymbol{v} + |-5\boldsymbol{v}|^2 && (1.42)\\
&= |\boldsymbol{u}|^2 - 10|\boldsymbol{u}||\boldsymbol{v}|\cos(\pi/3) + 25|\boldsymbol{v}|^2 && (1.43)\\
&= 4 - 60\cos(\pi/3) + 225 = 4 - 30 + 225 = 199 && (1.44)
\end{aligned}$$

Alltså: $|\boldsymbol{u} - 5\boldsymbol{v}| = \sqrt{199}$. Observera den dubbla produkten i (1.41).



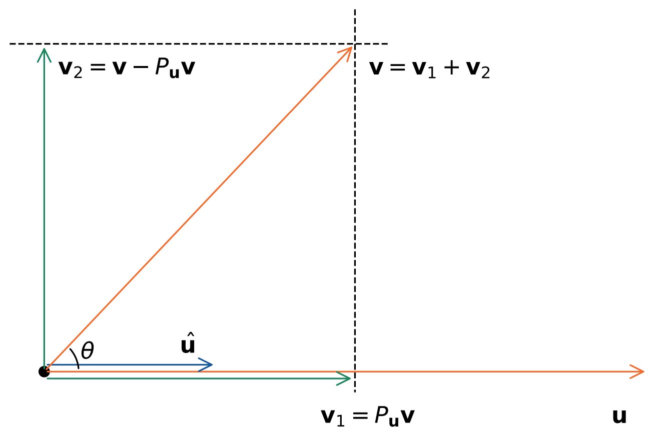
Figur 1.9: Cosinussatsen.

Exempel 1.4 (Cosinussatsen) En triangel har sidor med längderna a, b, c och vinkeln mellan sidorna a och b är θ. Då gäller

$$c^2 = a^2 + b^2 - 2ab\cos(\theta) \qquad (1.45)$$

Bevis: Bilda vektorer $\boldsymbol{a}$, $\boldsymbol{b}$ och $\boldsymbol{c}$ längs de tre sidorna så att $\boldsymbol{a} = \boldsymbol{b} + \boldsymbol{c}$ och $a = |\boldsymbol{a}|$, $b = |\boldsymbol{b}|$, $c = |\boldsymbol{c}|$ (se figur 1.9). Då fås

$$\begin{aligned}
c^2 = |\boldsymbol{c}|^2 = |\boldsymbol{a} - \boldsymbol{b}|^2 = (\boldsymbol{a} - \boldsymbol{b}) \cdot (\boldsymbol{a} - \boldsymbol{b}) &= \boldsymbol{a} \cdot \boldsymbol{a} - 2\boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{b} \cdot \boldsymbol{b} && (1.46)\\
= |\boldsymbol{a}|^2 - 2\boldsymbol{a} \cdot \boldsymbol{b} + |\boldsymbol{b}|^2 &= a^2 + b^2 - 2ab\cos(\theta) && (1.47)
\end{aligned}$$



Figur 1.10: Ortogonal uppdelning.

Exempel 1.5 (Uppdelning i ortogonala komposanter) Antag $\boldsymbol{u} \neq \mathbf{0}$. Varje vektor $\boldsymbol{v}$ kan delas upp i två unika komposanter, $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$, där $\boldsymbol{v}_1$ är kolinjär med $\boldsymbol{u}$ och $\boldsymbol{v}_2$ är ortogonal mot $\boldsymbol{u}$. Komposanterna är $\boldsymbol{v}_1 = P_{\boldsymbol{u}}(\boldsymbol{v}) = (\boldsymbol{v} \cdot \hat{\boldsymbol{u}})\hat{\boldsymbol{u}}$ och $\boldsymbol{v}_2 = \boldsymbol{v} - P_{\boldsymbol{u}}(\boldsymbol{v})$, dvs den ortogonal uppdelningen är (se figur 1.10)

$$\boldsymbol{v} = P_{\boldsymbol{u}}(\boldsymbol{v}) + (\boldsymbol{v} - P_{\boldsymbol{u}}(\boldsymbol{v})) \tag{1.48}$$

Bevis: Att $\boldsymbol{v}_1$ är kolinjär med $\boldsymbol{u}$ betyder att $\boldsymbol{v}_1 = t\boldsymbol{u}$ för någon skalär t. Vi bestämmer t så att $\boldsymbol{v}_2 = \boldsymbol{v} - t\boldsymbol{u}$ blir ortogonal mot $\boldsymbol{u}$, dvs

$$0 = (\boldsymbol{v} - t\boldsymbol{u}) \cdot \boldsymbol{u} = \boldsymbol{v} \cdot \boldsymbol{u} - t\boldsymbol{u} \cdot \boldsymbol{u} = \boldsymbol{v} \cdot \boldsymbol{u} - t|\boldsymbol{u}|^2 \tag{1.49}$$

$$t = \frac{\boldsymbol{v} \cdot \boldsymbol{u}}{|\boldsymbol{u}|^2} = \boldsymbol{v} \cdot \hat{\boldsymbol{u}} \tag{1.50}$$

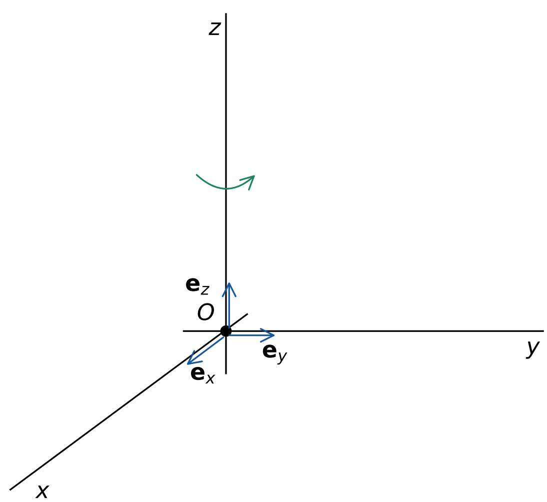
Alltså:

$$\boldsymbol{v}_1 = t\boldsymbol{u} = \frac{\boldsymbol{v} \cdot \boldsymbol{u}}{|\boldsymbol{u}|^2}\boldsymbol{u} = (\boldsymbol{v} \cdot \hat{\boldsymbol{u}})\hat{\boldsymbol{u}} = P_{\boldsymbol{u}}(\boldsymbol{v}) \tag{1.51}$$

$$\boldsymbol{v}_2 = \boldsymbol{v} - P_{\boldsymbol{u}}(\boldsymbol{v}) \tag{1.52}$$

1.4 Koordinatsystem

Vi inför nu ett koordinatsystem i rummet. Det består av tre axlar som är parvis ortogonala och som skär varandra i en punkt O, som kallas origo. Axlarna kallas x-, y- och z-axlarna. På respektive axel placerar vi enhetsvektorer $\boldsymbol{e}_x$, $\boldsymbol{e}_y$, $\boldsymbol{e}_z$ som pekar ut från origo och definierar en positiv riktning till respektive axel. Dessa vektorer kallas *basvektorer*. Ett sådant koordinatsystem med ortogonala och normerade basvektorer kallas *Cartesiskt*. Vi antar dessutom att det utgör ett *högersystem*, dvs z-axeln är riktad i den riktning som en högergängad skruv rör sig då den vrids den kortaste vägen från x-axeln till y-axeln. Alternativt kan vi säga att x-, y- och z-axlarna är orienterade så som tumme, pekfinger respektive långfinger på höger hand. Vi har alltså ett Cartesiskt högersystem (se figur 1.11).

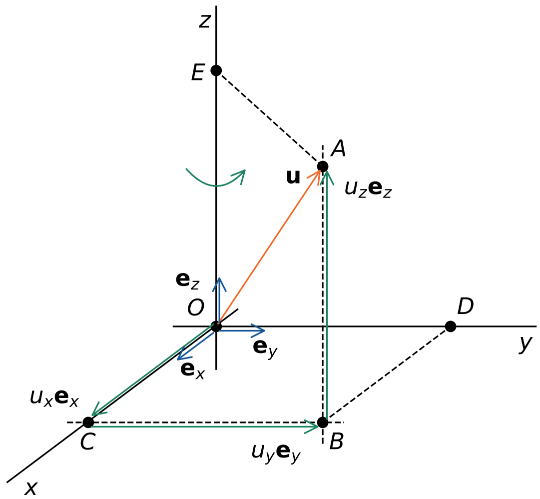


Figur 1.11: Cartesiskt högersystem.

Att basvektorerna är ortogonala och normerade, eller kortfattat *ortonormerade*, kan formuleras så här:

$$\begin{aligned} \boldsymbol{e}_x \cdot \boldsymbol{e}_x = \boldsymbol{e}_y \cdot \boldsymbol{e}_y = \boldsymbol{e}_z \cdot \boldsymbol{e}_z = 1 \quad &\text{(normerade)} \\ \boldsymbol{e}_x \cdot \boldsymbol{e}_y = \boldsymbol{e}_x \cdot \boldsymbol{e}_z = \boldsymbol{e}_y \cdot \boldsymbol{e}_z = 0 \quad &\text{(ortogonala)} \end{aligned} \tag{1.53}$$

Vi ska nu dela upp en godtycklig vektor i komposanter längs basvektorerna. Att vektorer kan uttryckas på detta vis med precis tre basvektorer innebär att rummet är tre-dimensionellt.



Figur 1.12: Uppdelning i komposanter.

Sats 1.4 (Uppdelning i komposanter) Varje vektor $\boldsymbol{u}$ kan skrivas

$$\boldsymbol{u} = u_x\boldsymbol{e}_x + u_y\boldsymbol{e}_y + u_z\boldsymbol{e}_z \tag{1.54}$$

där skalärerna u_x, u_y, u_z är unika. De ges av

$$u_x = \boldsymbol{u}\cdot\boldsymbol{e}_x,\ u_y = \boldsymbol{u}\cdot\boldsymbol{e}_y,\ u_z = \boldsymbol{u}\cdot\boldsymbol{e}_z \tag{1.55}$$

Vektorerna $u_x\boldsymbol{e}_x$, $u_y\boldsymbol{e}_y$, $u_z\boldsymbol{e}_z$ kallas *komposanter* medan talen (koefficienterna) u_x, u_y, u_z kallas *komponenter*.

Bevis. Vi visar först att sådana tal finns. Låt $\boldsymbol{u} = \overrightarrow{OA}$, dvs den riktade sträckan från origo till någon punkt A (se figur 1.12). Dra sedan en rät linje genom A parallellt med z-axeln. Den skär xy-planet i en punkt B. Genom B drar vi räta linjer parallella med x-axeln respektive y-axeln som skär axlarna i punkter C respektive D. Då har vi

$$\boldsymbol{u} = \overrightarrow{OA} = \overrightarrow{OC} + \overrightarrow{CB} + \overrightarrow{BA} \tag{1.56}$$

Eftersom $\overrightarrow{OC}$, $\overrightarrow{CB}$, $\overrightarrow{BA}$ är kolinjära med basvektorerna $\boldsymbol{e}_x$, $\boldsymbol{e}_y$, respektive $\boldsymbol{e}_z$, så finns tal u_x, u_y, u_z sådana att

$$\overrightarrow{OC} = u_x\boldsymbol{e}_x, \quad \overrightarrow{CB} = u_y\boldsymbol{e}_y, \quad \overrightarrow{BA} = u_z\boldsymbol{e}_z \tag{1.57}$$

och (1.54) följer.

Genom parellförflyttning ser vi att komposanterna $u_y\boldsymbol{e}_y = \overrightarrow{OD}$ och $u_z\boldsymbol{e}_z = \overrightarrow{OE}$ så att vi har även den ortogonala uppdelningen

$$\boldsymbol{u} = \overrightarrow{OA} = \overrightarrow{OC} + \overrightarrow{OD} + \overrightarrow{OE} \tag{1.58}$$

Komponenterna u_x, u_y och u_z är alltså avstånden (med tecken) från origo längs axlarna till punkterna C, D respektive E.

För att visa att komponenterna är unika antar vi att (1.54) gäller. Vi bildar skalärprodukten av (1.54) med $\boldsymbol{e}_x$:

$$\boldsymbol{u} \cdot \boldsymbol{e}_x = u_x \boldsymbol{e}_x \cdot \boldsymbol{e}_x + u_y \boldsymbol{e}_y \cdot \boldsymbol{e}_x + u_z \boldsymbol{e}_z \cdot \boldsymbol{e}_x = u_x \tag{1.59}$$

där vi använde att basvektorerna är ortonormerade, (1.53). Alltså: talet

$$u_x = \boldsymbol{u} \cdot \boldsymbol{e}_x \tag{1.60}$$

är den skalära projektionen av $\boldsymbol{u}$ på $\boldsymbol{e}_x$. På liknande sätt får vi

$$u_y = \boldsymbol{u} \cdot \boldsymbol{e}_y, \; u_z = \boldsymbol{u} \cdot \boldsymbol{e}_z \tag{1.61}$$

Talen är alltså entydigt bestämda. □

Om vi har valt ett koordinatsystem så är det praktiskt att skriva vektorer med hjälp av komponenter, dvs vi skriver kortfattat

$$\boldsymbol{u} = (u_x, u_y, u_z) \tag{1.62}$$

istället för att skriva med komposanter $\boldsymbol{u} = u_x \boldsymbol{e}_x + u_y \boldsymbol{e}_y + u_z \boldsymbol{e}_z$. Skrivsättet (1.62) innebär att vi identifierar vektorn med en taltrippel, $(u_x, u_y, u_z) \in \mathbb{R}^3$. I senare kapitel ska vi arbeta med vektorer i n dimensioner, dvs i $\mathbb{R}^n$.

Vi ska nu se hur vi räknar med komponenter.

Sats 1.5 (Komponentvis räkning)

$$\boldsymbol{u} + \boldsymbol{v} = (u_x, u_y, u_z) + (v_x, v_y, v_z) = (u_x + v_x, u_y + v_y, u_z + v_z) \tag{1.63}$$

$$t\boldsymbol{u} = t(u_x, u_y, u_z) = (tu_x, tu_y, tu_z) \tag{1.64}$$

$$\boldsymbol{u} \cdot \boldsymbol{v} = (u_x, u_y, u_z) \cdot (v_x, v_y, v_z) = u_x v_x + u_y v_y + u_z v_z \tag{1.65}$$

$$|\boldsymbol{u}| = |(u_x, u_y, u_z)| = \sqrt{u_x^2 + u_y^2 + u_z^2} \tag{1.66}$$

Bevis. Bevisen av (1.63) och (1.64) lämnas som övning åt läsaren, se problem 1.9. För (1.65) använder vi räkneregler för skalärprodukt och (1.53):

$$\begin{aligned}
\boldsymbol{u} \cdot \boldsymbol{v} &= (u_x \boldsymbol{e}_x + u_y \boldsymbol{e}_y + u_z \boldsymbol{e}_z) \cdot (v_x \boldsymbol{e}_x + v_y \boldsymbol{e}_y + v_z \boldsymbol{e}_z) && (1.67)\\
&= (u_x \boldsymbol{e}_x) \cdot (v_x \boldsymbol{e}_x) + (u_x \boldsymbol{e}_x) \cdot (v_y \boldsymbol{e}_y) + (u_x \boldsymbol{e}_x) \cdot (v_z \boldsymbol{e}_z) + \dots && (1.68)\\
&= u_x v_x (\boldsymbol{e}_x \cdot \boldsymbol{e}_x) + u_x v_y (\boldsymbol{e}_x \cdot \boldsymbol{e}_y) + u_x v_z (\boldsymbol{e}_x \cdot \boldsymbol{e}_z) + \dots && (1.69)\\
&= u_x v_x + u_y v_y + u_z v_z && (1.70)
\end{aligned}$$

För (1.66) använder vi $|\boldsymbol{u}| = \sqrt{\boldsymbol{u} \cdot \boldsymbol{u}}$ och tar $\boldsymbol{v} = \boldsymbol{u}$ i (1.65). □

Exempel 1.6 (Räkning med komponenter) Låt $\boldsymbol{u} = (1, 2, 3)$ och $\boldsymbol{v} = (-4, 5, 6)$. Vi har då

$$|\boldsymbol{u}| = \sqrt{1+4+9} = \sqrt{14}, \tag{1.71}$$
$$|\boldsymbol{v}| = \sqrt{16+25+36} = \sqrt{77} \tag{1.72}$$
$$\boldsymbol{u} \cdot \boldsymbol{v} = -4+10+18 = 24 \tag{1.73}$$

Vinkeln ges då av sambandet $\boldsymbol{u} \cdot \boldsymbol{v} = |\boldsymbol{u}||\boldsymbol{v}|\cos(\theta)$:

$$\cos(\theta) = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{|\boldsymbol{u}||\boldsymbol{v}|} = \frac{24}{\sqrt{14 \cdot 77}}, \quad \theta = \arccos\left(\frac{24}{\sqrt{1078}}\right) \tag{1.74}$$

Den normerade vektorn är

$$\hat{\boldsymbol{u}} = \frac{\boldsymbol{u}}{|\boldsymbol{u}|} = \frac{(1,2,3)}{\sqrt{14}} = \left(\tfrac{1}{\sqrt{14}}, \tfrac{2}{\sqrt{14}}, \tfrac{3}{\sqrt{14}}\right) \tag{1.75}$$

Projektionen av $\boldsymbol{v}$ på $\boldsymbol{u}$ ges av

$$P_{\boldsymbol{u}}(\boldsymbol{v}) = (\boldsymbol{v} \cdot \hat{\boldsymbol{u}})\hat{\boldsymbol{u}} = \left((-4,5,6) \cdot \frac{(1,2,3)}{\sqrt{14}}\right)\frac{(1,2,3)}{\sqrt{14}} \tag{1.76}$$
$$= \frac{24}{14}(1,2,3) = \left(\tfrac{12}{7}, \tfrac{24}{7}, \tfrac{36}{7}\right) \tag{1.77}$$

För att ange läget för en punkt P bildar vi dess *ortsvektor* $\overrightarrow{OP}$. Efter uppdelning i komposanter har vi (se figur 1.13)

$$\overrightarrow{OP} = x\boldsymbol{e}_x + y\boldsymbol{e}_y + z\boldsymbol{e}_z \tag{1.78}$$

där vi kallar komponenterna x, y, z. Dessa kallas *koordinater* för punkten P och vi skriver kortfattat

$$P = (x, y, z) \tag{1.79}$$

Koordinaterna för punkten P är detsamma som komponenterna för ortsvektorn $\overrightarrow{OP}$. Vi betecknar ofta ortsvektorn med $\boldsymbol{r}$, dvs

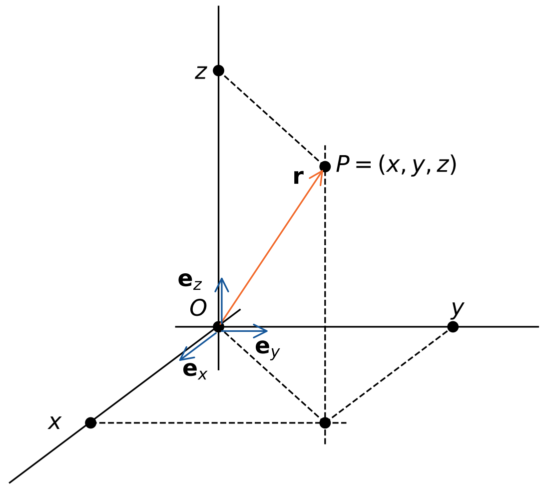
$$\boldsymbol{r} = \overrightarrow{OP} = (x, y, z) \tag{1.80}$$

Exempel 1.7 (Räkning med koordinater) Låt $P_1 = (x_1, y_1, z_1)$ och $P_2 = (x_2, y_2, z_2)$ vara två punkter. Då gäller

$$\overrightarrow{P_1P_2} = (x_2 - x_1, y_2 - y_1, z_2 - z_1) \tag{1.81}$$

och avståndet mellan P_1 och P_2 ges av

$$d = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2} \tag{1.82}$$



Figur 1.13: Ortsvektor och koordinater för en punkt.

Bevis: Genom att gå via origo får vi

$$\overrightarrow{P_1P_2} = \overrightarrow{P_1O} + \overrightarrow{OP_2} = \overrightarrow{OP_2} - \overrightarrow{OP_1} \tag{1.83}$$

$$= (x_2, y_2, z_2) - (x_1, y_1, z_1) = (x_2 - x_1, y_2 - y_1, z_2 - z_1) \tag{1.84}$$

Avståndsformeln följer nu direkt av $d = |\overrightarrow{P_1P_2}|$ och (1.66).

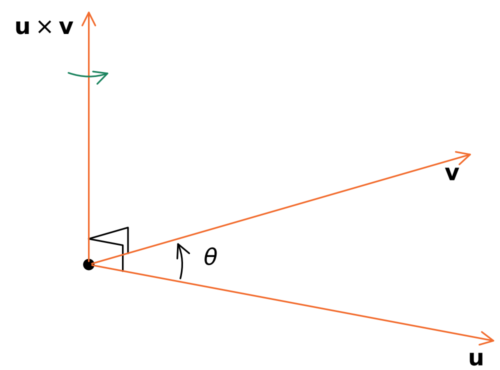
1.5 Kryssprodukt

Definition 1.8 (Kryssprodukt) Kryssprodukten $\boldsymbol{u} \times \boldsymbol{v}$ är den vektor som definieras entydigt av sin längd och riktning enligt följande (se figur 1.14):

- Längden av $\boldsymbol{u} \times \boldsymbol{v}$ är
$$|\boldsymbol{u} \times \boldsymbol{v}| = |\boldsymbol{u}||\boldsymbol{v}| \sin(\theta) \tag{1.85}$$
där θ är vinkeln mellan $\boldsymbol{u}$ och $\boldsymbol{v}$.

- Vektorn $\boldsymbol{u} \times \boldsymbol{v}$ är ortogonal mot både $\boldsymbol{u}$ och $\boldsymbol{v}$ och riktad i den riktning som en högergängad skruv rör sig då den vrids kortaste vägen från $\boldsymbol{u}$ till $\boldsymbol{v}$.

Till skillnad från skalärprodukten $\boldsymbol{u} \cdot \boldsymbol{v}$, som är en skalär, är kryssprodukten $\boldsymbol{u} \times \boldsymbol{v}$ en vektor. Den kallas därför också *vektoriell produkt*. Vi erinrar oss att $\boldsymbol{u} \cdot \boldsymbol{v} = 0$ om $\boldsymbol{u}$ och $\boldsymbol{v}$ är



Figur 1.14: Kryssprodukt.

ortogonala, ty då är $\cos(\theta) = 0$. På liknande sätt ser vi från (1.85) att om $\boldsymbol{u}$ och $\boldsymbol{v}$ är kolinjära (parallella eller antiparallella), så är $\boldsymbol{u} \times \boldsymbol{v} = \mathbf{0}$, ty då är $\sin(\theta) = 0$. Till skillnad från $\cos(\theta)$, som byter tecken, är $\sin(\theta) \geqslant 0$ för $\theta \in [0, \pi]$.

Vi noterar också att vi behöver alla tre dimensionerna för att definiera kryssprodukten. Om, till exempel, både $\boldsymbol{u}$ och $\boldsymbol{v}$ ligger i xy-planet, så blir $\boldsymbol{u} \times \boldsymbol{v}$ kolinjär med z-axeln, dvs den ligger utanför xy-planet.

Sats 1.6 (Räkneregler för kryssprodukt)

$$\boldsymbol{v} \times \boldsymbol{u} = -(\boldsymbol{u} \times \boldsymbol{v}) \quad \text{(antikommutativitet)} \tag{1.86}$$

$$\boldsymbol{u} \times \boldsymbol{u} = \mathbf{0} \tag{1.87}$$

$$(t\boldsymbol{u}) \times \boldsymbol{v} = t(\boldsymbol{u} \times \boldsymbol{v}) = \boldsymbol{u} \times (t\boldsymbol{v}) \quad \text{(distributiv lag)} \tag{1.88}$$

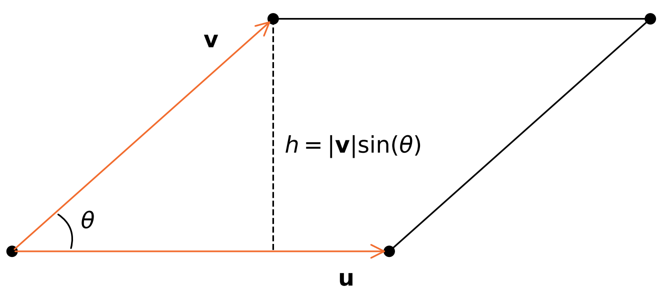
$$\boldsymbol{u} \times (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \times \boldsymbol{v} + \boldsymbol{u} \times \boldsymbol{w} \quad \text{(distributiv lag)} \tag{1.89}$$

Bevis. Reglerna (1.86), (1.87), (1.88) följer lätt från definition 1.8 och lämnas som problem 1.11. Beviset av distributiva lagen (1.89) är krångligare och vi utelämnar det. □

Sambandet (1.86) betyder att kommutativa lagen inte gäller för kryssprodukten, istället är den *antikommutativ*. Inte heller associativa lagen gäller, dvs i allmänhet har vi

$$(\boldsymbol{u} \times \boldsymbol{v}) \times \boldsymbol{w} \neq \boldsymbol{u} \times (\boldsymbol{v} \times \boldsymbol{w}) \tag{1.90}$$

se problem 1.12.



Figur 1.15: Area av parallellogram.

Att $\cos(\theta)$ och $\sin(\theta)$ ingår i definitionerna av skalärprodukt respektive kryssprodukt gör att dessa kan användas i vissa trigonometriska beräkningar. Vi visar detta i två exempel som handlar om area och volym.

Exempel 1.8 (Area av parallellogram) Antag att vektorerna $\boldsymbol{u} \neq \mathbf{0}$, $\boldsymbol{v} \neq \mathbf{0}$. Då spänner de upp en parallellogram vars area ges av basen $|\boldsymbol{u}|$ gånger höjden $|\boldsymbol{v}|\sin(\theta)$, dvs $|\boldsymbol{u}||\boldsymbol{v}|\sin(\theta)$ (se figur 1.15). Arean ges alltså av formeln

$$A = |\boldsymbol{u} \times \boldsymbol{v}| \tag{1.91}$$

dvs längden av kryssprodukten. Vektorn $\boldsymbol{u} \times \boldsymbol{v}$ är ortogonal mot parallellogrammen; den kallas därför *normalvektor*. Även den motriktade vektorn $-(\boldsymbol{u} \times \boldsymbol{v}) = \boldsymbol{v} \times \boldsymbol{u}$ är en normalvektor.

Exempel 1.9 (Volym av parallellepiped) Antag att vektorerna $\boldsymbol{u} \neq \mathbf{0}$, $\boldsymbol{v} \neq \mathbf{0}$, $\boldsymbol{w} \neq \mathbf{0}$. Då spänner de upp en parallellepiped vars volym ges av basytan gånger höjden, $V = hA$. Vi låter A vara arean av parallellogrammen som spänns upp av $\boldsymbol{v}$ och $\boldsymbol{w}$, dvs

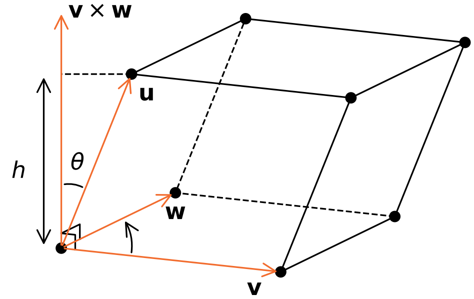
$$A = |\boldsymbol{v} \times \boldsymbol{w}| \tag{1.92}$$

Vektorn $\boldsymbol{v} \times \boldsymbol{w}$ är ortogonal mot basytan och bildar en vinkel θ med vektorn $\boldsymbol{u}$. Om vinkeln är spetsig, dvs $\cos(\theta) \geqslant 0$ (se figur 1.16), så är höjden den skalära projektionen av $\boldsymbol{u}$ på $\boldsymbol{v} \times \boldsymbol{w}$,

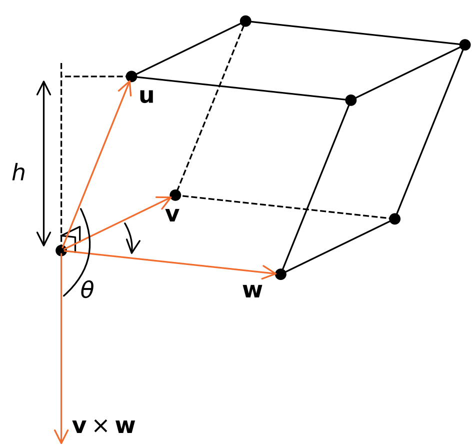
$$h = |\boldsymbol{u}|\cos(\theta) \tag{1.93}$$

och volymen blir

$$V = |\boldsymbol{u}||\boldsymbol{v} \times \boldsymbol{w}|\cos(\theta) = \boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) \tag{1.94}$$



Figur 1.16: Volym av parallellepiped. Här är $\cos(\theta) \geqslant 0$ och $h = |\boldsymbol{u}|\cos(\theta)$.



Figur 1.17: Volym av parallellepiped. Här är $\cos(\theta) \leqslant 0$ och $h = -|\boldsymbol{u}|\cos(\theta)$.

Om vinkeln är trubbig, dvs $\cos(\theta) \leqslant 0$ (se figur 1.17), så blir höjden $h = -|\boldsymbol{u}| \cos(\theta)$ och volymen

$$V = -|\boldsymbol{u}||\boldsymbol{v} \times \boldsymbol{w}| \cos(\theta) = -\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) \tag{1.95}$$

Vi drar slutsatsen att volymen i båda fallen ges av absolutbeloppet av $\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})$,

$$V = |\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})| \tag{1.96}$$

Anledningen till att absolutbelopp behövs är alltså att vi kan råka ta vektorerna i "fel ordning" så att $\boldsymbol{u}$ och $\boldsymbol{v} \times \boldsymbol{w}$ bildar trubbig vinkel. Eftersom volymen inte är negativ, rättar vi till felet genom att ta absolutbelopp.

Produkten $\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})$ kallas *skalär trippelprodukt* och har intressanta egenskaper men vi följer inte upp detta här.

Det är nödvändigt att kunna dela upp kryssprodukten i komposanter.

Sats 1.7 (Komposantuppdelning av kryssprodukten) Antag att $\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z$ är basvektorer i ett Cartesiskt högersystem. Då gäller

$$\boldsymbol{u} \times \boldsymbol{v} = (u_y v_z - u_z v_y)\boldsymbol{e}_x + (u_z v_x - u_x v_z)\boldsymbol{e}_y + (u_x v_y - u_y v_x)\boldsymbol{e}_z \tag{1.97}$$

Bevis. Med hjälp av distributiva lagen och sambanden $\boldsymbol{v} \times \boldsymbol{u} = -(\boldsymbol{u} \times \boldsymbol{v})$, $\boldsymbol{u} \times \boldsymbol{u} = \boldsymbol{0}$ från sats 1.6 får vi

$$\begin{aligned}
\boldsymbol{u} \times \boldsymbol{v} &= (u_x\boldsymbol{e}_x + u_y\boldsymbol{e}_y + u_z\boldsymbol{e}_z) \times (v_x\boldsymbol{e}_x + v_y\boldsymbol{e}_y + v_z\boldsymbol{e}_z) && (1.98)\\
&= u_xv_x(\boldsymbol{e}_x \times \boldsymbol{e}_x) + u_xv_y(\boldsymbol{e}_x \times \boldsymbol{e}_y) + u_xv_z(\boldsymbol{e}_x \times \boldsymbol{e}_z) && (1.99)\\
&\quad + u_yv_x(\boldsymbol{e}_y \times \boldsymbol{e}_x) + u_yv_y(\boldsymbol{e}_y \times \boldsymbol{e}_y) + u_yv_z(\boldsymbol{e}_y \times \boldsymbol{e}_z) && (1.100)\\
&\quad + u_zv_x(\boldsymbol{e}_z \times \boldsymbol{e}_x) + u_zv_y(\boldsymbol{e}_z \times \boldsymbol{e}_y) + u_zv_z(\boldsymbol{e}_z \times \boldsymbol{e}_z) && (1.101)\\
&= (u_yv_z - u_zv_y)(\boldsymbol{e}_y \times \boldsymbol{e}_z) + (u_zv_x - u_xv_z)(\boldsymbol{e}_z \times \boldsymbol{e}_x) && (1.102)\\
&\quad + (u_xv_y - u_yv_x)(\boldsymbol{e}_x \times \boldsymbol{e}_y) && (1.103)
\end{aligned}$$

Eftersom vi har ett högersystem, så gäller

$$\boldsymbol{e}_y \times \boldsymbol{e}_z = \boldsymbol{e}_x, \ \boldsymbol{e}_z \times \boldsymbol{e}_x = \boldsymbol{e}_y, \ \boldsymbol{e}_x \times \boldsymbol{e}_y = \boldsymbol{e}_z \tag{1.104}$$

och beviset är klart. □

För att lättare komma ihåg sambandet (1.97) inför vi ett räkneschema som kallas *determinant*. En *tvåradig determinant* tecknas $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ och definieras enligt

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \tag{1.105}$$

En *treradig determinant* tecknas och definieras

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & j \end{vmatrix} = a \begin{vmatrix} e & f \\ h & j \end{vmatrix} - b \begin{vmatrix} d & f \\ g & j \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix} \tag{1.106}$$

$$= a(ej - fh) - b(dj - fg) + c(dh - eg) \tag{1.107}$$

Notera hur den treradiga determinanten utvecklas i en summa av tre tvåradiga underdeterminanter multiplicerade med talen från första raden med alternerande tecken. Underdeterminanterna fås genom att stryka den rad och den kolumn som möts i respektive tal. Strecken som används i determinantens beteckning kallas determinantstreck och ska inte förväxlas med absolutbelopp eller längd. Vi ska studera determinanter mer utförligt i sektion 3.3.

Om vi sätter in basvektorerna i första raden och komponenterna för $\boldsymbol{u}$ och $\boldsymbol{v}$ i andra respektive tredje raden, så får vi

$$\begin{vmatrix} \boldsymbol{e}_x & \boldsymbol{e}_y & \boldsymbol{e}_z \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix} = \begin{vmatrix} u_y & u_z \\ v_y & v_z \end{vmatrix} \boldsymbol{e}_x - \begin{vmatrix} u_x & u_z \\ v_x & v_z \end{vmatrix} \boldsymbol{e}_y + \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix} \boldsymbol{e}_z \tag{1.108}$$

$$= (u_y v_z - u_z v_y)\boldsymbol{e}_x - (u_x v_z - u_z v_x)\boldsymbol{e}_y + (u_x v_y - u_y v_x)\boldsymbol{e}_z \tag{1.109}$$

Detta är detsamma som (1.97). Alltså:

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{vmatrix} \boldsymbol{e}_x & \boldsymbol{e}_y & \boldsymbol{e}_z \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix} \tag{1.110}$$

Exempel 1.10 (Area av triangel) De tre punkterna $O = (0,0,0)$, $A = (1,1,1)$, $B = (1,2,3)$ är hörn i en triangel. Vi beräknar dess area.

Triangeln spänns upp av vektorerna $\boldsymbol{u} = \overrightarrow{OA} = (1,1,1)$, $\boldsymbol{v} = \overrightarrow{OB} = (1,2,3)$. Dess area är hälften av arean av motsvarande parallellogram. Enligt (1.91) har vi då

$$A = \frac{1}{2}|\boldsymbol{u} \times \boldsymbol{v}| \tag{1.111}$$

och enlig (1.110)

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{vmatrix} \boldsymbol{e}_x & \boldsymbol{e}_y & \boldsymbol{e}_z \\ 1 & 1 & 1 \\ 1 & 2 & 3 \end{vmatrix} = (1, -2, 1) \tag{1.112}$$

Arean är $A = \frac{1}{2}\sqrt{1^2 + (-2)^2 + 1^2} = \frac{1}{2}\sqrt{6}$.

Exempel 1.11 (Volym av tetraeder) De fyra punkterna är $O=(0,0,0)$, $A=(3,-1,2)$, $B=(4,0,5)$, $C=(1,3,2)$ är hörn i en tetraeder. Vi beräknar dess volym.

Tetraedern spänns upp av vektorerna $\boldsymbol{u}=\overrightarrow{OA}=(3,-1,2)$, $\boldsymbol{v}=\overrightarrow{OB}=(4,0,5)$, $\boldsymbol{w}=\overrightarrow{OC}=(1,3,2)$. Dess volym är en sjättedel av volymen av motsvarande parallellepiped, se även problem 1.13. Enligt (1.96) har vi då

$$V=\frac{1}{6}|\boldsymbol{u}\cdot(\boldsymbol{v}\times\boldsymbol{w})| \tag{1.113}$$

och enlig (1.110) (se även problem 1.14)

$$\boldsymbol{v}\times\boldsymbol{w}=\begin{vmatrix}\boldsymbol{e}_x & \boldsymbol{e}_y & \boldsymbol{e}_z\\ 4 & 0 & 5\\ 1 & 3 & 2\end{vmatrix}=(-15,-3,12) \tag{1.114}$$

$$\boldsymbol{u}\cdot(\boldsymbol{v}\times\boldsymbol{w})=(3,-1,2)\cdot(-15,-3,12)=-18 \tag{1.115}$$

Volymen är $V=\frac{1}{6}|-18|=3$.

1.6 Räta linjen och planet

Vi ska nu bestämma ekvationer för räta linjen och planet. Vi använder ett Cartesiskt högersystem (O,x,y,z).

En *rät linje* L i rummet bestäms entydigt av en punkt $P_0=(x_0,y_0,z_0)$ på linjen och en *riktningsvektor* $\boldsymbol{v}=(v_x,v_y,v_z)$ som pekar ut linjens riktning (se figur 1.18). (Obs att $\boldsymbol{v}\neq\boldsymbol{0}$ annars har vi ingen riktning.) En punkt $P=(x,y,z)$ ligger på linjen om och endast om vektorn $\overrightarrow{P_0P}$ är kolinjär (parallell eller antiparallell) med $\boldsymbol{v}$, dvs om det finns ett tal t så att $\overrightarrow{P_0P}=t\boldsymbol{v}$. Men $\overrightarrow{P_0P}=\boldsymbol{r}-\boldsymbol{r}_0$, där $\boldsymbol{r}$ och $\boldsymbol{r}_0$ är punkternas ortsvektorer. Alltså har vi $\boldsymbol{r}-\boldsymbol{r}_0=t\boldsymbol{v}$, dvs

$$\boldsymbol{r}=\boldsymbol{r}_0+t\boldsymbol{v},\quad t\in\mathbb{R} \tag{1.116}$$

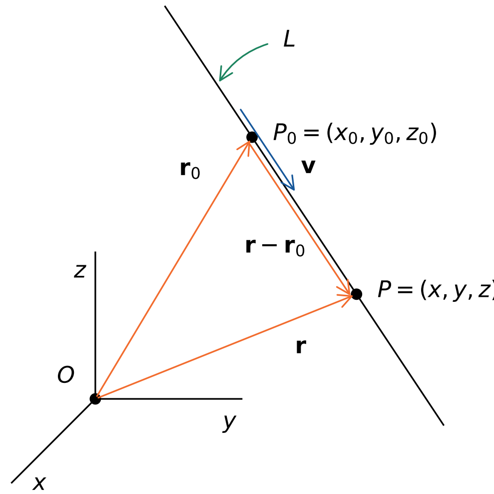
eller med koordinater

$$\begin{cases}x=x_0+tv_x,\\ y=y_0+tv_y, \qquad t\in\mathbb{R}\\ z=z_0+tv_z,\end{cases} \tag{1.117}$$

Ekvationerna (1.116) och (1.117) kallas *räta linjens ekvationer på parameterform* skrivna på vektorform respektive koordinatform. När parametern t genomlöper intervallet $(-\infty,\infty)$ genomlöper punkten P den räta linjen.

Genom att eliminera parametern t ur (1.117) får vi *räta linjens ekvationer på parameterfri form*

$$\frac{x-x_0}{v_x}=\frac{y-y_0}{v_y}=\frac{z-z_0}{v_z}\quad\bigl(=t\bigr) \tag{1.118}$$



Figur 1.18: En rät linje L bestäms av en punkt P_0 och en vektor $\boldsymbol{v}$.

förutsatt att alla komponenter v_x, v_y, v_z är skilda från 0. Om en av dem är 0, till exempel $v_z = 0$, så får vi istället

$$\frac{x - x_0}{v_x} = \frac{y - y_0}{v_y}, \; z = z_0 \tag{1.119}$$

Exempel 1.12 (Räta linjen genom två punkter) Vi bestämmer ekvationer för den räta linjen som går genom punkterna $A = (1, 2, 3)$ och $B = (3, 2, 1)$.

Vi väljer $P_0 = A$ och en riktningsvektor $\boldsymbol{v} = \overrightarrow{AB} = (3, 2, 1) - (1, 2, 3) = (2, 0, -2)$. Vi får

$$\begin{cases} x = 1 + 2t, & \\ y = 2, & t \in \mathbb{R} \\ z = 3 - 2t, & \end{cases} \tag{1.120}$$

Efter elimination av t:

$$\frac{x - 1}{2} = \frac{z - 3}{-2}, \; y = 2 \tag{1.121}$$

Obs att den andra punkten $B = (3, 2, 1)$ svarar mot $t = 1$ i (1.120).

Två godtyckliga räta linjer skär normalt inte varandra. Men det kan inträffa och i så fall skär de varandra i en unik punkt, som följande exempel visar.

Exempel 1.13 (Skärning mellan två räta linjer) Vi visar att de räta linjerna

$$x = y - 1 = \frac{z+1}{2} \quad \text{och} \quad \frac{x-3}{2} = \frac{y-1}{-1} = z - 2 \tag{1.122}$$

skär varandra och bestämmer skärningspunkten.

Vi skriver linjerna på parameterform:

$$\begin{cases} x = t \\ y = 1 + t \\ z = -1 + 2t \end{cases} \qquad \begin{cases} x = 3 + 2s \\ y = 1 - s \\ z = 2 + s \end{cases} \tag{1.123}$$

Obs att vi använder olika parametrar s och t. Vi likställer x, y, z-koordinaterna och får ekvationssystemet

$$\begin{cases} t = 3 + 2s \\ 1 + t = 1 - s \\ -1 + 2t = 2 + s \end{cases} \tag{1.124}$$

Vi har 3 ekvationer i 2 obekanta s och t. Det är inte självklart att det finns någon lösning. Vi undersöker. Den andra ekvationen ger $t = -s$, vilket sättes in i den första ekvationen: $-s = 3 + 2s$ och vi får $s = -1$ och sedan $t = -s = 1$. Dessa måste uppfylla även den tredje ekvationen. Vi sätter in: $-1 + 2t = -1 + 2 = 1$ och $2 + s = 2 - 1 = 1$, vilket stämmer. Vi har en unik lösning: $t = 1$, $s = -1$. Med dessa parametervärden får vi $x = 1$, $y = 2$, $z = 1$, dvs skärningspunkten $P = (1, 2, 1)$.

I nästa kapitel ska vi lära oss en systematisk metod för att undersöka lösbarheten för sådana ekvationssystem.

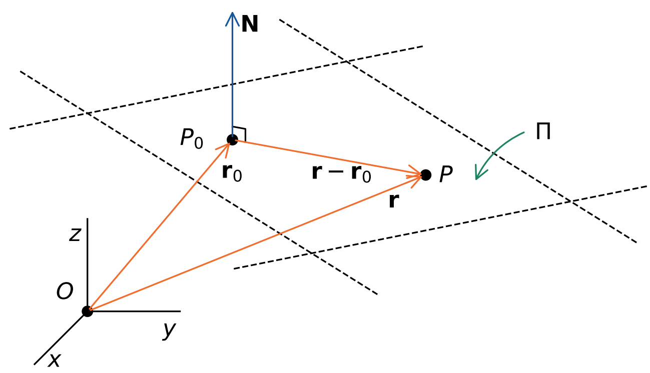
Det är viktigt att känna igen den räta linjens ekvationer (1.116), (1.117) och (1.118). Ekvationerna kan ta lite olika form genom omskrivning. Till exempel är P_0 och $\boldsymbol{v}$ inte unika. I (1.120) och (1.121) identifierar vi genast en punkt på linjen $P_0 = (1, 2, 3)$ och en riktningsvektor $\boldsymbol{v} = (2, 0, -2)$. Men, till exempel, $P_0 = (5, 2, -1)$ och $\boldsymbol{v} = (-4, 0, 4)$ kan också användas och ekvationerna blir då

$$\begin{cases} x = 5 - 4t, \\ y = 2, & t \in \mathbb{R} \\ z = -1 + 4t, \end{cases} \tag{1.125}$$

Vi går nu över till planet. Ett *plan* bestäms entydigt av en punkt $P_0 = (x_0, y_0, z_0)$ i planet och en *normalvektor* $\boldsymbol{N} = (N_x, N_y, N_z)$ som är ortogonal mot planet (se figur 1.19). En punkt $P = (x, y, z)$ är i planet om och endast om vektorn $\boldsymbol{r} - \boldsymbol{r}_0 = \overrightarrow{P_0P}$ är ortogonal mot normalvektorn $\boldsymbol{N}$. Detta ger ekvationen

$$\boldsymbol{N} \cdot (\boldsymbol{r} - \boldsymbol{r}_0) = 0 \tag{1.126}$$

Uttryckt i komponenter blir detta:



Figur 1.19: Ett plan Π bestäms av en punkt P_0 och en normalvektor $\boldsymbol{N}$.

$$N_x(x-x_0)+N_y(y-y_0)+N_z(z-z_0)=0 \tag{1.127}$$

Denna ekvation kan också skrivas på formen

$$Ax+By+Cz=D \tag{1.128}$$

där $A=N_x$, $B=N_y$, $C=N_z$, $D=N_xx_0+N_yy_0+N_zz_0$.

Det är viktigt att kunna känna igen planets ekvation (1.127) eller (1.128). Till exempel, om vi ser en ekvation på formen (1.128), så förstår vi genast att detta är ekvationen för ett plan med normalvektor $\boldsymbol{N}=(A,B,C)$ och som går genom punkten P_0, som vi får genom att välja två av x_0, y_0, z_0 och räkna ut den tredje ur (1.128).

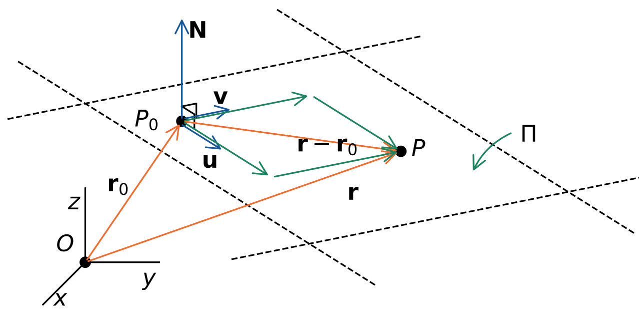
Ett sätt att hitta en normalvektor till planet är att ta två vektorer $\boldsymbol{u}$, $\boldsymbol{v}$ som är parallella med planet och bilda $\boldsymbol{N}=\boldsymbol{u}\times\boldsymbol{v}$. Det är nödvändigt att $\boldsymbol{u}$ och $\boldsymbol{v}$ inte är kolinjära, för då blir ju $\boldsymbol{u}\times\boldsymbol{v}=\boldsymbol{0}$. Sedan kan vi skriva planets ekvation enligt ovan (se figur 1.20).

Ett annat sätt att beskriva planet är att notera att en punkt med ortsvektor $\boldsymbol{r}$ är i planet om och endast om $\boldsymbol{r}-\boldsymbol{r}_0$ är parallell med planet, dvs $\boldsymbol{r}-\boldsymbol{r}_0=s\boldsymbol{u}+t\boldsymbol{v}$ med $\boldsymbol{u}$, $\boldsymbol{v}$ som ovan och med parametrar $s,t\in\mathbb{R}$. Detta är *planets ekvation på parameterform*:

$$\boldsymbol{r}=\boldsymbol{r}_0+s\boldsymbol{u}+t\boldsymbol{v},\quad s,t\in\mathbb{R} \tag{1.129}$$

eller

$$\begin{cases} x=x_0+su_x+tv_x, \\ y=y_0+su_y+tv_y, \quad s,t\in\mathbb{R} \\ z=z_0+su_z+tv_z, \end{cases} \tag{1.130}$$



Figur 1.20: Ett plan Π bestäms av en punkt P_0 och två vektorer $\boldsymbol{u}$ och $\boldsymbol{v}$ som är parallella med planet.

Exempel 1.14 (Plan genom tre punkter) Vi bestämmer en ekvation för planet som går genom punkterna $A = (1,1,1)$, $B = (1,2,3)$, $C = (-1,2,1)$.

Vektorerna

$$\boldsymbol{u} = \overrightarrow{AB} = (1,2,3) - (1,1,1) = (0,1,2) \tag{1.131}$$

$$\boldsymbol{v} = \overrightarrow{AC} = (-1,2,1) - (1,1,1) = (-2,1,0) \tag{1.132}$$

är parallella med planet och vektorn $\boldsymbol{N} = \boldsymbol{u} \times \boldsymbol{v}$ är en normalvektor. Vi beräknar den med determinantformeln (1.110):

$$\boldsymbol{N} = \boldsymbol{u} \times \boldsymbol{v} = \begin{vmatrix} \boldsymbol{e}_x & \boldsymbol{e}_y & \boldsymbol{e}_z \\ 0 & 1 & 2 \\ -2 & 1 & 0 \end{vmatrix} = \cdots = (-2,-4,2) \tag{1.133}$$

Med $P_0 = A = (1,1,1)$ kan vi nu skriva ned planets ekvation

$$-2(x-1) - 4(y-1) + 2(z-1) = 0 \tag{1.134}$$

Å andra sidan, om vi ser denna ekvation känner vi igen ekvationen för ett plan och vi kan genast avläsa en normalvektor $\boldsymbol{N} = (-2,-4,2)$ och en punkt i planet $P_0 = (1,1,1)$. Eventuellt kan vi vilja förenkla till

$$x + 2y - z = 2 \tag{1.135}$$

Ur denna ekvation kan vi direkt avläsa att $\boldsymbol{N}_1 = (1,2,-1)$ är en normalvektor. Denna är antiparallell med $\boldsymbol{N}$, ty $\boldsymbol{N}_1 = -2\boldsymbol{N}$, och vi inser att planets normalvektor är inte unik. En punkt i planet skulle vi kunna finna genom att sätta in $x = y = 0$ i (1.135) och räkna ut $z = -2$. Alltså är $P_1 = (0, 0, -2)$ en punkt i planet (i tillägg till de punkter vi startade med).

Planets ekvation på parameterform, (1.129), är

$$\boldsymbol{r} = (1, 1, 1) + s(0, 1, 2) + t(-2, 1, 0), \quad s, t \in \mathbb{R} \tag{1.136}$$

eller

$$\begin{cases} x = 1 - 2t, \\ y = 1 + s + t, \quad s, t \in \mathbb{R} \\ z = 1 + 2s, \end{cases} \tag{1.137}$$

Obs att punkten $B = (1, 2, 3)$ svarar mot $s = 1$, $t = 0$ och $C = (-1, 2, 1)$ fås för $s = 0, t = 1$.

Övningar

1.1 Geometriska vektorer

Övning 1.1 Illustrera med en figur.
(a) $\overrightarrow{AB} = \overrightarrow{CD}$ (b) $\overrightarrow{AB} \neq \overrightarrow{CD}$ (c) $\overrightarrow{AA}$ (d) $\overrightarrow{BA} = -\overrightarrow{AB}$

1.2 Vektoralgebra

Övning 1.2 Illustrera vektorerna med vektordiagram som i figur 1.2 och figur 1.3.
(a) $\boldsymbol{u} + 2\boldsymbol{v}$ (b) $\boldsymbol{u} + \boldsymbol{v} + \boldsymbol{w}$ (c) $\boldsymbol{u} - \boldsymbol{v}$ (d) $\boldsymbol{u} + \boldsymbol{v} - (\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{0}$

1.3 Skalärprodukt, ortogonalitet och projektion

Övning 1.3 Vektorerna $\boldsymbol{u}$ och $\boldsymbol{v}$ har längderna 2 respektive 3 och bildar vinkeln θ. Beräkna längden av vektorn $\boldsymbol{u} - 5\boldsymbol{v}$.
(a) $\theta = \pi/6$ (b) $\theta = \pi/2$ (c) $\theta = \pi/4$ (d) $\theta = \pi$

Övning 1.4 Vektorerna $\boldsymbol{u}$ och $\boldsymbol{v}$ har längderna 2 respektive 3 och bildar vinkeln θ. Beräkna skalära projektionen och vektorprojektionen av $\boldsymbol{v}$ på $\boldsymbol{u}$.
(a) $\theta = \pi/6$ (b) $\theta = \pi/2$ (c) $\theta = \pi/4$ (d) $\theta = \pi$

Övning 1.5 Vektorerna $\boldsymbol{u}$, $\boldsymbol{v}$ och $\boldsymbol{u}+\boldsymbol{v}$ har längderna 3, 4 respektive 2. Beräkna vinkeln mellan $\boldsymbol{u}$ och $\boldsymbol{v}$.

1.4 Koordinatsystem

Övning 1.6 Låt $\boldsymbol{u} = (-1, 2, 3)$, $\boldsymbol{v} = (2, 1, -2)$ och $\boldsymbol{w} = (4, 5, 6)$. Beräkna följande.
(a) $\boldsymbol{u} + 2\boldsymbol{v}$ (b) $\boldsymbol{u} + \boldsymbol{v} + \boldsymbol{w}$ (c) $\boldsymbol{u} - \boldsymbol{v}$ (d) $\hat{\boldsymbol{v}}$

Övning 1.7 Beräkna vinkeln mellan $\boldsymbol{u}$ och $2\boldsymbol{v} - \boldsymbol{u}$, då $\boldsymbol{u} = (2, -3, 4)$ och $\boldsymbol{v} = (3, 4, 0)$.

Övning 1.8 Bestäm t så att vektorerna $(t, 2t^2, 3t)$ och $(-1, 1, t)$ blir ortogonala.

Övning 1.9 Normera vektorn.
(a) $(1, 1, 1)$ (b) $(0, 0, 0)$ (c) $(-1, 1, -1)$ (d) $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$

Övning 1.10 Låt $\boldsymbol{u} = (-1, 2, 3)$ och $\boldsymbol{v} = (2, 1, -2)$.
(a) Beräkna längderna av vektorerna.
(b) Beräkna skalärprodukten mellan vektorerna.
(c) Beräkna vinkeln mellan vektorerna.
(d) Beräkna vektorprojektionen av $\boldsymbol{v}$ på $\boldsymbol{u}$.

Övning 1.11 Dela upp vektorn $\boldsymbol{u}$ i ortogonala komponenter där den ena är parallell/antiparallell med $\boldsymbol{v} = (1, 1, 1)$.
(a) $\boldsymbol{u} = (0, 1, 1)$ (b) $\boldsymbol{u} = (1, 1, 0)$

Övning 1.12 Beräkna vinkeln vid hörnet A i triangeln med hörnen $A = (2, -1, 3)$, $B = (3, 1, 1)$, $C = (4, 3, 2)$.

Övning 1.13 Avgör om vinkeln mellan vektorerna är rät, trubbig eller spetsig.
(a) $\boldsymbol{u} = (2, -1, 3)$, $\boldsymbol{v} = (3, 4, 1)$
(b) $\boldsymbol{u} = (2, 4, 1)$, $\boldsymbol{v} = (-3, 2, -2)$
(c) $\boldsymbol{u} = (2, 0, 3)$, $\boldsymbol{v} = (-2, 5, 1)$
(d) $\boldsymbol{u} = (1, t, 2)$, $\boldsymbol{v} = (-4t, 2, 3)$

1.5 Kryssprodukt

Övning 1.14 Antag att $|\boldsymbol{u}| = 3$, $|\boldsymbol{v}| = 4$ och att vinkeln mellan vektorerna är θ. Beräkna längden av $\boldsymbol{u} \times \boldsymbol{v}$.
(a) $\theta = 0$ (b) $\theta = \pi/6$ (c) $\theta = \pi/2$ (d) $\theta = \pi$

Övning 1.15 Beräkna arean av triangeln med hörnen i punkterna A, B, C.
(a) $A = (1, 3, 2)$, $B = (0, -1, 1)$, $C = (2, -1, 2)$
(b) $A = (1, 2, 3)$, $B = (2, -3, 1)$, $C = (3, 1, 2)$

Övning 1.16 Beräkna volymen av tetraedern med hörnen $A = (2, 2, 1)$, $B = (4, 2, 1)$, $C = (3, 5, 1)$, $D = (3, 3, 2)$. Tips: problem 1.13.

Övning 1.17 Bestäm alla enhetsvektorer $\hat{\boldsymbol{N}}$ som är ortogonala mot $\boldsymbol{u}$ och $\boldsymbol{v}$.
(a) $\boldsymbol{u} = (1, -3, 2)$, $\boldsymbol{v} = (2, 1, -3)$ (b) $\boldsymbol{u} = (2, 1, 3)$, $\boldsymbol{v} = (-1, 2, 1)$

Övning 1.18 En kraft $\boldsymbol{F}$ angriper i punkten P. Kraftens *vridmoment* kring punkten A är $\boldsymbol{M} = \overrightarrow{AP} \times \boldsymbol{F}$. Beräkna vridmomentet.
$\boldsymbol{F} = (4, 1, -3)$ $[\mathtt{N}]$, $P = (2, -3, 1)$, $A = (0, 1, 2)$ $[\mathtt{m}]$

Övning 1.19 Bestäm ekvationen för den räta linjen som går genom punkterna A och B. Skriv ekvationen dels på parameterform (1.116) och dels på parameterfri form (1.117).
(a) $A = (1, 3, 2)$, $B = (0, -1, 1)$ (b) $A = (1, 2, 3)$, $B = (2, -3, 1)$
(c) $A = (2, 2, 3)$, $B = (2, -3, 1)$

Övning 1.20 Bestäm de punkter på räta linjen i övning 1.19 som ligger på avståndet 1 från punkten A.

Övning 1.21 Skriv z-axelns ekvation på parameterform.

Övning 1.22 Ange en riktningsvektor och en punkt på den räta linjen.
(a) $\dfrac{x-1}{3} = y - 2 = \dfrac{z+3}{-2}$
(b) $\begin{cases} x = 2, \\ y = 2 - 5t, & -\infty < t < \infty \\ z = -3 + 4t, \end{cases}$
(c) $2x = y + 1 = 3 - 3z$

Övning 1.23 Beräkna avståndet mellan punkten $(1, 3, 2)$ och linjen $x - 1 = \frac{y-5}{2} = \frac{z-2}{3}$. Tips: problem 1.15.

Övning 1.24 Bestäm en punkt i planet och en normalvektor till planet.
(a) $5(x + 1) - 2(y - 3) + z = 0$ (b) $x + 2y + 3z = 6$
(c) $x/3 - y/4 + z = 0$ (d) $11x - 4y - z = 6$

Övning 1.25 Bestäm ekvationen för planet genom punkterna $(1, 1, 1)$, $(2, 3, 4)$, $(0, -2, 2)$.

Övning 1.26 Finn skärningspunkten mellan planet $x - y + 2z = 4$ och den räta linjen genom punkterna $(2, 1, 3)$ och $(4, 5, 7)$.

Problem

1.1 Geometriska vektorer

Problem 1.1 Hitta fler exempel på fysiska vektorer som i exempel 1.1.

1.2 Vektoralgebra

Problem 1.2 Bevisa sats 1.2

Problem 1.3 Visa att diagonalerna i en parallellogram skär varandra mitt itu.

Problem 1.4 En median i en triangel är en sträcka som förbinder ett hörn med mittpunkten på motstående sida. Visa att medianerna skär varandra i en punkt som delar medianerna i förhållandet 1:2. Punkten kallas triangelns *tyngdpunkt*.

1.3 Skalärprodukt, ortogonalitet och projektion

Problem 1.5 Bevisa Pythagoras sats med hjälp av vektorer.

Problem 1.6 Bevisa att $\boldsymbol{u} \cdot \boldsymbol{v} = \frac{1}{4}(|\boldsymbol{u}+\boldsymbol{v}|^2 - |\boldsymbol{u}-\boldsymbol{v}|^2)$.

Problem 1.7 Antag att $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u} \cdot \boldsymbol{w}$. Följer det att $\boldsymbol{v} = \boldsymbol{w}$? Ge bevis eller motexempel.

Problem 1.8 Visa med vektorer att sidorna i en liksidig triangel bildar vinkeln $\pi/3$.

1.4 Koordinatsystem

Problem 1.9 Bevisa (1.63) och (1.64).

Problem 1.10 Kraften $\boldsymbol{F} = (3, -4, 2)$ N verkar på en kropp som rör sig från punkten $A = (-1, 3, 2)$ m till punkten B. Hur mycket ändras kroppens rörelseenergi?
(a) $B = (1, 4, 5)$ (b) $B = (-3, 2, 3)$ (c) $B = (0, 5, 3)$

1.5 Kryssprodukt

Problem 1.11 Bevisa (1.86), (1.87), (1.88) i sats 1.6.

Problem 1.12 Bevisa (1.90). Tips: låt $\boldsymbol{u} \neq \boldsymbol{0}$, $\boldsymbol{v} \neq \boldsymbol{0}$ och låt $\boldsymbol{w} \neq \boldsymbol{0}$ vara ortogonal mot $\boldsymbol{u}$.

Problem 1.13 Bevisa formeln $V = \frac{1}{6}|\boldsymbol{u}\cdot(\boldsymbol{v}\times\boldsymbol{w})|$ för volymen av tetraedern som spänns upp av vektorerna $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$.

Problem 1.14 Visa att

$$\boldsymbol{u}\cdot(\boldsymbol{v}\times\boldsymbol{w}) = \begin{vmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix} \tag{1.138}$$

Tips: använd (1.110).

1.6 Räta linjen och planet

Problem 1.15 Visa att avståndet från punkten P till den räta linjen genom punkten P_0 med riktningsvektorn $\boldsymbol{v}$ är

$$d = \frac{|\overrightarrow{P_0P}\times\boldsymbol{v}|}{|\boldsymbol{v}|} \tag{1.139}$$

Problem 1.16 Visa att de räta linjerna $x = y - 1 = (z+1)/2$ och $\frac{x-3}{2} = \frac{y-1}{-1} = \frac{z-2}{1}$ skär varandra och beräkna vinkeln mellan dem.

Problem 1.17 En ljusstråle med riktningsvektorn $(1, 2, 2)$ reflekteras i planet $3x + 4y + z = 0$. Bestäm en riktningsvektor för den reflekterade strålen.

Datorövningar

Datorövning 1.1 I MATLAB skrivs vektorer antingen som radmatriser `v=[1,2,3]` eller som kolonnmatriser `v=[1;2;3]`. Kolonnformen rekommenderas med tanke på hur vi kommer att skriva senare i boken. Gör övning 1.6 och 1.10 med MATLAB. Tips: I MATLAB beräknar vi lätt skalärprodukt med `u'*v` om vektorerna är på kolonnform. Prova också att använda funktionen `dot()`.

Datorövning 1.2 Skriv en funktion som beräknar skalärprodukt utan att använda `dot()`. Den ska ha deklarationen `function s=skalarprodukt(u,v)`.

Datorövning 1.3 Skriv en funktion som beräknar den skalära projektionen av en vektor `v` längs vektorn `u`. Den ska ha deklarationen `function s=sproj(u,v)`.

Datorövning 1.4 Skriv en funktion som beräknar vektorprojektionen av en vektor `v` längs vektorn `u`. Den ska ha deklarationen `function w=proj(u,v)`.

Datorövning 1.5 Skriv en funktion som beräknar kryssprodukten utan att använda `cross`. Den ska ha deklarationen `function w=kryss(u,v)`. Testa med några exempel och jämför med `cross()`.

Datorövning 1.6 Skriv en funktion som testar om två vektorer är ortogonala. Den ska ha deklarationen `function a=ortotest(u,v)`, där a är en boolesk variabel `a=0` eller `a=1` om falskt eller sant. Tips: skalärprodukt och $\cos(\theta) = 0$ om vektorerna är ortogonala. Testa med några exempel.

Datorövning 1.7 Skriv en funktion som testar om två vektorer är parallella eller antiparallella. Den ska ha deklarationen `function a=paratest(u,v)`, där a är en boolesk variabel `a=0` eller `a=1` om falskt eller sant. Tips: kryssprodukt och $\sin(\theta) = 0$ om vektorerna är parallella eller antiparallella. Testa med några exempel.

Datorövning 1.8 Låt `v=[1;-2;4]`. Vad blir `length(v)`, `size(v)`, `abs(v)`, `norm(v)` och `sqrt(v'*v)`? Vad gör funktionerna `length()`, `size()`, `abs()` respektive `norm()`?

Datorövning 1.9 Ortogonal uppdelning. Skriv en funktion som delar upp vektorn `v` som `v=v1+v2`, där `v2` och `v2` är ortogonala och `v1` är parallell med vektorn `u`. Funktionen ska ha deklarationen `function [v1,v2]=ortodekomposition(u,v)`. Testa med övning 1.11.

Datorövning 1.10 Area av triangel. Skriv en funktion som beräknar arean av triangeln med hörnen i punkterna A, B, C. Deklaration `function a=triangelarea(A,B,C)`. Testa med övning 1.15.

2. Linjära ekvationssystem

Vi löser linjära ekvationssystem med Gauss eliminationsmetod.

2.1 Gauss eliminationsmetod

En *linjär ekvation* i n variabler $x_1, x_2, \dots, x_n$ har formen

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{2.1}$$

Talen $a_1, a_2, \dots, a_n, b$ kallas koefficienter och är givna. Vänsterledet är en linjär kombination av variablerna; därav beteckningen linjär ekvation. Ett *linjärt ekvationssystem* är en uppsättning av m linjära ekvationer:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{2.2}$$

Exempel 2.1 (Linjärt ekvationssystem) I detta exempel har vi $m = 2$, $n = 3$.

$$\begin{cases} -3x_1 + x_2 - \frac{5}{2}x_3 = 7 \\ x_1 + 8x_3 = 4 \end{cases} \tag{2.3}$$

Vårt mål är att lösa linjära ekvationssystem, dvs att finna värden på variablerna $x_1, \dots, x_n$ som uppfyller alla ekvationerna samtidigt. Med två eller tre variabler kan vi resonera geometriskt som i följande exempel.

Exempel 2.2 (Geometrisk tolkning av linjärt ekvationssystem) Vi erinrar oss från avsnitt 1.6 att

$$ax + by + cz = d \tag{2.4}$$

är ekvationen för ett plan i rummet. Att lösa ett system av sådana ekvationer innebär att söka skärningen mellan flera plan. Till exempel, två plan kan skära varandra längs en rät linje (oändligt många lösningar) eller inte alls om de är parallella (ingen lösning). Om tre plan skär varandra så är skärningen en rät linje (oändligt många lösningar) eller en enda punkt (unik lösning).

En linjär ekvation i två variabler,

$$ax + by = c \tag{2.5}$$

är ekvationen för en rät linje i planet. Ekvationssystemet

$$\begin{cases} 1x - 1y = 0 \\ 0x + 1y = 1 \end{cases} \tag{2.6}$$

betyder de två räta linjerna $y = x$ och $y = 1$, som skär varandra i punkten $(1, 1)$, dvs vi har en unik lösning $x = 1$, $y = 1$. På liknande sätt ser vi att ekvationssystemet

$$\begin{cases} x - y = 0 \\ x - y = -1 \end{cases} \tag{2.7}$$

betyder de parallella linjerna $y = x$ och $y = x+1$, som inte skär varandra, dvs systemet har ingen lösning. Ekvationerna

$$\begin{cases} x - y = 0 \\ -x + y = 0 \end{cases} \tag{2.8}$$

beskriver båda samma räta linje $y = x$, dvs oändligt många lösningar nämligen alla x, y med $y = x$.

Dessa exempel visar att ett linjärt ekvationssystem kan ha *ingen lösning*, *unik lösning* eller *oändligt många lösningar*. Vi ska se att detta gäller för linjära ekvationssystem i allmänhet.

Definition 2.1 (Konsistent ekvationssystem) Ett linjärt ekvationssystem (2.2) kallas *konsistent* om det har någon lösning, annars *inkonsistent*. *Lösningsmängden* är mängden av alla lösningar, dvs alla $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ som uppfyller ekvationssystemet.

Vi ska nu presentera en lösningsmetod för linjära ekvationssystem som är *systematisk* (dvs en algoritm) och som transformerar systemet till en standardform, *reducerad trappstegsform*, som avslöjar lösningsmängden. Vi beskriver metoden först genom ett räkneexempel.

Exempel 2.3 (Gauss eliminationsmetod) Vi löser det linjära ekvationssystemet

$$
\left\{\begin{array}{rrrcr}
 & 4x_2 & -4x_3 & = & 2 \\
2x_1 & -2x_2 & -2x_3 & = & 8 \\
x_1 & +3x_2 & -3x_3 & = & 2
\end{array}\right. \tag{2.9}
$$

Gauss eliminationsmetod går ut på att systematiskt eliminera variablerna ur ekvationerna. För att kunna starta måste vi ha en term med x_1 i den första ekvationen. Därför låter vi den första och den sista ekvationen byta plats, så som pilarna markerar. Vi får

$$
\left\{\begin{array}{rrrcrl}
x_1 & +3x_2 & -3x_3 & = & 2 & {\scriptstyle -2} \\
2x_1 & -2x_2 & -2x_3 & = & 8 & {\scriptstyle +} \\
 & 4x_2 & -4x_3 & = & 2 &
\end{array}\right. \tag{2.10}
$$

För att eliminera x_1 ur ekvation 2 multiplicerar vi ekvation 1 med -2 och adderar resultatet till ekvation 2. Notera hur vi markerar denna operation med en pil. Variabeln x_1 förekommer ej i ekvation 3 och behöver därför inte elimineras. Resultatet blir

$$
\left\{\begin{array}{rrrcrl}
x_1 & +3x_2 & -3x_3 & = & 2 & \\
 & -8x_2 & +4x_3 & = & 4 & {\scriptstyle \frac{1}{2}} \\
 & 4x_2 & -4x_3 & = & 2 & {\scriptstyle +}
\end{array}\right. \tag{2.11}
$$

Vi ska senare bevisa att ekvationssystemen (2.9) och (2.11) är ekvivalenta i den meningen att de har samma lösningsmängd (sats 2.1). Vi fortsätter genom att använda ekvation 2 för att eliminera x_2 ur ekvation 3 med den markerade operationen. Vi får

$$
\left\{\begin{array}{rrrcr}
x_1 & +3x_2 & -3x_3 & = & 2 \\
 & -8x_2 & +4x_3 & = & 4 \\
 & & -2x_3 & = & 4
\end{array}\right. \tag{2.12}
$$

Vi har nu eliminerat x_1 ur ekvation 2 och 3 och x_2 ur ekvation 3. Detta kallas *framåtelimination*. Ekvationssystemet har *trappstegsform* och kan lösas enkelt: vi ser genast från ekvation 3 att $x_3 = -2$, insättning i ekvation 2 ger sedan $x_2 = -3/2$ och med kända värden på x_2 och x_3 ger ekvation 1 till sist $x_1 = 1/2$. Detta förfarande kallas *bakåtsubstitution*.

Vi nöjer oss inte med detta utan fortsätter genom att eliminera bakåt. Vi eliminerar x_3 ur ekvation 1 och 2:

$$
\left\{\begin{array}{rrrcrll}
x_1 & +3x_2 & -3x_3 & = & 2 & {\scriptstyle +} & \\
 & -8x_2 & +4x_3 & = & 4 & & {\scriptstyle +} \\
 & & -2x_3 & = & 4 & {\scriptstyle -\frac{3}{2}} & {\scriptstyle 2}
\end{array}\right. \tag{2.13}
$$

Vi får

$$
\left\{\begin{array}{rrrcrl}
x_1 & +3x_2 & & = & -4 & {\scriptstyle +} \\
 & -8x_2 & & = & 12 & {\scriptstyle \frac{3}{8}} \\
 & & -2x_3 & = & 4 &
\end{array}\right. \tag{2.14}
$$

vilket leder till

$$\begin{cases} x_1 & & & = & \frac{1}{2} \\ & -8x_2 & & = & 12 \mid -\frac{1}{8} \\ & & -2x_3 & = & 4 \mid -\frac{1}{2} \end{cases} \tag{2.15}$$

Slutligen multiplicerar vi ekvation 2 och 3 med $-1/8$ respektive $-1/2$ för att få den utlovade standardformen:

$$\begin{cases} x_1 & & & = & \frac{1}{2} \\ & x_2 & & = & -\frac{3}{2} \\ & & x_3 & = & -2 \end{cases} \tag{2.16}$$

Vi avläser nu enkelt att vi har en unik lösning: $x_1 = 1/2$, $x_2 = -3/2$, $x_3 = -2$. Lösningsmängden består av en enda punkt $(1/2, -3/2, -2)$.

Vi inser att det är onödigt att skriva ut variablerna i ovanstående räkningar; det räcker att ange vilka operationer som görs på koefficienterna i ekvationssystemet. Vi erinrar oss först begreppet matris.

Definition 2.2 (Matris) En *matris* A av *typ* $m \times n$ är ett rektangulärt schema av reella tal ordnade i m rader och n kolonner. Vi skriver $A \in \mathbb{R}^{m \times n}$ och

$$A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \quad a_k = \begin{bmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{bmatrix}, \ k = 1, \ldots, n \tag{2.17}$$

Talen a_{jk} kallas matriselement och a_k är matrisens kolonner.

En matris $v \in \mathbb{R}^{1 \times n} = \mathbb{R}^n$ av typ $1 \times n$ består av en enda rad och kallas *radvektor*. En matris $v \in \mathbb{R}^{n \times 1} = \mathbb{R}^n$ av typ $n \times 1$ består av en enda kolonn och kallas *kolonnvektor*.

Obs att $\mathbb{R}^n$ kan beteckna både rad- och kolonnvektorer i sammanhang där det inte är viktigt vilken typ av vektor det är. Vi ska mestadels tänka på $\mathbb{R}^n$ som kolonnvektorer. Vi upprepar nu vårt räkneexempel med matrisbeteckningar.

Exempel 2.4 (Gauss eliminationsmetod på matrisform) Vi löser det linjära ekvationssystemet

$$\begin{cases} & 4x_2 & -4x_3 & = & 2 \\ 2x_1 & -2x_2 & -2x_3 & = & 8 \\ x_1 & +3x_2 & -3x_3 & = & 2 \end{cases} \tag{2.18}$$

Vi bildar *koefficientmatrisen* och *högerledsvektorn*

$$A = \begin{bmatrix} 0 & 4 & -4 \\ 2 & -2 & -2 \\ 1 & 3 & -3 \end{bmatrix} \in \mathbb{R}^{3 \times 3}, \quad b = \begin{bmatrix} 2 \\ 8 \\ 2 \end{bmatrix} \in \mathbb{R}^{3 \times 1} \tag{2.19}$$

och *totalmatrisen*

$$\begin{bmatrix} A & b \end{bmatrix} = \begin{bmatrix} 0 & 4 & -4 & 2 \\ 2 & -2 & -2 & 8 \\ 1 & 3 & -3 & 2 \end{bmatrix} \in \mathbb{R}^{3\times 4} \tag{2.20}$$

Räkningarna från förra exemplet kan nu skrivas på kompakt form som operationer på totalmatrisen:

$$\begin{bmatrix} 0 & 4 & -4 & 2 \\ 2 & -2 & -2 & 8 \\ 1 & 3 & -3 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -3 & 2 \\ 2 & -2 & -2 & 8 \\ 0 & 4 & -4 & 2 \end{bmatrix} \tag{2.21}$$

$$\sim \begin{bmatrix} 1 & 3 & -3 & 2 \\ 0 & -8 & 4 & 4 \\ 0 & 4 & -4 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -3 & 2 \\ 0 & -8 & 4 & 4 \\ 0 & 0 & -2 & 4 \end{bmatrix} \tag{2.22}$$

$$\sim \begin{bmatrix} 1 & 3 & 0 & -4 \\ 0 & -8 & 0 & 12 \\ 0 & 0 & -2 & 4 \end{bmatrix} \sim \begin{array}{c} \begin{bmatrix} 1 & 0 & 0 & \frac{1}{2} \\ 0 & -8 & 0 & 12 \\ 0 & 0 & -2 & 4 \end{bmatrix} \end{array} \begin{array}{l} \\ \mid -\frac{1}{8} \\ \mid -\frac{1}{2} \end{array} \tag{2.23}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & -\frac{3}{2} \\ 0 & 0 & 1 & -2 \end{bmatrix} \tag{2.24}$$

Den sista matrisen är totalmatrisen för ekvationssystemet

$$\begin{cases} x_1 & = & \frac{1}{2} \\ x_2 & = & -\frac{3}{2} \\ x_3 & = & -2 \end{cases} \tag{2.25}$$

vars lösning är uppenbar.

Vi har använt radoperationer av tre slag.

Definition 2.3 **(Elementära radoperationer)** En matris kan transformeras till en annan matris genom en av följande tre elementära radoperationer.

1. Addera en multipel av en rad till en annan rad.
2. Multiplicera en rad med en konstant som inte är 0 (skalning).
3. Låt två rader byta plats (permutation).

Det är klart att elementära radoperationer är inverterbara, dvs vi kan transformera tillbaka med samma typ av radoperation.

Definition 2.4 (Radekvivalens) Två matriser A och B är radekvivalenta om de kan transformeras till varandra genom elementära radoperationer. Vi skriver $A \sim B$.

Sats 2.1 (Radekvivalens bevarar lösningsmängd) Om totalmatriserna för två linjära ekvationssystem är radekvivalenta, så har de två systemen samma lösningsmängd.

Bevis. Det är klart att radoperationer av typ 2 och 3 inte ändrar lösningsmängden. För radoperation av typ 1 resonerar vi så här. Antag att $x_1, \dots, x_n$ löser de två ekvationerna

$$\begin{cases} a_1x_1 + \cdots + a_nx_n - b = 0 \\ c_1x_1 + \cdots + c_nx_n - d = 0 \end{cases} \tag{2.26}$$

Efter radoperation av typ 1 får vi det transformerade systemet

$$\begin{cases} a_1x_1 + \cdots + a_nx_n - b = 0 \\ (c_1x_1 + \cdots + c_nx_n - d) + \alpha(a_1x_1 + \cdots + a_nx_n - b) = 0 \end{cases} \tag{2.27}$$

Eftersom (2.26) innebär att de två parenteserna är lika med 0, så ser vi att det transformerade systemet också är uppfyllt. Å andra sidan: om det transformerade ekvationssystemet är uppfyllt, så kan vi transformera tillbaka med en transformation av typ 1 (multiplicera rad 1 med $-\alpha$ och addera till rad 2) och vi konstaterar att det ursprungliga systemet är uppfyllt. $\square$

Vi gör ett räkneexempel till.

Exempel 2.5 (Gauss eliminationsmetod) Vi löser det linjära ekvationssystemet

$$\begin{cases} 2x_1 & -3x_2 & +2x_3 & = & 1 \\ & x_2 & -4x_3 & = & 8 \\ 4x_1 & -8x_2 & +12x_3 & = & 1 \end{cases} \tag{2.28}$$

Vi gör framåtelemination på matrisform:

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 4 & -8 & 12 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & -2 & 8 & -1 \end{bmatrix} \tag{2.29}$$

$$\sim \begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 15 \end{bmatrix} \tag{2.30}$$

Här är den tredje ekvationen $0x_1 + 0x_2 + 0x_3 = 15$. Det betyder att det finns ingen lösning, dvs ekvationssystemet (2.28) är inkonsistent.

Om vi ändrar den tredje ekvationen i (2.28) till $4x_1 - 8x_2 + 12x_3 = -14$, så får vi istället

$$\begin{cases} 2x_1 & -3x_2 & +2x_3 & = & 1 \\ & x_2 & -4x_3 & = & 8 \\ 4x_1 & -8x_2 & +12x_3 & = & -14 \end{cases} \tag{2.31}$$

och

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 4 & -8 & 12 & -14 \end{bmatrix} \sim \begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & -2 & 8 & -16 \end{bmatrix} \tag{2.32}$$

$$\sim \begin{bmatrix} 1 & -\frac{3}{2} & 1 & \frac{1}{2} \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -5 & \frac{25}{2} \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.33}$$

Det resulterande ekvationssystemet är

$$\begin{cases} x_1 & & -5x_3 & = & \frac{25}{2} \\ & x_2 & -4x_3 & = & 8 \end{cases} \tag{2.34}$$

Det är konsistent för vi kan välja x_3 fritt och sedan lösa ut x_1, x_2:

$$\begin{cases} x_1 & = & \frac{25}{2} + 5x_3 \\ x_2 & = & 8 + 4x_3 \end{cases} \tag{2.35}$$

Vi har alltså oändligt många lösningar. Vi säger att variabeln x_3 är *fri* medan variablerna x_1, x_2 är *bundna*. Med $x_3 = t$, där $t \in \mathbb{R}$ är en parameter, kan vi skriva lösningsmängden på parameterform:

$$\begin{cases} x_1 & = & \frac{25}{2} + 5t \\ x_2 & = & 8 + 4t \\ x_3 & = & t \end{cases}, \quad t \in \mathbb{R} \tag{2.36}$$

Vi säger att matrisen i (2.30) har *trappstegsform* och att den sista matrisen i (2.33) har *reducerad trappstegsform*. Från den reducerade trappstegsformen kan vi enkelt läsa av ekvationssystemets lösningsmängd. Vi ska nu precisera dessa beteckningar.

Definition 2.5 (Trappstegsform) Det första nollskilda elementet i en rad kallas radens *pivotelement*. En matris har *trappstegsform* om den uppfyller följande tre villkor.

1. Alla rader som inte är enbart består av nollor är placerade över alla rader med enbart nollor.
2. I varje rad är pivotelementet placerat i en kolonn som är till höger om pivotelementet i raden ovanför.
3. Alla matriselement under ett pivotelement är noll.

En trappstegsmatris har *reducerad trappstegsform* om den har ytterligare två egenskaper:

4. Alla pivotelement är lika med 1.
5. Varje pivotelement är det enda nollskilda elementet i sin kolonn.

Varje matris kan transformeras till trappstegsform med hjälp av radoperationer, men trappstegsformen är inte unik utan beror på vilka radoperationer som används. Den reducerade trappstegsformen som vi får är dock unik enligt nästa sats.

Sats 2.2 (Reducerad trappstegsform är unik) Varje matris är radekvivalent med en unik reducerad trappstegsmatris.

Bevis. Existensen av en radekvivalent reducerad trappstegsmatris följer av algoritmen som presenteras i ruta 2.1 nedan. Beviset av entydigheten är svårare och vi hoppar över det. □

Definition 2.6 (Pivotposition) Pivotpositionerna i en matris är de positioner som svarar mot pivotelementen i dess reducerade trappstegsmatris. Motsvarande kolonner kallas pivotkolonner.

Enligt exempel 2.5 har vi

$$\begin{bmatrix} \mathbf{2} & -3 & 2 & 1 \\ 0 & \mathbf{1} & -4 & 8 \\ 4 & -8 & 12 & -14 \end{bmatrix} \sim \begin{bmatrix} \mathbf{1} & 0 & -5 & \frac{25}{2} \\ 0 & \mathbf{1} & -4 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.37}$$

Här är pivotpositionerna i båda matriserna markerade med fetstil och pivotkolonnerna i den ursprungliga matrisen är

$$\begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}, \quad \begin{bmatrix} -3 \\ 1 \\ -8 \end{bmatrix} \tag{2.38}$$

Ruta 2.1 (Transformation till reducerad trappstegsform) Denna metod transformerar en given matris till en radekvivalent reducerad trappstegsmatris. Vi börjar med framåtelimination.

1. Låt rad 1 vara den aktuella raden.
2. Se till så att pivotelementet i den aktuella raden inte är till höger om något annat pivotelement genom att eventuellt låta den byta plats med en rad som är längre ned (radoperation av typ 3).
3. Skapa nollor under det aktuella pivotelementet genom att göra radoperationer av typ 1.
4. Låt nästa rad vara den aktuella raden.
5. Upprepa steg 2 och 3 tills inga pivotelement återstår.

Vi har nu en trappstegsmatris. Vi avslutar med bakåtelimination och skalning.

6. Börja med det pivotelement som är längst till höger och arbeta åt vänster och skapa nollor ovanför alla pivotelement med radoperationer av typ 1.
7. Skala om raderna så att alla pivotelement blir 1 (radoperationer av typ 2).

Sats 2.3 (Existens och entydighet) Ett linjärt ekvationssystem är konsistent om och endast om den sista kolonnen i totalmatrisen inte är en pivotkolonn. Om det är konsistent så består lösningsmängden antingen av en unik lösning eller av oändligt många lösningar.

Bevis. Den sista kolonnen i totalmatrisen är en pivotkolonn om och endast om den har en trappstegsform med en rad på formen $[0 \; \cdots \; 0 \; b]$ med $b \neq 0$. Detta svarar mot ett inkonsistent system (se (2.30)). Alltså är systemet konsistent om och endast om den sista kolonnen i totalmatrisen inte är en pivotkolonn. Om systemet är konsistent, så kan vi läsa av lösningsmängden från den radekvivalenta reducerade trappstegsmatrisen. Vi säger att en variabel som hör till en pivotposition är en *bunden variabel*; övriga variabler kallas *fria variabler*. Det finns alltid minst en bunden variabel. Om alla variabler är bundna så har vi en unik lösning. Om det finns fria variabler, så kan vi välja dessa fritt och sedan lösa ut de bundna variablerna. Då har vi oändligt många lösningar. $\square$

Ruta 2.2 (Lösning av linjärt ekvationssystem med radreducering)

1. Skriv ned ekvationssystemets totalmatris.
2. Transformera totalmatrisen till en radekvivalent trappstegsmatris och använd denna för att avgöra om ekvationssystemet är konsistent eller ej.
3. Om det är konsistent så fortsätt att transformera till reducerad trappstegsform. Identifiera vilka variabler som är fria respektive bundna.
4. Skriv ned det ekvationssystem som svarar mot den reducerade trappstegsmatrisen. Sätt eventuella fria variabler lika med varsin parameter och lös ut de bundna variablerna.

2.2 Vektorekvation och matrisekvation

Vi introducerar två alternativa sätt att betrakta ett linjärt ekvationssystem: dels som en vektorekvation och dels som en matrisekvation. Vi börjar med att introducera vektorer i $\mathbb{R}^n$ och deras algebraiska egenskaper, dvs hur vi räknar med vektorer.

Definition 2.7 (Vektorrummet $\mathbb{R}^n$) En *vektor* i $\mathbb{R}^n$ (*kolonnvektor*) är en ordnad lista av n reella tal, som vi skriver på kolonnform

$$v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}, \quad v \in \mathbb{R}^n \tag{2.39}$$

På liknande sätt kan vi definiera *radvektor* $v = [v_1 \ \cdots \ v_n]$. I fortsättningen kommer vi huvudsakligen att tänka på $\mathbb{R}^n$ som en mängd av kolonnvektorer även om radvektorer också kommer att dyka upp ibland. Då $n = 3$ kan vi identifiera vektorer i $\mathbb{R}^3$ med geometriska vektorer i rummet, se sektion 1.4. Många böcker skriver vektorer i $\mathbb{R}^n$ med fetstil för att skilja dem från skalärer, men vi gör inte det. Vi skriver endast geometriska vektorer med fetstil.

Definition 2.8 (Linjära räkneoperationer i $\mathbb{R}^n$) Låt u och v vara vektorer i $\mathbb{R}^n$ och α en skalär, dvs

$$u = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, \; v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}, \; \alpha \in \mathbb{R} \tag{2.40}$$

Vi gör följande definitioner.

$$u = v \text{ om } u_i = v_i, \quad i = 1, \ldots, n \qquad \text{(likhet)} \tag{2.41}$$

$$u + v = \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix} \qquad \text{(addition)} \tag{2.42}$$

$$\alpha u = u\alpha = \begin{bmatrix} \alpha u_1 \\ \vdots \\ \alpha u_n \end{bmatrix} \qquad \text{(multiplikation med skalär)} \tag{2.43}$$

$$0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \text{(nollvektorn)} \tag{2.44}$$

$$-u = \begin{bmatrix} -u_1 \\ \vdots \\ -u_n \end{bmatrix} \qquad \text{(motsatta vektorn)} \tag{2.45}$$

Notera att räkneoperationerna definieras elementvis. Vi skriver helst skalären före vektorn i (2.43), dvs αu. Med hjälp av motsatta vektorn kan vi även definiera *subtraktion*: $u - v = u + (-v)$. Detta är alltså inte en grundläggande operation utan följer av addition och multiplikation med skalär.

Sats 2.4 (Räkneregler i $\mathbb{R}^n$)

$$u + v = v + u \qquad \text{(kommutativ lag)} \tag{2.46}$$

$$(u + v) + w = u + (v + w) \qquad \text{(associativ lag)} \tag{2.47}$$

$$u + 0 = u \qquad \text{(addition av nollvektorn)} \tag{2.48}$$

$$u + (-u) = 0 \qquad \text{(addition av motsatta vektorn)} \tag{2.49}$$

$$\alpha(u + v) = \alpha u + \alpha v \qquad \text{(distributiv lag)} \tag{2.50}$$

$$(\alpha + \beta)u = \alpha u + \beta u \qquad \text{(distributiv lag)} \tag{2.51}$$

$$(\alpha\beta)u = \alpha(\beta u) \qquad \text{(associativ lag)} \tag{2.52}$$

$$1u = u \tag{2.53}$$

Med ett mer rigoröst språkbruk innebär (2.48) att nollvektorn är en *additiv enhet* och (2.49) att $-u$ är en *additiv invers*.

Bevis. Skriv på elementform och använd räkneregler för reella tal. □

Observera att dessa räkneregler är desamma som för geometriska vektorer, se sats 1.1 och sats 1.2. En mängd som har räkneoperationerna addition och multiplikation med skalär och som uppfyller dessa räkneregler kallas *vektorrum*. Vi återkommer till detta senare.

Med hjälp av addition och multiplikation med skalär kan vi bilda nya vektorer: vi säger att vi bildar linjär kombination.

Definition 2.9 (Linjär kombination i $\mathbb{R}^n$) Låt $\{v_k\}_{k=1}^p$ vara en mängd av vektorer i $\mathbb{R}^n$ och $\{c_k\}_{k=1}^p$ vara en mängd av reella tal. Vektorn

$$u = c_1 v_1 + \cdots + c_p v_p = \sum_{k=1}^{p} c_k v_k \tag{2.54}$$

kallas *linjär kombination* av vektorerna v_k med koefficienterna (vikterna) c_k.

Här vore det praktiskt att skriva vektorer med fetstil, dvs

$$\boldsymbol{u} = c_1 \boldsymbol{v}_k + \cdots + c_p \boldsymbol{v}_p \tag{2.55}$$

Vi gör ändå inte det. Det bör framgå av sammanhanget vilka symboler som syftar på vektorer (u, v_k) respektive skalärer (c_k). Vi noterar också att elementen i vektor nummer k bör numreras enligt

$$v_k = \begin{bmatrix} v_{1k} \\ \vdots \\ v_{nk} \end{bmatrix} \tag{2.56}$$

för att stämma överens med numreringen av matriselement i (2.17). *Linjärkombinationen i* (2.54) *blir då*

$$\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = c_1 \begin{bmatrix} v_{11} \\ \vdots \\ v_{n1} \end{bmatrix} + \cdots + c_p \begin{bmatrix} v_{1p} \\ \vdots \\ v_{np} \end{bmatrix} = \sum_{k=1}^{p} c_k \begin{bmatrix} v_{1k} \\ \vdots \\ v_{nk} \end{bmatrix} \tag{2.57}$$

Det följer av definition 2.8 att vektorn u i (2.54) också är en vektor i $\mathbb{R}^n$. Linjär kombination är ett viktigt sätt att skapa nya vektorer. Omvänt kan vi fråga oss om en given vektor alltid kan genereras som en linjär kombination av givna vektorer. Det vill säga: givet vektorer $v_1, \ldots, v_p$ och b i $\mathbb{R}^n$, kan vi skriva b som en linjär kombination av $\{v_k\}_{k=1}^p$?

Exempel 2.6 (Vektorekvation) Låt

$$v_1 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}, \; v_2 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}, \; b = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} \tag{2.58}$$

Kan vi skriva $c_1 v_1 + c_2 v_2 = b$ för några koefficienter c_1, c_2? Dvs

$$c_1 \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} \tag{2.59}$$

vilket är detsamma som

$$\begin{cases} c_1 + 2c_2 = 7 \\ -2c_1 + 5c_2 = 4 \\ -5c_1 + 6c_2 = -3 \end{cases} \tag{2.60}$$

Detta är ett linjärt ekvationssystem för c_1, c_2. Vi löser det genom att skriva upp totalmatrisen och göra en radreducering:

$$\begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \tag{2.61}$$

Vi ser att systemet är konsistent (inget pivotelement i sista kolonnen) med unik lösning $c_1 = 3$, $c_2 = 2$. Det betyder att $b = 3v_1 + 2v_2$.

Mer allmänt kan vi säga följande: Vektorekvationen

$$c_1 v_2 + \cdots + c_p v_p = b \tag{2.62}$$

har samma lösningsmängd som det linjära ekvationssystemet med totalmatrisen $[v_1 \; \cdots \; v_p \; b]$. *Vektorn* b *är en linjär kombination av* $\{v_k\}_{k=1}^{p}$ *om och endast om detta linjära ekvationssystem är konsistent. Detta motiverar följande definition.*

Definition 2.10 (Linjärt hölje) Mängden av alla linjärkombinationer av vektorerna $\{v_k\}_{k=1}^p$ kallas det *linjära höljet* av $\{v_k\}_{k=1}^p$. Mängden betecknas[a]

$$\operatorname{span}\{v_1, \ldots, v_p\} = \operatorname{span}\{v_k\}_{k=1}^p \tag{2.63}$$

och ges alltså av

$$\operatorname{span}\{v_k\}_{k=1}^p = \left\{u \in \mathbb{R}^n \;\middle|\; u = \sum_{k=1}^p c_k v_k \text{ för några } c_k \in \mathbb{R}\right\} \tag{2.64}$$

Vi säger också att det linjära höljet *spänns upp* av vektorerna $\{v_k\}_{k=1}^p$.

a. Beteckningen span kommer från engelskans "span" = "spänna upp". Lämpligt svenskt ord saknas.

Obs att ett linjärt hölje alltid innehåller nollvektorn, $0 \in \operatorname{span}\{v_k\}_{k=1}^p$, (välj alla $c_k = 0$).

Exempel 2.7 (Geometrisk tolkning av linjärt hölje) Vektorer i $\mathbb{R}^3$ kan identifieras med geometriska vektorer i rummet. Linjära höljet av en vektor $v \in \mathbb{R}^3$ blir då

$$\operatorname{span}\{v\} = \{u \in \mathbb{R}^3 \mid u = tv,\ t \in \mathbb{R}\} \tag{2.65}$$

vilket kan tolkas som ekvationen för *en rät linje genom origo* med riktningsvektor v (förutsatt att $v \neq 0$, se (1.116)). Linjära höljet av två vektorer $u, v \in \mathbb{R}^3$ blir

$$\operatorname{span}\{u, v\} = \{w \in \mathbb{R}^3 \mid w = su + tv,\ s, t \in \mathbb{R}\} \tag{2.66}$$

vilket är ekvationen för *ett plan genom origo* på parameterform (såvida inte u och v är kolinjära, dvs $u = cv$ för något $c \in \mathbb{R}$, se (1.129)). Alltså: en vektor spänner (typiskt) upp en rät linje genom origo och två vektorer spänner (typiskt) upp ett plan genom origo.

Vår tidigare fråga om lösbarhet av vektorekvationen (2.62) *blir nu ekvivalent med frågan om* $b \in \operatorname{span}\{v_1, \ldots, v_p\}$*?*

Vi ska nu se hur vi kan formulera ett linjärt ekvationssystem som en matrisekvation på formen $Ax = b$.

Definition 2.11 (Matris-vektorprodukt) Låt $A \in \mathbb{R}^{m\times n}$ vara en matris med kolonner $a_1, \ldots, a_n \in \mathbb{R}^m$ och $x \in \mathbb{R}^n$. Produkten av A med x är

$$Ax = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 a_1 + \cdots + x_n a_n = \sum_{k=1}^{n} x_k a_k \tag{2.67}$$

dvs den linjära kombinationen av kolonnerna a_k med koefficienterna x_k.

Exempel 2.8 (Matris-vektorprodukt) Vi demonstrerar en produkt:

$$\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 7 \\ 8 \end{bmatrix} = 7 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 8 \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 39 \\ 54 \\ 69 \end{bmatrix} \tag{2.68}$$

Sats 2.5 (Räkneregler för matris-vektorprodukt) Låt $A \in \mathbb{R}^{m\times n}$, $u, v \in \mathbb{R}^n$ och $c \in \mathbb{R}$. Då gäller

$$A(u+v) = Au + Av \tag{2.69}$$

$$A(cu) = c(Au) \tag{2.70}$$

Bevis. Skriv på elementform och använd räkneregler för reella tal. □

Sats 2.5 innebär att matris-vektormultiplikation är en linjär operation, *dvs den bevarar linjär kombination:*

$$A(c_1 v_1 + \cdots + c_p v_p) = c_1(Av_1) + \cdots + c_p(Av_p) \tag{2.71}$$

Låt $A = \begin{bmatrix} a_1 & \ldots & a_n \end{bmatrix} \in \mathbb{R}^{m\times n}$ och $b \in \mathbb{R}^m$. Vi kan nu säga att följande ekvationer har samma lösningsmängd:

1. *Linjärt ekvationssystem med totalmatrisen $\begin{bmatrix} a_1 & \cdots & a_n & b \end{bmatrix}$.*
2. *Vektorekvationen $x_1 a_1 + \cdots + x_n a_n = b$.*
3. *Matrisekvationen $Ax = b$.*

Vi avslutar med en sats som ger fyra ekvivalenta villkor som garanterar att vårt ekvationsystem är lösbart för varje högerledsvektor.

Sats 2.6 (Existens av lösningar) Låt $A = [a_1 \ \dots \ a_n] \in \mathbb{R}^{m \times n}$. Följande påståenden är ekvivalenta.

1. Ekvationen $Ax = b$ har lösning för *varje* $b \in \mathbb{R}^m$.
2. *Varje* $b \in \mathbb{R}^m$ är linjär kombination av $a_1, \dots, a_n$, dvs $b \in \operatorname{span}\{a_1, \dots, a_n\}$.
3. Kolonnerna i A spänner upp *hela* $\mathbb{R}^m$, dvs $\operatorname{span}\{a_1, \dots, a_n\} = \mathbb{R}^m$.
4. A har en pivotposition i *varje* rad.

Bevis. Det är klart från våra definitioner att $1 \iff 2 \iff 3$. Det återstår att visa $1 \iff 4$. Låt totalmatrisen för ekvationen $Ax = b$ ha en radekvivalent reducerad trappstegsmatris på formen $[A\ b] \sim [U\ d]$.

Vi visar först $4 \implies 1$. Om 4 är sann så har vi ingen rad på formen $[0 \ \cdots \ 0]$ i U. Då är $Ux = d$ lösbar för alla d och därmed är även $Ax = b$ lösbar för alla b, dvs 1 är sann, (se sats 2.3).

Vi visar nu $1 \implies 4$, dvs $\neg 4 \implies \neg 1$ (icke 4 medför icke 1). Om 4 är falsk så är sista raden i U endast nollor. Tag då d med 1 som det sista elementet, så att $Ux = d$ blir inkonsistent. Då är också $Ax = b$ inkonsistent och 1 är falsk. $\square$

2.3 Lösningsmängden

Definition 2.12 (Homogent och inhomogent ekvationssystem) Ett linjärt ekvationssystem $Ax = b$ kallas *homogent* om $b = 0$, dvs $Ax = 0$. Om $b \neq 0$, så kallas systemet *inhomogent*.

Det är klart att nollvektorn $x = 0$ är lösning till det homogena systemet $Ax = 0$. Denna lösning kallas den *triviala lösningen*. Vi kan nu fråga oss om det finns andra lösningar, dvs om det finns icke-triviala lösningar. Från metoden i ruta 2.2 får vi svaret:

$Ax = 0$ har icke-trivial lösning om och endast om systemet har minst en fri variabel.

Vi illustrerar detta i ett exempel.

Exempel 2.9 (Homogent ekvationssystem) Vi löser det homogena systemet

$$\begin{cases} x_1 + 3x_2 - 3x_3 + 7x_4 = 0 \\ 2x_1 + 7x_2 - 10x_3 + 19x_4 = 0 \end{cases} \tag{2.72}$$

Vi skriver totalmatrisen och transformerar till reducerad trappstegsform:

$$\begin{bmatrix} 1 & 3 & -3 & 7 & 0 \\ 2 & 7 & -10 & 19 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -3 & 7 & 0 \\ 0 & 1 & -4 & 5 & 0 \end{bmatrix} \tag{2.73}$$

$$\sim \begin{bmatrix} 1 & 0 & 9 & -8 & 0 \\ 0 & 1 & -4 & 5 & 0 \end{bmatrix} \tag{2.74}$$

Motsvarande ekvationssystem är

$$\begin{cases} x_1 + 9x_3 - 8x_4 = 0 \\ x_2 - 4x_3 + 5x_4 = 0 \end{cases} \tag{2.75}$$

Vi ser att x_3 och x_4 är fria och vi sätter $x_3 = s$, $x_4 = t$. Sedan får vi $x_1 = -9s + 8t$, $x_2 = 4s - 5t$, dvs lösningen är på parameterform

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -9s + 8t \\ 4s - 5t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} -9 \\ 4 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 8 \\ -5 \\ 0 \\ 1 \end{bmatrix}, \quad s, t \in \mathbb{R} \tag{2.76}$$

Vi har alltså oändligt många icke-triviala lösningar. Med

$$u = \begin{bmatrix} -9 \\ 4 \\ 1 \\ 0 \end{bmatrix}, \quad v = \begin{bmatrix} 8 \\ -5 \\ 0 \\ 1 \end{bmatrix} \tag{2.77}$$

har vi $x = su + tv$, $s, t \in \mathbb{R}$, dvs lösningsmängden är linjära höljet av u och v, $\operatorname{span}\{u, v\}$. Eftersom u inte är en multipel av v (kolla detta, se exempel 2.12), så behövs båda vektorerna för att beskriva hela lösningsmängden. Hellre än att säga att vi har oändligt många icke-triviala lösningar, säger vi att *lösningsmängden spänns upp av de två vektorerna* u och v. Obs att triviala lösningen $x = 0$ ingår också i $\operatorname{span}\{u, v\}$.

Nu tittar vi på motsvarande inhomogena system.

Exempel 2.10 (Inhomogent ekvationssystem) Vi väljer en högerledsvektor $b = \begin{bmatrix} 3 \\ 10 \end{bmatrix}$ och löser det inhomogena systemet

$$\begin{cases} x_1 + 3x_2 - 3x_3 + 7x_4 = 3 \\ 2x_1 + 7x_2 - 10x_3 + 19x_4 = 10 \end{cases} \tag{2.78}$$

Vi skriver totalmatrisen och transformerar till reducerad trappstegsform:

$$\begin{bmatrix} 1 & 3 & -3 & 7 & 3 \\ 2 & 7 & -10 & 19 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -3 & 7 & 3 \\ 0 & 1 & -4 & 5 & 4 \end{bmatrix} \tag{2.79}$$

$$\sim \begin{bmatrix} 1 & 0 & 9 & -8 & -9 \\ 0 & 1 & -4 & 5 & 4 \end{bmatrix} \tag{2.80}$$

Motsvarande ekvationssystem är

$$\begin{cases} x_1 + 9x_3 - 8x_4 = -9 \\ x_2 - 4x_3 + 5x_4 = 4 \end{cases} \tag{2.81}$$

Vi ser att x_3 och x_4 är fria och vi sätter $x_3 = s, x_4 = t$. Sedan får vi $x_1 = -9 - 9s + 8t$, $x_2 = 4 + 4s - 5t$, dvs lösningen är på parameterform

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -9 - 9s + 8t \\ 4 + 4s - 5t \\ s \\ t \end{bmatrix} = \begin{bmatrix} -9 \\ 4 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -9 \\ 4 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 8 \\ -5 \\ 0 \\ 1 \end{bmatrix}, \quad s, t \in \mathbb{R} \tag{2.82}$$

Med

$$u = \begin{bmatrix} -9 \\ 4 \\ 1 \\ 0 \end{bmatrix}, \quad v = \begin{bmatrix} 8 \\ -5 \\ 0 \\ 1 \end{bmatrix}, \quad w = \begin{bmatrix} -9 \\ 4 \\ 0 \\ 0 \end{bmatrix} \tag{2.83}$$

har vi $x = w + su + tv$, $s, t \in \mathbb{R}$. Vi noterar att $Aw = b$ och $A(su + tv) = 0$. Lösningsmängden ges alltså på formen $x = x_{\mathrm{p}} + x_{\mathrm{h}}$, där $x_{\mathrm{p}} = w$ är en speciell lösning till det inhomogena systemet och $x_{\mathrm{h}} = su + tv \in \operatorname{span}\{u, v\}$ är en godtycklig lösning (allmänna lösningen) till det homogena systemet. Vektorn x_{p} kallas en *partikulärlösning* medan x_{h} kallas *homogenlösning*.

Denna observation om lösningsmängdens form gäller allmänt. Vi sammanfattar det i nästa sats.

Sats 2.7 (Lösningsmängden till konsistent system) Antag att $Ax = b$ är konsistent med en lösning x_{p}. Då ges lösningsmängden av $x = x_{\mathrm{p}} + x_{\mathrm{h}}$, där x_{h} är en godtycklig lösning till det homogena systemet $Ax = 0$. Lösningen är unik om och endast om det homogena systemet endast har triviala lösningen $x_{\mathrm{h}} = 0$.

Bevis. Antag att x är en annan lösning än x_{p}, dvs vi har både $Ax_{\mathrm{p}} = b$ och $Ax = b$. Bilda $u = x - x_{\mathrm{p}}$. Då gäller, enligt sats 2.5, att $Au = A(x - x_{\mathrm{p}}) = Ax - Ax_{\mathrm{p}} = b - b = 0$, dvs u är en homogenlösning. Alltså är $x = x_{\mathrm{p}} + u$, partikulärlösning plus homogenlösning. Å andra sidan: om u är en homogenlösning, så har vi $A(x_\mathrm{p}+u) = Ax_\mathrm{p} + Au = b + 0 = b$, dvs $x = x_\mathrm{p} + u$ är en lösning. Påståendet om unik lösning är nu självklart. $\square$

Vi har redan sett att lösningen ges av partikulärlösning plus homogenlösning i samband med linjära differentialekvationer, se sats 3.7 i del II.

2.4 Linjärt oberoende

Vi har definierat linjära höljet $\mathrm{span}\{v_k\}_{k=1}^p$ av en mängd vektorer $\{v_k\}_{k=1}^p$ som mängden av alla linjära kombinationer

$$c_1 v_1 + \cdots + c_p v_p = \sum_{k=1}^{p} c_k v_k \tag{2.84}$$

som kan bildas genom att variera koefficienterna c_k. Vi kan nu fråga oss om alla vektorerna v_k behövs för att spänna upp $\mathrm{span}\{v_k\}_{k=1}^p$. Om till exempel en av v_k redan är en linjär kombination av de övriga, så kan vi ju ta bort den och de övriga vektorerna spänner upp samma mängd. För att kunna precisera denna tanke gör vi en definition.

Definition 2.13 (Linjärt oberoende) En mängd $\{v_1, \ldots, v_p\}$ av vektorer i $\mathbb{R}^n$ är *linjärt oberoende* om vektorekvationen

$$c_1 v_1 + \cdots + c_p v_p = 0 \tag{2.85}$$

endast har den triviala lösningen, $c_k = 0$ för $k = 1, \ldots, p$. Mängden är annars *linjärt beroende*, dvs om det finns $\{c_k\}_{k=1}^p$ med något $c_j \neq 0$ så att (2.85) gäller. Ekvationen (2.85) ger då *linjära samband* mellan vektorerna.

Vi noterar att testet för linjärt oberoende, vektorekvationen (2.85), är ett homogent linjärt ekvationssystem. Det har alltid den triviala lösningen och ibland även icke-triviala lösningar.

Exempel 2.11 (Linjärt beroende mängd) Låt

$$v_1 = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} \tag{2.86}$$

Är mängden $\{v_1, v2, v3\}$ linjärt oberoende? Vi ställer upp och löser vektorekvationen $c_1v_1 + c_2v_2 + c_3v_3 = 0$. Totalmatrisen är

$$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 3 & 4 & 5 & 0 \\ 5 & 6 & 7 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & -2 & -4 & 0 \\ 0 & -4 & -8 & 0 \end{bmatrix} \quad | -\tfrac{1}{2} \tag{2.87}$$

$$\sim \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.88}$$

Vi ser att c_3 är en fri variabel. Det finns alltså icke-triviala lösningar och mängden är linjärt beroende. Tag till exempel $c_3 = 1$ och räkna ut $c_2 = -2c_3 = -2$, $c_1 = c_3 = 1$, så att $v_1 - 2v_2 + v_3 = 0$ är ett linjärt samband. Vi ser till exempel att $v_3 = -v_1 + 2v_2$, så att $\operatorname{span}\{v_1, v_2, v_3\} = \operatorname{span}\{v_1, v_2\}$. Mängden $\{v_1, v_2\}$ är dock linjärt oberoende, eftersom det inte finns fler fria variabler, och vi kan därför inte ta bort någon av dem utan att det linjära höljet blir mindre.

Exempel 2.12 (Linjärt oberoende mängd) Låt

$$u = \begin{bmatrix} -9 \\ 4 \\ 1 \\ 0 \end{bmatrix}, \quad v = \begin{bmatrix} 8 \\ -5 \\ 0 \\ 1 \end{bmatrix} \tag{2.89}$$

vara homogenlösningarna från exempel 2.9. Är de en linjärt oberoende mängd? Vi löser vektorekvationen $su + tv = 0$. Totalmatrisen är

$$\begin{bmatrix} -9 & 8 & 0 \\ 4 & -5 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{2.90}$$

Inga fria variabler, endast triviala lösningen, mängden är linjärt oberoende. (Vi ser redan från ekvation 3 och 4 i det ursprungliga systemet att $s = 0$, $t = 0$ utan radreducering.) Båda vektorerna behövs för att spänna upp hela lösningsmängden $\operatorname{span}\{u, v\}$.

Sats 2.8 (Linjärt beroende är ekvivalent med linjära samband) En mängd av vektorer är linjärt beroende om och endast om en av vektorerna är en linjär kombination av de övriga.

Bevis. Låt mängden vara $\{v_k\}_{k=1}^p$. Antag att v_j är en linjär kombination av de övriga vektorerna, dvs

$$v_j = \sum_{k=1,\, k\neq j}^{p} c_k v_k \tag{2.91}$$

Vi flyttar v_j till andra sidan och får $\sum_{k=1}^p c_k v_k = 0$ med $c_j = -1 \neq 0$, dvs minst en koefficient är skiljd från 0, och mängden är linjärt oberoende. Å andra sidan: om mängden är linjärt oberoende, dvs $\sum_{k=1}^p c_k v_k = 0$ med minst en $c_j \neq 0$, så kan vi lösa ut v_j:

$$c_j v_j = -\sum_{k=1,\, k\neq j}^{p} c_k v_k, \quad v_j = -\frac{1}{c_j}\sum_{k=1,\, k\neq j}^{p} c_k v_k \tag{2.92}$$

och vi ser att den är en linjär kombination av de övriga. □

> Sats 2.9 (För många vektorer är linjärt beroende) En mängd av fler än n vektorer i $\mathbb{R}^n$ är linjärt beroende.

Bevis. Låt mängden vara $\{v_k\}_{k=1}^p$ med $p > n$ och bilda matrisen $A = [v_1 \ \ldots \ v_p] \in \mathbb{R}^{n\times p}$. Testet för linjärt oberoende är $Ax = 0$. Detta system har fler variabler än ekvationer och därför minst en fri variabel. Alltså finns icke-trivial lösning och mängden är linjärt beroende. □

2.5 Linjär funktion

Vi börjar med att erinra oss begreppet funktion från definition 2.1 i del I.

> Definition 2.14 (Funktion, definitionsmängd, målmängd, värdemängd) En funktion $f\colon X \to Y$ är en *regel* som för varje *argument* x i en mängd $X = \mathcal{D}(f)$ (funktionens *definitionsmängd*) bestämmer ett *unikt värde* $y = f(x)$ i en mängd Y (funktionens *målmängd*). Mängden $\mathcal{R}(f) = \{f(x) \mid x \in X\} \subseteq Y$ av alla värden kallas funktionens *värdemängd*.

Vi skriver även

$$\begin{aligned} f\colon X &\to Y \\ x &\mapsto y = f(x) \end{aligned} \tag{2.93}$$

Vi såg i sats 2.5 att matris-vektorprodukten $x \mapsto Ax$ är en linjär operation, dvs den bevarar linjär kombination. Om $A \in \mathbb{R}^{m\times n}$, så blir $L(x) = Ax$ en funktion från $X = \mathbb{R}^n$ till $Y = \mathbb{R}^m$. Detta motiverar följande definition.

Definition 2.15 (Linjär funktion $\mathbb{R}^n \to \mathbb{R}^m$) En *linjär funktion* $L\colon \mathbb{R}^n \to \mathbb{R}^m$ är en funktion L, som har definitionsmängd $\mathbb{R}^n$ och målmängd $\mathbb{R}^m$ och som uppfyller, för alla $u, v \in \mathbb{R}^n$ och $c \in \mathbb{R}$,

$$L(u+v) = L(u) + L(v), \quad L(cu) = cL(u) \tag{2.94}$$

Ekvivalent kan vi säga att L bevarar linjär kombination:

$$L(c_1v_1 + \cdots + c_pv_p) = c_1L(v_1) + \cdots + c_pL(v_p) \tag{2.95}$$

En linjär funktion kallas också *linjär operator*, *linjär transformation* eller *linjär avbildning*. En funktion som inte är linjär kallas *ickelinjär*.

Observera att för en linjär funktion är definitionsmängden lika med hela $\mathbb{R}^n$, dvs $\mathcal{D}(L) = \mathbb{R}^n$, medan värdemängden inte behöver vara hela målmängden, $\mathcal{R}(L) \subseteq \mathbb{R}^m$. Observera också att nollvektorn i $\mathbb{R}^n$ avbildas på nollvektorn i $\mathbb{R}^m$, $L(0) = 0$ (tag $c = 0$ i (2.94)). För en linjär funktion av typen $L(x) = Ax$ med $A = [a_1 \dots a_n]$ är värdemängden $\mathcal{R}(L) = \operatorname{span}\{a_1, \dots, a_n\}$ enligt definition 2.11.

Vi ska nu se att varje linjär funktion $L\colon \mathbb{R}^n \to \mathbb{R}^m$ har formen $L\colon x \mapsto Ax$. Matrisen $A \in \mathbb{R}^{m\times n}$ bestäms genom att vi tittar på vad funktionen gör med basvektorerna i $\mathbb{R}^n$. Vi illustrerar detta först då $n = 3$ för enkelhets skull.

Basvektorerna i $\mathbb{R}^3$ är

$$e_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}, \; e_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}, \; e_3 = \begin{bmatrix}0\\0\\1\end{bmatrix} \tag{2.96}$$

Varje vektor $x \in \mathbb{R}^3$ kan skrivas som en linjär kombination av basvektorerna:

$$x = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = x_1\begin{bmatrix}1\\0\\0\end{bmatrix} + x_2\begin{bmatrix}0\\1\\0\end{bmatrix} + x_3\begin{bmatrix}0\\0\\1\end{bmatrix} = x_1e_1 + x_2e_2 + x_3e_3 \tag{2.97}$$

dvs $\mathbb{R}^3 = \operatorname{span}\{e_1, e_2, e_3\}$. Mängden $\{e_1, e_2, e_3\}$ är linjärt oberoende, det är lätt att visa. Alltså behövs alla tre basvektorerna för att spänna upp hela $\mathbb{R}^3$. Om nu $L\colon \mathbb{R}^3 \to \mathbb{R}^m$ är en linjär funktion, så har vi

$$L(x) = L(x_1e_1 + x_2e_2 + x_3e_3) = x_1L(e_1) + x_2L(e_2) + x_3L(e_3) \tag{2.98}$$

$$= \begin{bmatrix}L(e_1) & L(e_2) & L(e_3)\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = Ax \tag{2.99}$$

Matrisen $A = [L(e_1)\; L(e_2)\; L(e_3)] \in \mathbb{R}^{m\times 3}$ består av de tre kolonnvektorerna $L(e_k) \in \mathbb{R}^m$, $k = 1, 2, 3$. Kolonnerna i A är alltså bilderna av basvektorerna under avbildningen L.

Detta bevis utvidgas lätt till allmänt $\mathbb{R}^n$. Basvektorerna i $\mathbb{R}^n$ är

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, e_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots\ e_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{2.100}$$

där basvektorn $e_j \in \mathbb{R}^m$ har en etta i rad j och nollor för övrigt. Vi formulerar resultatet i en sats.

Sats 2.10 (Matrisen för linjär funktion) Om $L\colon \mathbb{R}^n \to \mathbb{R}^m$ är en linjär funktion, så är $L(x) = Ax$ med $A = [L(e_1)\ \cdots\ L(e_n)] \in \mathbb{R}^{m\times n}$.

Exempel 2.13 (Rotation i planet) Avbildningen $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ roterar ortsvektorn för varje punkt i planet moturs med en vinkel φ. Genom att rita en figur där vektorerna u, v, $u+v$ och αu utgår från origo och rotera denna, så inser vi att L är en linjär operator. Vi bestämmer dess matris genom att rotera basvektorerna:

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad L(e_1) = \begin{bmatrix} \cos(\varphi) \\ \sin(\varphi) \end{bmatrix} \tag{2.101}$$

$$e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad L(e_2) = \begin{bmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{bmatrix} \tag{2.102}$$

Matrisen är alltså

$$A = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix} \tag{2.103}$$

Exempel 2.14 (Ickelinjär funktion) Vi tittar på några envariabelfunktioner, $f\colon \mathbb{R} \to \mathbb{R}$. Funktionen $f(x) = x^3$ är inte linjär, för $f(2x) = 8x^3 \neq 2x^3 = 2f(x)$ om $x \neq 0$. Vi har även $(x+y)^3 \neq x^3 + y^3$ utom för speciella värden på x, y.

Funktionen $f(x) = kx + m$ är inte linjär om $m \neq 0$, för $f(0) = m \neq 0$. I själva verket är $f(x) = kx$ (rät linje genom origo) den enda envariabelfunktion som är linjär enligt sats 2.10. Men ändå säger vi ibland (felaktigt) att $f(x) = kx + m$ är en linjär funktion, därför att dess graf är en rät linje. Den korrekta benämningen är *affin funktion*, dvs linjär funktion plus konstant.

Övningar

2.1 Gauss eliminationsmetod

Övning 2.1 Lös ekvationssystemet med Gauss eliminationsmetod.

(a) $\begin{cases} 3x_1 + 2x_2 = 4 \\ 6x_1 + 5x_2 = 1 \end{cases}$ (b) $\begin{cases} x_1 + x_2 + x_3 = 2 \\ 2x_1 + 3x_3 = 3 \\ 3x_1 + x_2 + 2x_3 = 4 \end{cases}$ (c) $\begin{cases} 2x_1 + 3x_2 - 4x_3 = -2 \\ -6x_1 - 5x_2 = 4 \\ x_1 + 2x_2 + 4x_3 = 1 \end{cases}$

(d) $\begin{cases} 6x_1 + 3x_2 + 12x_3 = -3 \\ -2x_1 + x_2 - 20x_3 = 3 \\ 3x_1 + 2x_2 - 10x_3 = 2 \end{cases}$

Övning 2.2 Lös ekvationssystemet med Gauss eliminationsmetod.

(a) $\begin{cases} x_1 + 2x_2 + 3x_3 = 4 \\ 5x_1 + 5x_3 = 10 \\ 3x_1 + x_2 = -2 \end{cases}$ (b) $\begin{cases} -2x_1 + 4x_2 + 4x_3 = 1 \\ 5x_1 + 3x_2 - 7x_3 = \frac{5}{2} \\ 4x_1 + 5x_2 - x_3 = 1 \end{cases}$

(c) $\begin{cases} -7x_1 + 3x_2 + 14x_3 = 1 \\ 4x_1 - x_2 - 3x_3 = -1 \\ 8x_1 - 4x_2 + 4x_3 = 0 \end{cases}$ (d) $\begin{cases} -\frac{1}{2}x_1 + 2x_2 + 3x_3 = 1 \\ x_1 - 2x_2 + \frac{3}{2}x_3 + 6x_4 = 2 \\ -2x_1 + x_2 - 3x_3 + 3x_4 = -10 \\ 8x_1 - 10x_2 + 12x_3 - 10x_4 = 0 \end{cases}$

Övning 2.3 Lös ekvationssystemet med Gauss eliminationsmetod.

(a) $\begin{cases} 6x_1 - 4x_2 + 2x_3 = -3 \\ 3x_1 + x_2 - x_3 = 1 \\ 3x_1 - 5x_2 + 3x_3 = 0 \end{cases}$ (b) $\begin{cases} 3x_1 + 6x_2 - 9x_3 = 3 \\ x_1 + 3x_2 + 4x_3 = -3 \\ -2x_1 - 3x_2 + 13x_3 = -6 \end{cases}$

(c) $\begin{cases} 8x_1 - 2x_2 + 6x_3 = 4 \\ 6x_1 + 5x_2 - x_3 = 2 \\ x_1 + 3x_2 - 2x_3 = 2 \end{cases}$ (d) $\begin{cases} \frac{3}{2}x_1 - 3x_2 + \frac{9}{4}x_3 = 3 \\ \frac{11}{2}x_1 - \frac{5}{2}x_2 + \frac{3}{4}x_3 = 0 \\ 2x_1 + \frac{1}{4}x_2 - \frac{3}{4}x_3 = -\frac{7}{2} \end{cases}$

Övning 2.4 Lös ekvationssystemet med Gauss eliminationsmetod.

(a) $\begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ 2x_1 + 3x_2 + 4x_3 = 2 \\ 3x_1 + 4x_2 + 5x_3 = 3 \end{cases}$ (b) $\begin{cases} 6x_1 - \frac{7}{2}x_2 + 2x_3 = -\frac{3}{2} \\ 3x_1 - 4x_2 + 2x_3 = 1 \\ \frac{3}{2}x_1 + \frac{5}{2}x_2 - x_3 = -3 \end{cases}$

(c) $\begin{cases} 2x_1 - 5x_2 + 4x_3 = -\frac{3}{2} \\ 4x_1 - 10x_2 + 11x_3 = 1 \\ 12x_1 - 30x_2 + \frac{57}{2}x_3 = -3 \end{cases}$ (d) $\begin{cases} 2x_1 + 3x_2 + 7x_3 + 4x_4 = 2 \\ -2x_1 + 2x_2 - 6x_3 - 4x_4 = 0 \\ 3x_1 - 7x_2 + 2x_3 = -3 \\ 7x_1 - 14x_2 + x_3 - 4x_4 = 1 \end{cases}$

Övning 2.5 Lös ekvationssystemet med Gauss eliminationsmetod.

(a) $\begin{cases} -x_1 + 3x_2 + 2x_3 = 6 \\ 5x_1 - 3x_3 = -4 \\ \frac{5}{3}x_1 + x_2 - \frac{8}{15}x_3 = \frac{2}{5} \end{cases}$ (b) $\begin{cases} \frac{3}{5}x_1 + \frac{2}{7}x_2 + \frac{1}{5}x_3 = 1 \\ -\frac{3}{5}x_1 + \frac{2}{7}x_2 - \frac{2}{5}x_3 = 1 \\ \frac{8}{5}x_1 - \frac{2}{7}x_2 + \frac{7}{5}x_3 = 1 \end{cases}$

(c) $\begin{cases} -4x_1 + \frac{1}{2}x_2 + \frac{1}{3}x_3 = 2 \\ -9x_1 + 2x_2 + x_3 = \frac{2}{3} \\ \frac{2}{3}x_1 - \frac{1}{12}x_2 - \frac{1}{36}x_3 = \frac{1}{12} \end{cases}$ (d) $\begin{cases} 5x_1 + 6x_2 + 2x_3 + 7x_4 = 0 \\ 4x_1 + 6x_2 + 7x_3 + 2x_4 = 6 \\ 7x_1 + 7x_2 + 6x_3 + 6x_4 = 0 \\ 2x_1 + 9x_2 + 2x_3 + 7x_4 = 0 \end{cases}$

Övning 2.6 Lös ekvationssystemet med Gauss eliminationsmetod.

(a) $\begin{cases} x_1 - x_2 + 3x_3 + \frac{1}{2}x_4 = 4 \\ 4x_1 + 2x_2 + 10x_3 - 6x_4 = 3 \\ 2x_1 - 8x_2 + 8x_3 + 3x_4 = 2 \\ x_1 + x_3 = 6 \end{cases}$ (b) $\begin{cases} \frac{2}{3}x_1 + \frac{4}{5}x_2 - 2x_3 - 2x_4 = -1 \\ x_1 + \frac{7}{10}x_2 - x_3 + \frac{1}{2}x_4 = 2 \\ \frac{3}{2}x_1 + \frac{3}{5}x_2 + \frac{1}{10}x_3 + \frac{29}{10}x_4 = \frac{83}{20} \\ -2x_1 + \frac{1}{5}x_2 - \frac{12}{5}x_3 - \frac{5}{2}x_4 = 6 \end{cases}$

(c) $\begin{cases} -11x_1 + 3x_2 + 8x_3 - 9x_4 = 1 \\ -22x_1 + 3x_2 + 9x_3 - 7x_4 = 1 \\ \frac{11}{2}x_1 + 7x_2 + \frac{7}{3}x_3 + \frac{1}{3}x_4 = \frac{7}{3} \\ -33x_1 + 14x_2 + \frac{95}{3}x_3 + \frac{23}{3}x_4 = -\frac{1}{3} \end{cases}$

(d) $\begin{cases} 7x_1 + 2x_2 - 2x_3 + 14x_4 = 2 \\ -63x_1 - 8x_2 - 4x_3 - 68x_4 = -17 \\ 7x_1 - 37x_2 + \frac{14}{5}x_3 - \frac{11}{5}x_4 = \frac{11}{10} \\ 14x_1 + \frac{23}{2}x_2 - 25x_3 - \frac{11}{6}x_4 = \frac{59}{12} \end{cases}$

2.2 Vektorekvation och matrisekvation

Övning 2.7 Låt $u = \begin{bmatrix} 3 \\ -5 \\ 2 \end{bmatrix}$, $v = \begin{bmatrix} -2 \\ -1 \\ 4 \end{bmatrix}$. Beräkna uttrycket.

(a) $u + v$ (b) $-\frac{1}{2}v$ (c) $2u - 3v$ (d) $5u - 2(3u + v)$

Övning 2.8 Beräkna produkten.

(a) $\begin{bmatrix} 6 & -3 & 2 \\ 2 & 4 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 1 \\ -5 & 4 \\ 8 & 4 \\ -2 & -7 \end{bmatrix} \begin{bmatrix} 4 \\ -3 \end{bmatrix}$

(c) $\begin{bmatrix} 3 & -2 & -7 \\ 1 & -5 & -3 \\ 7 & -4 & 3 \\ 4 & -2 & -2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$ (d) $\begin{bmatrix} 9 & 7 & 4 & 2 \\ 3 & 2 & 6 & 1 \\ 7 & 4 & 2 & 2 \\ 5 & 3 & 1 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 5 \end{bmatrix}$

Övning 2.9 Beräkna produkten.

(a) $\begin{bmatrix} -2 & 3 & -3 \\ 9 & 0 & -1 \\ -1 & -3 & 2 \\ 5 & 2 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & -2 \\ 2 & 6 \\ 11 & -3 \\ 2 & -6 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 6 \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 9 & -13 & 2 \\ 4 & 6 & -9 & 17 \end{bmatrix} \begin{bmatrix} 3 \\ -6 \\ 4 \\ 5 \end{bmatrix}$ (d) $\begin{bmatrix} \frac{1}{3} & \frac{5}{2} & \frac{9}{7} & \frac{2}{7} \\ \frac{5}{6} & -\frac{6}{5} & -\frac{3}{7} & -\frac{3}{7} \\ \frac{11}{3} & 2 & \frac{3}{10} & \frac{1}{10} \end{bmatrix} \begin{bmatrix} 3 \\ 5 \\ 7 \\ 7 \end{bmatrix}$

Övning 2.10 Bestäm x så att $Ax = b$.

(a) $A = \begin{bmatrix} 2 & 3 & -6 \\ 1 & \frac{1}{2} & -2 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ (b) $A = \begin{bmatrix} 2 & 3 \\ 1 & \frac{1}{2} \\ 2 & 4 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}$

(c) $A = \begin{bmatrix} 2 & -2 & 3 & 4 \\ -6 & 3 & -6 & -2 \\ 4 & 2 & -11 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 1 \\ 4 \end{bmatrix}$

(d) $A = \begin{bmatrix} -2 & 3 & 1 \\ 4 & 2 & -6 \\ -8 & 10 & 7 \\ 10 & -7 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 2 \\ \frac{5}{2} \\ 0 \end{bmatrix}$

Övning 2.11 Undersök huruvida $b \in \text{span}\{a_1, a_2\}$, om $b = \begin{bmatrix} 6 \\ 15 \\ 3 \end{bmatrix}$.

(a) $a_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, a_2 = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$ (b) $a_1 = \begin{bmatrix} 3 \\ 10 \\ 1 \end{bmatrix}, a_2 = \begin{bmatrix} 0 \\ 5 \\ -1 \end{bmatrix}$

(c) $a_1 = \begin{bmatrix} -3 \\ -12 \\ -3 \end{bmatrix}, a_2 = \begin{bmatrix} 3 \\ -6 \\ -3 \end{bmatrix}$ (d) $a_1 = \begin{bmatrix} 6 \\ 15 \\ -3 \end{bmatrix}, a_2 = \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix}$

2.3 Lösningsmängden

Övning 2.12 Lös matrisekvationen $Ax = 0$. Finns icke-trivial lösning?

(a) $A = \begin{bmatrix} 2 & 2 & -3 \\ 5 & -2 & 2 \\ 3 & -4 & 5 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 3 & -5 \\ -2 & -1 & 3 \\ -1 & -2 & 2 \end{bmatrix}$ (c) $A = \begin{bmatrix} 2 & 9 & 3 \\ -2 & 0 & 3 \end{bmatrix}$

(d) $A = \begin{bmatrix} 8 & 3 \\ 4 & 4 \\ -2 & 3 \end{bmatrix}$

Övning 2.13 Lös matrisekvationen $Ax = b$ och visa hur lösningen x kan delas upp som $x_\text{h} + x_\text{p}$.

(a) $A = \begin{bmatrix} 2 & 2 & -3 \\ 5 & -2 & 2 \\ 3 & -4 & 5 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & 3 & -5 \\ -2 & -1 & 3 \\ -1 & -2 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ -2 \\ -1 \end{bmatrix}$

(c) $A = \begin{bmatrix} 2 & 9 & 3 \\ -2 & 0 & 3 \end{bmatrix}, \quad b = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$ (d) $A = \begin{bmatrix} 8 & 3 \\ 4 & 4 \\ -2 & 3 \end{bmatrix}, \quad b = \begin{bmatrix} -4 \\ -2 \\ 1 \end{bmatrix}$

Övning 2.14 Lös matrisekvationen och visa hur lösningen kan delas upp som $x_{\mathrm{h}} + x_{\mathrm{p}}$.

(a) $\begin{bmatrix} 1 & -3 & -2 & 4 \\ 2 & -1 & 4 & 1 \\ 3 & -4 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 1 \\ 1 \\ -4 \end{bmatrix}$

(b) $\begin{bmatrix} 2 & -3 & -1 & 3 \\ -4 & 3 & 4 & -9 \\ 5 & \frac{3}{2} & -\frac{5}{2} & \frac{3}{2} \\ 8 & 3 & 4 & -18 \end{bmatrix} x = \begin{bmatrix} -2 \\ -7 \\ 9 \\ -10 \end{bmatrix}$

(c) $\begin{bmatrix} 7 & 5 & -5 \\ -5 & 2 & 4 \\ 2 & 7 & -1 \\ 3 & -9 & -3 \end{bmatrix} x = \begin{bmatrix} 3 \\ -8 \\ -5 \\ 13 \end{bmatrix}$ (d) $\begin{bmatrix} 3 & -3 & 8 \\ 0 & 1 & 7 \\ 3 & -4 & 1 \\ -3 & 5 & 6 \end{bmatrix} x = \begin{bmatrix} 3 \\ 1 \\ 2 \\ -1 \end{bmatrix}$

Övning 2.15 Lös ekvationssystemet.

(a) $\begin{cases} 2x_1 + 4x_2 + 5x_3 - x_4 = 0 \\ -6x_1 - 7x_2 - 5x_3 + 4x_4 = 0 \end{cases}$ (b) $\begin{cases} -5x_1 + 3x_2 + 9x_3 + 4x_4 = 0 \\ 6x_1 + x_2 - 7x_3 - 2x_4 = 0 \end{cases}$

(c) $\begin{cases} -2x_1 + 2x_2 - 3x_3 - x_4 = 0 \\ 8x_1 - 7x_2 + 7x_3 + 2x_4 = 0 \\ 2x_1 - \frac{3}{2}x_2 + \frac{1}{2}x_3 = 0 \end{cases}$ (d) $\begin{cases} -x_1 + 4x_2 + 2x_3 + 5x_4 - x_5 = 0 \\ 4x_1 - 6x_2 + 2x_3 + 5x_4 + 9x_5 = 0 \end{cases}$

2.4 Linjärt oberoende

Övning 2.16 Utred huruvida mängden $\{u_i\}_{i=1}^{p}$ är linjärt oberoende.

(a) $u_1 = \begin{bmatrix} -1 \\ 4 \\ 2 \end{bmatrix}, u_2 = \begin{bmatrix} 2 \\ -6 \\ -4 \end{bmatrix}, u_3 = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}$

(b) $u_1 = \begin{bmatrix} 4 \\ 23 \\ 8 \end{bmatrix}, u_2 = \begin{bmatrix} -7 \\ -19 \\ -7 \end{bmatrix}, u_3 = \begin{bmatrix} 11 \\ 12 \\ -17 \end{bmatrix}, u_4 = \begin{bmatrix} 23 \\ -19 \\ 12 \end{bmatrix}$

(c) $u_1 = \begin{bmatrix} 4 \\ 2 \\ 11 \end{bmatrix}, u_2 = \begin{bmatrix} -9 \\ -4 \\ 1 \end{bmatrix}, u_3 = \begin{bmatrix} 5 \\ 2 \\ -12 \end{bmatrix}$

(d) $u_1 = \begin{bmatrix} 8 \\ 5 \\ 10 \end{bmatrix}, u_2 = \begin{bmatrix} -2 \\ -3 \\ 1 \end{bmatrix}, u_3 = \begin{bmatrix} 3 \\ -2 \\ 3 \end{bmatrix}$

Övning 2.17 Utred huruvida mängden $\{u_i\}_{i=1}^{p}$ är linjärt oberoende.

(a) $u_1 = \begin{bmatrix} 3 \\ 8 \\ -4 \end{bmatrix}, u_2 = \begin{bmatrix} -2 \\ -12 \\ -4 \end{bmatrix}$ (b) $u_1 = \begin{bmatrix} \frac{5}{2} \\ -\frac{3}{4} \\ 2 \end{bmatrix}, u_2 = \begin{bmatrix} -\frac{35}{12} \\ \frac{7}{8} \\ -\frac{7}{3} \end{bmatrix}$

(c) $u_1 = \begin{bmatrix} \frac{2}{3} \\ -3 \\ \frac{4}{3} \\ 0 \end{bmatrix}, u_2 = \begin{bmatrix} \frac{5}{3} \\ \frac{5}{2} \\ -\frac{7}{3} \\ -2 \end{bmatrix}, u_3 = \begin{bmatrix} \frac{1}{6} \\ -1 \\ 1 \\ 5 \end{bmatrix}$

(d) $u_1 = \begin{bmatrix} 3 \\ \frac{3}{2} \\ 11 \\ -7 \end{bmatrix}, u_2 = \begin{bmatrix} -2 \\ 3 \\ 3 \\ -5 \end{bmatrix}, u_3 = \begin{bmatrix} 7 \\ -\frac{9}{2} \\ 5 \\ 3 \end{bmatrix}$

Övning 2.18 Undersök om v_1 kan skrivas som en linjärkombination av v_2 och v_3.

(a) $v_1 = \begin{bmatrix} 2 \\ 5 \\ 4 \end{bmatrix}, v_2 = \begin{bmatrix} -3 \\ \frac{9}{2} \\ -4 \end{bmatrix}, v_3 = \begin{bmatrix} 11 \\ \frac{15}{2} \\ -3 \end{bmatrix}$

(b) $v_1 = \begin{bmatrix} 4 \\ -12 \\ 2 \end{bmatrix}, v_2 = \begin{bmatrix} -9 \\ 1 \\ 2 \end{bmatrix}, v_3 = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$

(c) $v_1 = \begin{bmatrix} 2 \\ -10 \\ 3 \\ 5 \end{bmatrix}, v_2 = \begin{bmatrix} -3 \\ 3 \\ -\frac{3}{2} \\ 1 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 3 \\ -\frac{1}{2} \\ -4 \end{bmatrix}$

(d) $v_1 = \begin{bmatrix} 4 \\ 5 \\ 2 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}, v_3 = \begin{bmatrix} -1 \\ -5 \\ 2 \end{bmatrix}$

Övning 2.19 Undersök om v_1 kan skrivas som en linjärkombination av v_2 och v_3.

(a) $v_1 = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}, v_2 = \begin{bmatrix} 6 \\ -6 \\ 3 \end{bmatrix}, v_3 = \begin{bmatrix} -2 \\ 7 \\ \frac{5}{2} \end{bmatrix}$

(b) $v_1 = \begin{bmatrix} 5 \\ 8 \\ 2 \end{bmatrix}, v_2 = \begin{bmatrix} 6 \\ 9 \\ -3 \end{bmatrix}, v_3 = \begin{bmatrix} -4 \\ 7 \\ 11 \end{bmatrix}$

(c) $v_1 = \begin{bmatrix} 1 \\ -6 \\ 9 \end{bmatrix}, v_2 = \begin{bmatrix} 4 \\ -12 \\ 2 \end{bmatrix}, v_3 = \begin{bmatrix} -2 \\ 6 \\ -1 \end{bmatrix}$

(d) $v_1 = \begin{bmatrix} 2 \\ 8 \\ 9 \\ 2 \end{bmatrix}, v_2 = \begin{bmatrix} -3 \\ -7 \\ 9 \\ \frac{3}{2} \end{bmatrix}, v_3 = \begin{bmatrix} 2 \\ 2 \\ -18 \\ -\frac{17}{5} \end{bmatrix}$

Övning 2.20 Bestäm c_1 och c_2 så att $v_3 = c_1 v_1 + c_2 v_2$.

(a) $v_1 = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}, v_2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, v_3 = \begin{bmatrix} 3 \\ 5 \\ 8 \end{bmatrix}$

(b) $v_1 = \begin{bmatrix} 4 \\ 2 \\ 7 \end{bmatrix}, v_2 = \begin{bmatrix} 2 \\ 2 \\ 6 \end{bmatrix}, v_3 = \begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix}$

(c) $v_1 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, v_2 = \begin{bmatrix} 3 \\ -6 \end{bmatrix}, v_3 = \begin{bmatrix} 4 \\ -8 \end{bmatrix}$

(d) $v_1 = \begin{bmatrix} 3 \\ -12 \\ -2 \end{bmatrix}, v_2 = \begin{bmatrix} -3 \\ 8 \\ 9 \end{bmatrix}, v_3 = \begin{bmatrix} 3 \\ -24 \\ 19 \end{bmatrix}$

2.5 Linjär funktion

Övning 2.21 Bestäm matrisen till den linjära funktionen $L\colon \mathbb{R}^n \to \mathbb{R}^m$.

(a) $L(x) = \begin{bmatrix} x_2 - x_1 \\ x_1 \end{bmatrix}$ (b) $L(x) = \begin{bmatrix} 2x_3 + x_1 \\ x_2 - x_3 \end{bmatrix}$

(c) $L(x) = x_1 - 5x_3$ (d) $L(x) = x$

Övning 2.22 Visa att funktionen inte är linjär.

(a) $f(x) = \begin{bmatrix} x_2 - x_1 \\ x_1 + 2 \end{bmatrix}$ (b) $f(x) = \begin{bmatrix} x_1 x_2 \\ x_1^2 \end{bmatrix}$ (c) $f(x) = \begin{bmatrix} |x_1| \\ \sin(x_2) \end{bmatrix}$

Övning 2.23 Den linjära funktionen $L\colon \mathbb{R}^2 \to \mathbb{R}^3$ avbildar $\begin{bmatrix}1\\2\end{bmatrix}$ på $\begin{bmatrix}1\\-2\\3\end{bmatrix}$ och $\begin{bmatrix}3\\1\end{bmatrix}$ på nollvektorn. Bestäm matrisen för L.

3. Matriser

Vi lär oss räkna med matriser.

3.1 Matrisalgebra

Vi har redan definierat begreppet matris och infört beteckningar för matriser i definition 2.2. Vi upprepar dem här och inför några fler beteckningar.

Definition 3.1 (Matrisbeteckningar)

1. $A \in \mathbb{R}^{m \times n}$ är en matris av typ $m \times n$.
2. Om $m = n$, så säger vi att $A \in \mathbb{R}^{n \times n}$ är en *kvadratisk matris*, annars är den en *rektangulär matris*.
3. $A = [a_{ij}]$, där $a_{ij} \in \mathbb{R}$ är matrisens element på plats (i, j), dvs i rad nummer i och kolonn nummer j.
4. $a_{ij} = (A)_{ij}$ är matrisens element på plats (i, j).
5. $a_{ii} = (A)_{ii}$ är matrisens diagonalelement.
6. $A = [a_1 \ \cdots \ a_n]$, där $a_k \in \mathbb{R}^m$ är matrisens kolonner.
7. Vi skriver för rektangulär matris

$$A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n} \tag{3.1}$$

och för kvadratisk matris

$$A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \in \mathbb{R}^{n \times n} \tag{3.2}$$

8. $O \in \mathbb{R}^{m \times n}$ eller $O_{m \times n}$ är *nollmatrisen* av typ $m \times n$ vars alla element är 0.
9. $I \in \mathbb{R}^{n \times n}$ eller I_n är *enhetsmatrisen* av typ $n \times n$ med elementen 1 på diagonalen och 0 utanför diagonalen, dvs

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}, \quad (I)_{ij} = \begin{cases} 1, & \text{om } i = j \\ 0, & \text{om } i \neq j \end{cases} \tag{3.3}$$

Vi skriver matriser med "hakparenteser" [ ]. I andra böcker är det vanligt att använda vanliga parenteser, dvs

$$A = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \tag{3.4}$$

En 1×1-matris är ett tal. Då skriver vi oftast inte ut matrisparenteserna, dvs $A = [a_{11}] = a_{11}$. Enhetsmatrisens element skrivs ibland med hjälp av *Kroneckers symbol* (Kroneckers delta):

$$\delta_{ij} = \begin{cases} 1, & \text{om } i = j \\ 0, & \text{om } i \neq j \end{cases} \tag{3.5}$$

dvs $I = [\delta_{ij}]$.

De linjära räkneoperationerna, dvs addition och multiplikation med skalär, kan nu definieras.

Definition 3.2 (Linjära räkneoperationer i $\mathbb{R}^{m\times n}$) Låt $A = [a_{ij}]$ och $B = [b_{ij}]$ vara matriser av typ $m \times n$ och α en skalär. Vi gör följande definitioner.

$$A = B \text{ om alla } a_{ij} = b_{ij} \qquad \text{(likhet)} \tag{3.6}$$

$$A + B = [a_{ij} + b_{ij}] \qquad \text{(addition)} \tag{3.7}$$

$$\alpha A = A\alpha = [\alpha a_{ij}] \qquad \text{(multiplikation med skalär)} \tag{3.8}$$

$$-A = [-a_{ij}] \qquad \text{(motsatt matris)} \tag{3.9}$$

Räkneoperationerna definieras alltså elementvis precis som för vektorer (kolonnmatriser) i definition 2.8. Observera att matriserna måste vara av samma typ. Vi kan nu bilda linjärkombination av matriser:

$$A = c_1 A_1 + \cdots + c_p A_p \tag{3.10}$$

Vi föredrar att skriva skalären före matrisen i (3.8), dvs αA.

Vi kan även definiera elementvis multiplikation: $A \diamond B = [a_{ij} b_{ij}]$. Detta kallas *Hadamardprodukt*[1]. Vi kommer inte att använda denna produkt, även om den är praktisk i vissa sammanhang, till exempel i datorprogrammering. Vi ska strax definiera en annan matrisprodukt.

De vanliga räknereglerna gäller enligt nästa sats.

Sats 3.1 (Räkneregler i $\mathbb{R}^{m\times n}$)

$$A + B = B + A \qquad \text{(kommutativ lag)} \tag{3.11}$$

$$(A + B) + C = A + (B + C) \qquad \text{(associativ lag)} \tag{3.12}$$

$$A + O = A \qquad \text{(addition av nollmatrisen)} \tag{3.13}$$

$$A + (-A) = O \qquad \text{(addition av motsatta matrisen)} \tag{3.14}$$

$$\alpha(A + B) = \alpha A + \alpha B \qquad \text{(distributiv lag)} \tag{3.15}$$

$$(\alpha + \beta)A = \alpha A + \beta A \qquad \text{(distributiv lag)} \tag{3.16}$$

$$(\alpha\beta)A = \alpha(\beta A) \qquad \text{(associativ lag)} \tag{3.17}$$

$$1A = A \tag{3.18}$$

Med ett mer rigoröst språkbruk innebär (3.13) att nollmatrisen är en *additiv enhet* och (3.14) att $-A$ är en *additiv invers*.

1. Den elementvisa produkten skrivs `A.*B` i MATLAB.

Bevis. Skriv på elementform och använd räkneregler för reella tal. □

Vi ska nu definiera produkten AB av två matriser A och B. Vi utgår från matris-vektor-produkten i definition 2.11. Låt $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times r}$ och $x \in \mathbb{R}^r$. Då är $Bx \in \mathbb{R}^n$ och $A(Bx) \in \mathbb{R}^m$. Vektorn Bx är en linjär kombination av kolonnerna i B:

$$Bx = x_1 b_1 + \cdots + x_r b_r = x_1 \begin{bmatrix} b_{11} \\ \vdots \\ b_{n1} \end{bmatrix} + \cdots + x_r \begin{bmatrix} b_{1r} \\ \vdots \\ b_{nr} \end{bmatrix} \tag{3.19}$$

Eftersom multiplikation med A är en linjär operation, får vi

$$A(Bx) = A(x_1 b_1 + \cdots + x_r b_r) = x_1(Ab_1) + \cdots + x_r(Ab_r) = \begin{bmatrix} Ab_1 \cdots Ab_r \end{bmatrix} x \tag{3.20}$$

Detta motiverar att vi definierar produktmatrisen AB som den matris som har kolonnerna Ab_j, så att $A(Bx) = (AB)x$.

Definition 3.3 (Matrisprodukt) Om $A \in \mathbb{R}^{m\times n}$ och $B \in \mathbb{R}^{n\times r}$, så definierar vi produkten $AB \in \mathbb{R}^{m\times r}$ enligt[a]

$$AB = [Ab_1 \cdots Ab_r] \tag{3.21}$$

a. Skrivs `A*B` i MATLAB.

Observera att produktmatrisen definieras endast om typerna stämmer överens enligt regeln

$$(m \times n) \cdot (n \times r) = m \times r \tag{3.22}$$

Exempel 3.1 (Matrismultiplikation) Enligt definitionen har vi

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \left[\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} \right] \tag{3.23}$$

$$= \left[\begin{bmatrix} 19 \\ 43 \end{bmatrix} \quad \begin{bmatrix} 22 \\ 50 \end{bmatrix} \right] = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix} \tag{3.24}$$

Matrismultiplikation utförs hellre enligt regeln i nästa sats.

Sats 3.2 (Rad-kolonn-regeln) Elementet på plats (i, j) i produktmatrisen ges av

$$(AB)_{ij} = \begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix} = \sum_{k=1}^{n} a_{ik} b_{kj} \tag{3.25}$$

dvs $\text{rad}_i(A)$ gånger $\text{kolonn}_j(B)$.

Bevis. Kolonn nummer j i AB är enligt (3.21) och (2.67)

$$Ab_j = \begin{bmatrix} a_1 \cdots a_n \end{bmatrix} \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix} = \begin{bmatrix} a_{11} \\ \vdots \\ a_{i1} \\ \vdots \\ a_{m1} \end{bmatrix} b_{1j} + \cdots + \begin{bmatrix} a_{1k} \\ \vdots \\ a_{ik} \\ \vdots \\ a_{mk} \end{bmatrix} b_{kj} + \cdots + \begin{bmatrix} a_{1n} \\ \vdots \\ a_{in} \\ \vdots \\ a_{mn} \end{bmatrix} b_{nj} \tag{3.26}$$

Här är rad nummer i lika med

$$a_{i1}b_{1j} + \cdots + a_{ik}b_{kj} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj} \tag{3.27}$$

Detta är elementet på plats (i, j) i AB. Å andra sidan är produktmatrisen

$$\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix} \tag{3.28}$$

av typ $(1 \times n) \cdot (n \times 1) = 1$, dvs ett tal, och lika med

$$a_{i1}b_{1j} + \cdots + a_{ik}b_{kj} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj} \tag{3.29}$$

dvs (3.27) kan skrivas som $\text{rad}_i(A)$ gånger $\text{kolonn}_j(B)$. □

Exempel 3.2 (Rad-kolonn-regeln) Vi beräknar

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} & \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} \\ \\ \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} & \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} \end{bmatrix} \tag{3.30}$$

$$= \begin{bmatrix} 1 \cdot 5 + 2 \cdot 7 & 1 \cdot 6 + 2 \cdot 8 \\ 3 \cdot 5 + 4 \cdot 7 & 3 \cdot 6 + 4 \cdot 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix} \tag{3.31}$$

Sats 3.3 (Räkneregler för matrisprodukt) Låt A, B, C vara matriser sådana att följande matrisprodukter är definierade och låt α vara en skalär. Då gäller

$$A(BC) = (AB)C \quad \text{(associativ lag)} \tag{3.32}$$

$$A(B + C) = AB + AC \quad \text{(distributiv lag)} \tag{3.33}$$

$$(B + C)A = BA + CA \quad \text{(distributiv lag)} \tag{3.34}$$

$$\alpha(AB) = (\alpha A)B = A(\alpha B) \quad \text{(distributiv lag)} \tag{3.35}$$

$$I_m A = A,\ AI_n = A, \quad A \in \mathbb{R}^{m\times n} \quad \text{(mult. med enhetsmatris)} \tag{3.36}$$

$$O_{p\times m}A = O_{p\times n},\ AO_{n\times p} = O_{m\times p},\ A \in \mathbb{R}^{m\times n} \quad \text{(mult. med nollmatris)} \tag{3.37}$$

Regeln (3.36) innebär att enhetsmatrisen I_m respektive I_n fungerar som *multiplikativ enhet* vid multiplikation från vänster respektive höger. Notera att vi behöver använda enhetsmatriser av olika storlek om A är rektangulär.

Bevis. Vi utelämnar beviset. □

Observera att vi har ingen kommutativ lag här. I själva verket är $AB \neq BA$ i allmänhet, dvs produkten är ej kommutativ. Om $AB = BA$, så säger vi att A och B *kommuterar*. I så fall måste $A, B \in \mathbb{R}^{n\times n}$ vara kvadratiska.

Associativa lagen (3.32) innebär att vi kan skriva

$$ABC = A(BC) = (AB)C \tag{3.38}$$

utan parenteser. Exempelvis skriver vi $A^3 = AAA = A(A(A))$ om $A \in \mathbb{R}^{n\times n}$ är kvadratisk. Mer allmänt kan vi bilda *matrispotensen* $A^k = A\cdots A$ (k faktorer) för kvadratisk matris A.

Likaså innebär (3.35) att vi kan ta bort parenteser och skriva till exempel

$$\alpha Ax = \alpha(Ax) = (\alpha A)x \tag{3.39}$$

Exempel 3.3 (Icke-kommutativa matriser)

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}, \ B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \tag{3.40}$$

Vi får

$$AB = \begin{bmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \\ 29 & 40 & 51 \end{bmatrix}, \ BA = \begin{bmatrix} 22 & 28 \\ 49 & 64 \end{bmatrix} \tag{3.41}$$

Här är AB av typ 3×3 och BA av typ 2×2, dvs de är inte ens av samma typ.

Exempel 3.4 (Icke-kommutativa matriser) Med matriserna

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \ B = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} \tag{3.42}$$

får vi

$$AB = \begin{bmatrix} 9 & 12 \\ 19 & 26 \end{bmatrix} \neq \begin{bmatrix} 7 & 10 \\ 19 & 28 \end{bmatrix} = BA \tag{3.43}$$

Här har AB och BA samma typ men de är inte lika, dvs A och B kommuterar ej.

Definition 3.4 (Transponat) Låt $A \in \mathbb{R}^{m\times n}$. *Transponatet* av A är matrisen $A^T \in \mathbb{R}^{n\times m}$ som ges av

$$(A^T)_{ij} = (A)_{ji} \tag{3.44}$$

Att transponera en matris innebär alltså att låta rader och kolonner byta plats, dvs $\text{kolonn}_j(A^T) = \text{rad}_j(A)$.

För att spara plats på textraden skriver vi ibland en kolonnvektor som transponatet av en radvektor, dvs $v = [v_1 \ \cdots \ v_n]^T$ istället för $v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$.

Exempel 3.5 (Transponat)

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}, \ A^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} \tag{3.45}$$

Sats 3.4 (Räkneregler för transponat)

$$(A^T)^T = A \tag{3.46}$$

$$(A+B)^T = A^T + B^T \tag{3.47}$$

$$(\alpha A)^T = \alpha(A^T) \tag{3.48}$$

$$(AB)^T = B^T A^T \tag{3.49}$$

Reglerna (3.47) och (3.48) innebär att transponering är en linjär operation.

Bevis. Vi bevisar (3.49). Enligt (3.44) och rad-kolonn-regeln (3.25) har vi

$$\left((AB)^T\right)_{ij} = (AB)_{ji} = \sum_{k=1}^{n} (A)_{jk}(B)_{ki} = \sum_{k=1}^{n} (B^T)_{ik}(A^T)_{kj} = (B^T A^T)_{ij} \tag{3.50}$$

Detta är (3.49). $\square$

En matris A kallas *symmetrisk* om $A^T = A$. En *antisymmetrisk* matris B uppfyller $B^T = -B$. En matris som är symmetrisk eller antisymmetrisk måste vara kvadratisk. Som exempel har vi

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix} \quad \text{respektive} \quad B = \begin{bmatrix} 0 & 2 & 3 \\ -2 & 0 & 5 \\ -3 & -5 & 0 \end{bmatrix} \tag{3.51}$$

Se problem 3.6.

3.2 Invers matris

Vi erinrar oss att $-A$ fungerar som *additiv invers*:

$$A + (-A) = (-A) + A = O \tag{3.52}$$

dvs addition av $-A$ upphäver addition med A. På liknande vis är $a^{-1} = 1/a$ en *multiplikativ invers* vid multiplikation av tal:

$$a \cdot a^{-1} = a^{-1} \cdot a = 1 \tag{3.53}$$

dvs multiplikation med a^{-1} upphäver multiplikation med a. Obs att a^{-1} existerar endast om $a \neq 0$. På liknande sätt definierar vi multiplikativ invers för matrismultiplikation.

Definition 3.5 (Invers matris) En kvadratisk matris $A \in \mathbb{R}^{n\times n}$ är *inverterbar* (eller *icke-singulär*) om det finns en matris $B \in \mathbb{R}^{n\times n}$ sådan att

$$AB = I,\ BA = I \tag{3.54}$$

I så fall är B unik, dvs det finns ingen annan matris som uppfyller (3.54), se problem 3.7. Vi säger då att B är *inversen* eller *inversa matrisen* till A. Inversa matrisen skrivs A^{-1}. Alltså har vi

$$AA^{-1} = A^{-1}A = I \tag{3.55}$$

Matrisen är *singulär* om den den inte är inverterbar.

Beteckningen A^{-1} uttalas "A invers", inte "A upphöjt till -1". Vi talar sällan om "matris-division". Det är opraktiskt eftersom vi måste skilja mellan division av B med A från vänster, $A^{-1}B$, och från höger, BA^{-1}; matrismultiplikation är ju inte kommutativ.[2]

Den inversa funktionen till den linjära funktionen $L(x) = Ax$ ges av multiplikation med A^{-1}, dvs $L^{-1}(x) = A^{-1}x$, om A är inverterbar. Det följer av att

$$L(L^{-1}(x)) = A(A^{-1}x) = (AA^{-1})x = Ix = x \tag{3.56}$$

$$L^{-1}(L(x)) = A^{-1}(Ax) = (A^{-1}A)x = Ix = x \tag{3.57}$$

2. Matlab har dock vänster- och högerdivision.

Vi frågar oss nu om det finns några inverterbara matriser och hur vi i så fall beräknar inversen. I fallet $n = 1$ har vi $A = a \in \mathbb{R}$ och $A^{-1} = 1/a$ om och endast om $A = a \neq 0$. Fallet $n = 2$ är också enkelt enligt nästa sats.

Sats 3.5 (Inversen av 2×2-matris) Låt $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ vara en 2×2-matris. Om $ad-bc \neq 0$, så är A inverterbar med

$$A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{3.58}$$

Om $ad - bc = 0$, så är A singulär.

Observera att uttrycket $ad-bc$ är inget annat än determinanten av matrisen A, se (1.105). Vi återkommer till detta i avsnitt 3.3.

Bevis. Vi multiplicerar matriserna på båda hållen:

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = (ad-bc)I \tag{3.59}$$

Vi ser att om $ad - bc \neq 0$, så är A inverterbar med den angivna inversen. Detta visar bara att villkoret $ad - bc \neq 0$ är tillräckligt för inverterbarhet. För att visa att det också är nödvändigt hänvisar vi till metoden i sats 3.8, se problem 3.8. □

Sats 3.6 (Egenskaper hos matrisinversen)

1. Om A är inverterbar, så är A^{-1} inverterbar med inversen $(A^{-1})^{-1} = A$.
2. Om A, B är inverterbara, så är AB inverterbar med inversen $(AB)^{-1} = B^{-1}A^{-1}$.
3. Om A är inverterbar, så är A^T inverterbar med inversen $(A^T)^{-1} = (A^{-1})^T$.

Bevis. Vi bevisar 2. Vi kollar att uppfyller $B^{-1}A^{-1}$ definitionen av invers i (3.54). Associativa lagen (3.32) ger

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I \tag{3.60}$$

och

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I \tag{3.61}$$

Bevisen av de övriga lämnas till läsaren, se problem 3.10. □

Hur beräknar vi inversa matrisen A^{-1} om den finns? Svaret är att vi radreducerar totalmatrisen som består av A utvidgad med enhetsmatrisen, dvs $[A\ I] \sim [I\ A^{-1}]$. För att bevisa detta ska vi skriva radoperationerna på matrisform. Vi presenterar detta först för 3×3-matriser.

Exempel 3.6 (Elementära radoperationer på matrisform) Låt

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \in \mathbb{R}^{3\times 3} \tag{3.62}$$

Som exempel på den första typen av radoperation i definition 2.3 tar vi

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \alpha & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \alpha \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{3.63}$$

Den första matrisen till höger är enhetsmatrisen. Multiplikation med den andra matrisen plockar ut rad 2 och lägger in den som rad 3 (kolla detta). Alltså blir multiplikation med E_1:

$$E_1 A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} + \alpha \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \tag{3.64}$$

dvs addera α gånger rad 2 till rad 3. Dess invers är att subtrahera α gånger rad 2 från rad 3, dvs

$$(E_1)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\alpha & 1 \end{bmatrix} \tag{3.65}$$

Multiplikation med matrisen

$$E_2 = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.66}$$

innebär att rad 1 multipliceras med faktorn $\alpha \neq 0$ (skalning). Dess invers är

$$(E_2)^{-1} = \begin{bmatrix} \alpha^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.67}$$

Multiplikation med matrisen

$$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{3.68}$$

innebär att rad 2 och 3 byter plats (permutation) (kolla detta). Den är sin egen invers $(E_3)^{-1} = E_3$.

Dessa tre slags matriser kallas *elementära matriser* och motsvarar våra tre slags elementära radoperationer i definition 2.3. Det klart hur de generaliseras till matriser av typ $n \times n$.

Sats 3.7 (Inverterbarhet av elementära matriser) Elementära matriser är inverterbara.

Bevis. Elementära radoperationer är inverterbara. Se även inverserna i exempel 3.6. □

Sats 3.8 (Invers matris kan beräknas med radreducering) Matrisen $A \in \mathbb{R}^{n\times n}$ är inverterbar om och endast om den är radekvivalent med enhetsmatrisen I. De radoperationer som reducerar A till I transformerar I till A^{-1}, dvs $[A\ I] \sim [I\ A^{-1}]$.

Bevis. 1. Antag att A är inverterbar. Då är ekvationen $Ax = b$ lösbar för varje b enligt sats 3.9 och därför har A pivotpositioner i varje kolonn enligt sats 2.6. Men detta innebär att $A \sim I$.

2. Antag att $A \sim I$, dvs

$$E_p \cdots E_2 E_1 A = I \tag{3.69}$$

för några elementära matriser. Låt $E = E_p \cdots E_1$, så att $EA = I$. Detta är "hälften" av definitionen av invers i (3.54) och vi måste visa att även $AE = I$. Men de elementära matriserna är inverterbara och enligt punkt 2 i sats 3.6 följer att E är inverterbar med inversen

$$E^{-1} = (E_p \cdots E_1)^{-1} = E_1^{-1} \cdots E_p^{-1} \tag{3.70}$$

Genom att multiplicera $EA = I$ med E^{-1} från vänster och med E från höger får vi nu

$$E^{-1}EAE = E^{-1}IE, \quad \text{dvs} \quad AE = I \tag{3.71}$$

och därmed är A inverterbar med inversen

$$A^{-1} = E = E_p \cdots E_1 \tag{3.72}$$

Genom att multiplicera den utvidgade matrisen $[A\ I]$ med E får vi

$$E[A\ I] = [EA\ EI] = [I\ A^{-1}], \quad \text{dvs} \quad [A\ I] \sim [I\ A^{-1}] \tag{3.73}$$

Det betyder att de radoperationer som reducerar A till I transformerar också I till A^{-1}. □

Ruta 3.1 (Metod för beräkning av invers matris med radreducering)
Radreducera $[A\ I]$. Detta ger antingen $[I\ A^{-1}]$ eller visar att A är singulär.

Exempel 3.7 (Beräkna invers) Med $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ använder vi elementära matrisen $E = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ (addera -1 gånger rad 1 till rad 2) och vi får

$$E \begin{bmatrix} A & I \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} \tag{3.74}$$

Vi ser att $A \sim \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ har bara 1 pivotposition och A är därför singulär. Som en jämförelse ser vi att villkoret i sats 3.5 ger $ad - bc = 0$, dvs singulär matris.

Exempel 3.8 (Beräkna invers) Vi tar ett lätt exempel till. Vi gör beräkningarna på samma sätt som i beviset av sats 3.8.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{matrix} -3 \\ + \end{matrix} \sim \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix} \begin{matrix} + \\ 1 \end{matrix} \;\Big|\; -\tfrac{1}{2} \sim \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \tag{3.75}$$

Detta motsvarar följande elementära matriser:

$$E_1 = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}, \; E_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \; E_3 = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \tag{3.76}$$

Ovanstående räkningar blir på matrisform (genomför beräkningarna!)

$$E_3 E_2 E_1 A = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \tag{3.77}$$

Det betyder att A är inverterbar med inversen

$$A^{-1} = E_3 E_2 E_1 = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} \tag{3.78}$$

Sats 3.5 ger förstås samma resultat:

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{-2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} \tag{3.79}$$

Exempel 3.9 (Beräkna invers) Vi tar en större matris.

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & -1 & -1 \end{bmatrix} \tag{3.80}$$

$$\sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I \tag{3.81}$$

Matrisen A är alltså inverterbar. Vi gör nu samma operationer på matrisen I, dvs multiplicerar med motsvarande elementära matriser (i rätt ordning):

$$A^{-1} = E_4E_3E_2E_1 = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix} \tag{3.82}$$

$$= \frac{1}{3} \begin{bmatrix} 7 & 2 & 2 \\ -17 & -1 & -4 \\ 2 & 1 & 1 \end{bmatrix} \tag{3.83}$$

För att spara plats har vi här slagit ihop de två första och de två sista radoperationerna i en matris vardera, dvs E_1 och E_4. Vi inser nu att här behövs datorberäkning; att multiplicera flera 3×3-matriser blir nästan ogörligt med penna och papper.

Sats 3.9 (Linjärt ekvationssystem löses med invers matris) Om $A \in \mathbb{R}^{n\times n}$ är inverterbar, så har ekvationen $Ax = b$ en *unik* lösning för *varje* $b \in \mathbb{R}^n$. Lösningen ges av $x = A^{-1}b$.

Bevis. Låt $b \in \mathbb{R}^n$ och sätt $x = A^{-1}b$. Då fås

$$Ax = A(A^{-1}b) = (AA^{-1})b = Ib = b \tag{3.84}$$

Alltså har vi minst en lösning $x = A^{-1}b$ för varje $b \in \mathbb{R}^n$.

Antag att det finns en annan lösning y, dvs $Ay = b$. Genom att multiplicera denna ekvation med A^{-1} från vänster, så får vi

$$Ay = b, \quad A^{-1}(Ay) = A^{-1}b, \quad (A^{-1}A)y = A^{-1}b, \quad Iy = x \tag{3.85}$$

dvs $y = x$. Alltså är lösningen unik. $\square$

Invers matris definieras endast för kvadratisk matris, dvs av typ $n \times n$. Om inversen existerar, så får vi en formel för lösningen till det linjära ekvationssystemet, dvs $x = A^{-1}b$. Detta är viktigt när vi utvecklar teori och bevisar satser. Men vi har sett att det är tidskrävande att beräkna matrisinversen redan för måttligt stora matriser ($n = 3, 4$), än värre för stora matriser ($n = 10^k$ med $k = 1, 2, 3, 4$ eller mer förekommer i tekniska beräkningar). Därför använder vi inte lösningsformeln $x = A^{-1}b$ i praktiska beräkningar. I stället löser vi linjära ekvationssystem genom att radreducera $[A\ b]$. Andra beräkningsmetoder kommer att presenteras i kapitel 6.

3.3 Determinant

Vi ska nu definiera determinanten $\det(A)$ för kvadratisk matris $A \in \mathbb{R}^{n \times n}$. Vi har redan använt tvåradig och treradig determinant ($n = 2, 3$) i samband med kryssprodukt, se (1.105) och (1.106). Vi har också sett att den tvåradiga determinanten förekommer i villkoret för inverterbarhet i sats 3.5. Vi upprepar dessa här[3].

För $n = 1$:

$$A = [a_{11}], \quad \det(A) = a_{11} \tag{3.86}$$

För $n = 2$:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \tag{3.87}$$

För $n = 3$:

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \tag{3.88}$$

$$= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \tag{3.89}$$

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \tag{3.90}$$

Observera att för $n = 3$ har vi

$$\det(A) = a_{11} \det(A_{11}) - a_{12} \det(A_{12}) + a_{13} \det(A_{13}) \tag{3.91}$$

där $A_{ij} \in \mathbb{R}^{2 \times 2}$ betecknar den delmatris av A där vi strukit rad i och kolonn j. Även fallet $n = 2$ har denna form:

$$\det(A) = a_{11} \det(A_{11}) - a_{12} \det(A_{12}) = a_{11}a_{22} - a_{12}a_{21} \tag{3.92}$$

för delmatriserna är här av typ 1×1, nämligen $A_{11} = [a_{22}]$, $A_{12} = [a_{21}]$. Vi säger att vi utvecklar här determinanten efter rad 1 som en summa av underdeterminanter $\det(A_{ij})$ av

3. Strecken i dessa formler kallas determinantstreck och ska inte förväxlas med absolutbelopp.

storlek 1 mindre än $\det(A)$. På samma vis kan vi utveckla $\det(A_{ij})$ i mindre underdeterminanter. Detta leder till följande rekursiva definition.

Definition 3.6 (Determinant) Låt $A \in \mathbb{R}^{n\times n}$ vara en kvadratisk matris och låt $A_{ij} \in \mathbb{R}^{(n-1)\times(n-1)}$ beteckna den delmatris av A där vi strukit rad i och kolonn j. Determinanten $\det(A)$ definieras rekursivt på följande vis.

För $n = 1$:

$$\det(A) = \det([a_{11}]) = a_{11} \tag{3.93}$$

För $n \geqslant 2$:

$$\begin{aligned} \det(A) &= a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + \cdots + (-1)^{1+n} a_{1n}\det(A_{1n}) && (3.94)\\ &= \sum_{j=1}^{n} (-1)^{1+j} a_{1j}\det(A_{1j}) && (3.95) \end{aligned}$$

I (3.95) utvecklar vi alltså determinanten efter rad 1. Nästa sats säger att vi kan utveckla determinanten efter vilken rad eller kolonn som helst.

Sats 3.10 (Determinanten kan utvecklas efter godtycklig rad eller kolonn) För fixt i respektive fixt j har vi

$$\begin{aligned} \det(A) &= \sum_{j=1}^{n} (-1)^{i+j} a_{ij}\det(A_{ij}) && (3.96)\\ &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij}\det(A_{ij}) && (3.97) \end{aligned}$$

Formel (3.96) innebär utveckling efter rad i medan (3.97) är utveckling efter kolonn j.

Bevis. Beviset är ganska krångligt. Vi utelämnar det. $\square$

Observera att termerna i utvecklingen har alternerande plus- och minustecken enligt schemat

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ \vdots & & & & \end{bmatrix}, \quad \text{dvs } \left[(-1)^{i+j}\right] \tag{3.98}$$

Exempel 3.10 (Utveckla determinant) Vi beräknar följande determinant genom att utveckla efter kolonn nummer 3 (för att dra fördel av nollorna):

$$\begin{vmatrix} 1 & 5 & 0 \\ 2 & 4 & 3 \\ 0 & -2 & 0 \end{vmatrix} = a_{13}\det(A_{13}) + a_{23}(-1)\det(A_{23}) + a_{33}\det(A_{33}) \tag{3.99}$$

$$= 0\begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix} + 3(-1)\begin{vmatrix} 1 & 5 \\ 0 & -2 \end{vmatrix} + 0\begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix} = 6 \tag{3.100}$$

Det är krångligt att beräkna determinant i allmänhet, men om matrisen är triangulär är det lätt.

Definition 3.7 (Triangulär matris) En kvadratisk matris $A \in \mathbb{R}^{n\times n}$ är *uppåt triangulär* eller *övertriangulär* om $a_{ij} = 0$ för $i > j$, dvs alla element under diagonalen är noll:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix} \tag{3.101}$$

Matrisen kallas *nedåt triangulär* eller *undertriangulär* om $a_{ij} = 0$ för $i < j$.

Sats 3.11 (Determinanten av triangulär matris) Determinanten av en triangulär matris är produkten av diagonalelementen, dvs

$$\det(A) = a_{11}a_{22}\cdots a_{nn} = \prod_{i=1}^{n} a_{ii} \tag{3.102}$$

Bevis. Antag att matrisen är övertriangulär. Vi utvecklar efter första kolonnen eftersom den innehåller många nollor. Vi får

$$\det(A) = a_{11}\det(A_{11}) = a_{11}\Big(a_{22}\det((A_{11})_{11})\Big) = \cdots = a_{11}a_{22}\cdots a_{nn} \tag{3.103}$$

där vi utnyttjat att delmatrisen A_{11} också är övertriangulär och motsvarande underdeterminant utvecklats efter första kolonnen och så vidare. Undertriangulär determinant behandlas på liknande vis. $\square$

Sats 3.12 (Egenskaper hos determinant) Låt $A, B \in \mathbb{R}^{n \times n}$ vara kvadratiska matriser. Då har vi

$$\det(A^T) = \det(A) \tag{3.104}$$

$$\det(AB) = \det(A)\det(B) \tag{3.105}$$

Bevis. Vi utelämnar beviset. □

För att ta fram en systematisk metod för beräkning av determinant återvänder vi till radreducering. Vi erinrar oss hur vi skriver elementära radoperationer på matrisform från exempel 3.6.

Exempel 3.11 (Elementära matriser igen) Ett exempel på första typen av radoperation ges av matrisen

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \alpha & 1 \end{bmatrix} \tag{3.106}$$

dvs addera α gånger rad 2 till rad 3. Den är undertriangulär med determinanten

$$\det(E_1) = 1 \tag{3.107}$$

Skalning av rad 1 med faktorn $\alpha \neq 0$ ges av

$$E_2 = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.108}$$

med $\det(E_2) = \alpha$.
Som exempel på permutation av två rader tar vi

$$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{3.109}$$

dvs rad 2 och 3 byter plats. Determinanten fås genom utveckling efter rad 1:

$$\det(E_3) = 1 \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1 \tag{3.110}$$

Dessa resultat generaliseras enkelt till elementära matriser av typ $n \times n$.

Tillsammans med produktregeln i (3.105) visar detta att determinanten ändras nästan inte under elementära radoperationer. Vi sammanfattar i nästa sats.

Sats 3.13 (Determinant under elementära radoperationer) Om E är matrisen för en elementär radoperation så gäller

$$\det(EA) = \det(E)\det(A) = c\det(A) \tag{3.111}$$

där c beror på typ av radoperation:

$$c = \begin{cases} 1 & \text{addera multipel av en rad till en annan rad} \\ -1 & \text{permutera två rader} \\ \alpha & \text{skala en rad med faktorn } \alpha \end{cases} \tag{3.112}$$

På grund av transponeringsregeln i (3.104) gäller samma sak för *elementära kolonnoperationer*.

Vi formulerar nu en metod för beräkning av determinant. Vi kan alltid transformera en kvadratisk matris $A \in \mathbb{R}^{n\times n}$ till trappstegsform U med elementära radoperationer utan skalning:

$$U = E_p \cdots E_1 A \tag{3.113}$$

så att

$$\det(U) = \det(E_p)\cdots\det(E_1)\det(A) = (-1)^k\det(A) \tag{3.114}$$

där k är antalet permutationer. Eftersom U är uppåt triangulär så får vi

$$\det(A) = (-1)^k\det(U) = (-1)^k u_{11}\cdots u_{nn} \tag{3.115}$$

Ruta 3.2 (Metod för beräkning av determinant med radreducering)
Transformera A till trappstegsform U med elementära radoperationer utan skalning. Då fås

$$\det(A) = (-1)^k u_{11}\cdots u_{nn} \tag{3.116}$$

där k är antalet permutationer som använts.

Detta ger också ett villkor för inverterbarhet, som generaliserar villkoret i sats 3.5.

Sats 3.14 (Villkor för inverterbarhet) En kvadratisk matris $A \in \mathbb{R}^{n\times n}$ är inverterbar om och endast om $\det(A) \neq 0$.

Bevis. Vi transformerar A till trappstegsform U med elementära radoperationer utan skalning, så att (3.116) gäller. Vi vet att A inverterbar om och endast om den har pivotpositioner i varje rad, dvs om och endast om alla $u_{ii} \neq 0$. Detta bevisar påståendet. □

Exempel 3.12 (Beräkning av determinant) Vi använder metoden i ruta 3.2:

$$\begin{vmatrix}0&1&1\\3&4&5\\6&7&8\end{vmatrix}\begin{matrix}\leftarrow\\ \leftarrow\\ \end{matrix} = (-1)\begin{vmatrix}3&4&5\\0&1&1\\6&7&8\end{vmatrix}\begin{matrix}-2\\ \\ \leftarrow +\end{matrix} = (-1)\begin{vmatrix}3&4&5\\0&1&1\\0&-1&-2\end{vmatrix}\begin{matrix}\\ 1\\ \leftarrow +\end{matrix} \tag{3.117}$$

$$= (-1)\begin{vmatrix}3&4&5\\0&1&1\\0&0&-1\end{vmatrix} = (-1)\cdot 3\cdot 1\cdot(-1) = 3 \tag{3.118}$$

Vi gjorde en permutation, vilket ger en faktor (-1).

Vi kan även använda skalning. Formeln i (3.111) ger

$$\det(A) = \frac{1}{c}\det(EA) \tag{3.119}$$

Vi fortsätter med skalning och får

$$(-1)\begin{vmatrix}3&4&5\\0&1&1\\0&0&-1\end{vmatrix}\begin{matrix}\mid \frac{1}{3}\\ \\ \mid -1\end{matrix} = (-1)\cdot 3\cdot(-1)\begin{vmatrix}1&\frac{4}{3}&\frac{5}{3}\\0&1&1\\0&0&1\end{vmatrix} = 3 \tag{3.120}$$

Vi kan se det som att vi har brutit ut 3 ur rad 1 och -1 ur rad 3.

Eftersom determinanten inte är lika med noll drar vi slutsatsen att matrisen är inverterbar. För att beräkna inversen skulle vi behöva fortsätta och transformera till reducerad trappstegsform enligt metoden i ruta 3.1.

Övningar

3.1 Matrisalgebra

Övning 3.1 Låt $A = \begin{bmatrix} 2 & 0 & 1 \\ 4 & 2 & -2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 4 & 7 \\ 0 & 3 & 1 \end{bmatrix}$, $C = \begin{bmatrix} 4 & 8 \\ 0 & 4 \\ 2 & -4 \end{bmatrix}$. Beräkna

(a) $A + 2B$ (b) $\frac{1}{2}A$ (c) $C - 2A^T$ (d) $B^T + \frac{1}{2}C$

Övning 3.2 Med A, B, C som i övning 3.1 lös ekvationen $2X^T + C = 2B^T$.

Övning 3.3 Beräkna $u^T v$ och uv^T. (Se problem 3.3.)

(a) $u = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $v = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$ (b) $u = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$, $v = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{bmatrix}$ (c) $u = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$, $v = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}$

Övning 3.4 Beräkna AB och BA.

(a) $A = \begin{bmatrix} -1 & 2 \\ 5 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$ (b) $A = \begin{bmatrix} -1 & 2 \\ 5 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 4 & 5 \\ 6 & 7 & 8 \end{bmatrix}$

Övning 3.5 Vilka matriser kommuterar med A?

(a) $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}$ (c) $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ (d) $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 1 & 2 \end{bmatrix}$

3.2 Invers matris

Övning 3.6 Beräkna A^{-1}. Använd både sats 3.5 och sats 3.8.
(a) $A = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$ (b) $A = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$ (c) $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (d) $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$

Övning 3.7 Beräkna A^{-1}.
(a) $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 3 & 2 \\ -1 & 0 & 2 \\ 2 & 1 & 1 \end{bmatrix}$
(c) $A = \begin{bmatrix} 1 & 3 & 2 \\ 2 & 6 & 4 \\ 2 & 1 & 1 \end{bmatrix}$ (d) $A = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix}$

Övning 3.8 Beräkna A^{-1} och använd den för att lösa ekvationen $Ax = b$.
$A = \begin{bmatrix} 2 & 2 & -1 \\ 1 & -1 & -2 \\ 1 & 0 & -1 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

Övning 3.9 Låt $A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 4 & 2 \\ 0 & -1 & 3 \\ 1 & 1 & 1 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 1 & 2 \\ -1 & 3 & 1 \end{bmatrix}$.
Lös matrisekvationen.
(a) $AXB = C$ (b) $AXB = XB + C$

3.3 Determinant

Övning 3.10 Beräkna $\det(A)$.
(a) $A = \begin{bmatrix} 5 & -3 \\ 4 & 0 \end{bmatrix}$ (b) $A = \begin{bmatrix} \cos(x) & -\sin(x) \\ \sin(x) & \cos(x) \end{bmatrix}$
(c) $A = \begin{bmatrix} -2 & 1 & 3 \\ 1 & 0 & 2 \\ -1 & 5 & 3 \end{bmatrix}$ (d) $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 7 & 5 & 9 & 2 \\ 9 & 3 & 5 & 9 \end{bmatrix}$

Övning 3.11 Beräkna $\det(A)$.
(a) $A = \begin{bmatrix} 1 & 5 & 8 & 10 \\ 0 & 2 & 6 & 9 \\ 0 & 0 & 3 & 7 \\ 0 & 0 & 0 & 4 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}$ (c) $A = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix}$

Övning 3.12 Beräkna $\det(A^3B^7)$ med $A = \begin{bmatrix} 4 & 3 \\ 2 & 4 \end{bmatrix}$ och $B = \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix}$

Problem

3.1 Matrisalgebra

Problem 3.1 Bevisa $(A+B)^2 = A^2 + AB + BA + B^2$. Motivera varje steg i beviset med hänvisning till räkneregler.

Problem 3.2 Bestäm A^k för alla heltal $k \geqslant 1$.
(a) $A = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$ (b) $A = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix}$

Problem 3.3 Låt $u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$. Vad är då kolonnerna i matrisen $A = uv^T \in \mathbb{R}^{m\times n}$? Visa att $Ax \in \operatorname{span}\{u\}$ om $x \in \mathbb{R}^n$. Matrisen $uv^T \in \mathbb{R}^{m\times n}$ kallas *tensorprodukten* av vektorerna u och v och tecknas $u \otimes v$. (Se övning 3.3.)

Problem 3.4 Låt $u \in \mathbb{R}^n$. Visa att matrisen $A = u \otimes u = uu^T \in \mathbb{R}^{n\times n}$ är symmetrisk, dvs att $A^T = A$.

Problem 3.5 Låt $A \in \mathbb{R}^{m\times n}$. Visa att matrisen $B = A^TA \in \mathbb{R}^{n\times n}$ är symmetrisk, dvs att $B^T = B$.

Problem 3.6 Antag att A är antisymmetrisk, dvs att $A^T = -A$. Visa att A är kvadratisk och att alla dess diagonalelement är 0, dvs alla $a_{ii} = 0$.

3.2 Invers matris

Problem 3.7 I definition 3.5 påstår vi att den inversa matrisen är unik. Bevisa detta. Mer precist: låt $A \in \mathbb{R}^{n\times n}$ och antag att det finns matriser $B, C \in \mathbb{R}^{n\times n}$ sådana att $BA = AB = I$ och $CA = AC = I$. Bevisa att $B = C$. Se även problem 3.12.

Problem 3.8 Använd sats 3.8 för att bevisa sats 3.5.

Problem 3.9 Visa att $A \in \mathbb{R}^{n\times n}$ är singulär om en av dess rader är en multipel av en annan rad.

Problem 3.10 Bevisa räknereglerna i sats 3.6.

3.3 Determinant

Problem 3.11 En kvadratisk matris A kallas *nilpotent* om $A^k = O$ för något heltal $k \geqslant 1$.

(a) Visa att $A = \begin{bmatrix} 0 & a & c \\ 0 & 0 & b \\ 0 & 0 & 0 \end{bmatrix}$ är nilpotent.

(b) Visa att determinanten för en nilpotent matris är 0.

Problem 3.12 Antag att A är kvadratisk och $AB = I$ för någon matris B. Visa att A är inverterbar med inversen B. (Om $AB = I$, så kallas B *högerinvers* till A. Påståendet är alltså att om en kvadratisk matris A har en högerinvers B, så är B en invers, dvs den är även en *vänsterinvers* : $BA = I$.)

Problem 3.13 Visa att $\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = (a-b)(b-c)(c-a)$.

4. Linjära rum

Vi har sett att i mängden av geometriska vektorer kan vi addera och multiplicera med skalär så att de vanliga räknereglerna gäller. Med andra ord: vi kan skapa nya vektorer genom linjär kombination. Vi kan också bilda skalärprodukt och ortogonal projektion. Vi ska nu se att dessa begrepp är fundamentala och kan användas i andra mängder också, till exempel i $\mathbb{R}^n$ och i funktionsrum. Sådana mängder kallas linjära rum *eller* vektorrum.

4.1 Vektorrum och underrum

Definition 4.1 (Linjärt rum) Ett *linjärt rum* (eller *vektorrum*) är en icke-tom mängd V av element (kallade *vektorer*) som vi kan *addera* och *multiplicera med skalär*. Följande räkneregler ska vara uppfyllda för alla vektorer $u, v \in V$ och skalärer $\alpha, \beta \in \mathbb{R}$.

1. $u + v \in V$	(sluten under addition)	(4.1)
2. $u + v = v + u$	(kommutativ lag)	(4.2)
3. $(u + v) + w = u + (v + w)$	(associativ lag)	(4.3)
4. $\exists\, 0 \in V \ \ (u + 0 = u)$	(det finns nollvektor)	(4.4)
5. $\exists -u \in V \ \ (u + (-u) = 0)$	(det finns motsatt vektor)	(4.5)
6. $\alpha u \in V$	(sluten under multiplikation med skalär)	(4.6)
7. $\alpha(u + v) = \alpha u + \alpha v$	(distributiv lag)	(4.7)
8. $(\alpha + \beta)u = \alpha u + \beta u$	(distributiv lag)	(4.8)
9. $(\alpha\beta)u = \alpha(\beta u)$	(associativ lag)	(4.9)
10. $1u = u$		(4.10)

Punkterna 1 och 6 innebär att mängden V är sluten under bildande av linjär kombination: en linjär kombination av vektorer i V är också en vektor i V (dvs en vektor av samma slag):

$$u = c_1 v_1 + \cdots + c_p v_p \in V \tag{4.11}$$

Det är ofta detta som är svårt att kolla medan övriga regler är naturliga och enkla.

Med ett *rum* menar vi en mängd som har någon extra struktur; i det här fallet *linjär struktur* (addition och multiplikation med skalär). Vi har då ett *linjärt rum* även kallat vektorrum.

Här definierar vi *reellt vektorrum* eftersom skalärerna är reella tal. Vi kan också definiera *komplext vektorrum* genom att låta skalärerna vara komplexa tal, $\alpha, \beta \in \mathbb{C}$.

Vi har redan tre exempel på linjära rum: geometriska vektorer (se sats 1.1 och sats 1.2), kolonnvektorer $\mathbb{R}^n$ (se sats 2.4) och matriser $\mathbb{R}^{m\times n}$ (se sats 3.1). Här ett nytt exempel.

Exempel 4.1 (Funktionsrum) Låt $V = C([0,1])$ vara mängden av alla kontinuerliga funktioner $f\colon [0,1] \to \mathbb{R}$. Addition och multiplikation med skalär definieras punktvis, dvs funktionerna $f+g$ och αf definieras genom

$$(f+g)(x) = f(x) + g(x),\ (\alpha g)(x) = \alpha g(x), \quad x \in [0,1] \tag{4.12}$$

Vi kollar reglerna i definition 4.1. Antag att $f, g \in C([0,1])$, dvs kontinuerliga funktioner på $[0,1]$. Då är $f+g$ och αf också kontinuerliga på $[0,1]$ enligt sats 3.4 i bokseriens del I. Det betyder att $f+g \in C([0,1])$ och $\alpha f \in C([0,1])$, så som krävs i punkterna 1 och 6. Detta är icke-triviala resultat från del I. Resten är enkelt: vi räknar punktvis med de vanliga räknereglerna för reella tal. Till exempel: nollfunktionen är $O(x) = 0$ för alla $x \in [0,1]$. Alltså är $C([0,1])$ ett vektorrum där vektorerna är funktioner. Därför kallar vi det för ett funktionsrum.

Definition 4.2 (Underrum) Ett *underrum* av ett vektorrum V är en delmängd U av V (dvs $U \subseteq V$) sådan att (med de linjära operationerna från V)

$$u, v \in U \implies u + v \in U \tag{4.13}$$

$$u \in U, \alpha \in \mathbb{R} \implies \alpha u \in U \tag{4.14}$$

dvs mängden U är sluten under linjär kombination.

Ett underrum U är självt ett vektorrum. Först noterar vi att nollvektorn tillhör U, $0 \in U$, genom att ta $\alpha = 0$ i (4.14). Sedan följer alla övriga räkneregler: de gäller i V och alla $u \in U$ är i V. Till exempel $u, v \in U$ medför att $u + v = v + u$ för detta gäller i V.

Alla vektorrum V har två *triviala underrum*:

$$U = \{0\} \text{ består endast av nollvektorn} \tag{4.15}$$

$$U = V \text{ hela rummet} \tag{4.16}$$

De är det minsta respektive det största underrummet; andra underrum ligger däremellan, $\{0\} \subseteq U \subseteq V$.

Exempel 4.2 (Underrum till $\mathbb{R}^3$) Rummet $\mathbb{R}^2$ är ej underrum till $\mathbb{R}^3$, för $\mathbb{R}^2$ är ej delmängd till $\mathbb{R}^3$. Men

$$U = \{v \in \mathbb{R}^3 \mid v_3 = 0\} = \{v \in \mathbb{R}^3 \mid v = \begin{bmatrix} v_1 & v_2 & 0 \end{bmatrix}^T\} \tag{4.17}$$

är underrum till $\mathbb{R}^3$, därför att v_3 förblir 0 under addition och multiplikation med skalär, dvs mängden U är sluten under linjär kombination. Geometriskt kan vi tolka U som x_1x_2-planet.

Mer allmänt: ett underrum till $\mathbb{R}^3$ är antingen en rät linje genom origo eller ett plan genom origo.

Ett viktigt sätt att bilda underrum är att bilda ett linjärt hölje. Detta är också ett viktigt sätt att beskriva ett underrum, nämligen att ange en mängd av vektorer som spänner upp rummet.

Sats 4.1 (Linjära höljet är underrum) Låt $v_1, \ldots, v_p \in V$ och $U = \operatorname{span}\{v_1, \ldots, v_p\}$. Då är U ett underrum till V.

Bevis. Låt $u, v \in U = \operatorname{span}\{v_1, \ldots, v_p\}$, dvs

$$u = a_1v_1 + \cdots + a_pv_p, \quad v = b_1v_1 + \cdots + b_pv_p \tag{4.18}$$

för några koefficienter a_k och b_k. Då har vi

$$u + v = (a_1 + b_1)v_1 + \cdots + (a_p + b_p)v_p \tag{4.19}$$
$$\alpha u = (\alpha a_1)v_1 + \cdots + (\alpha a_p)v_p \tag{4.20}$$

dvs $u + v \in U$ och $\alpha u \in U$. □

Om vi har en matris $A \in \mathbb{R}^{m\times n}$ så får vi två viktiga underrum till $\mathbb{R}^n$ respektive $\mathbb{R}^m$.

Definition 4.3 (Nollrum och kolonnrum) Låt $A = [a_1 \ \cdots \ a_n] \in \mathbb{R}^{m\times n}$. *Nollrummet* till A är

$$\mathcal{N}(A) = \{x \in \mathbb{R}^n \mid Ax = 0\} \tag{4.21}$$

Kolonnrummet till A är

$$\mathcal{C}(A) = \operatorname{span}\{a_1, \ldots, a_n\} \tag{4.22}$$

Nollrummet är alltså mängden av alla vektorer i $\mathbb{R}^n$ som avbildas på $0 \in \mathbb{R}^m$ av den linjära avbildningen $L\colon x \mapsto Ax$. Vi erinrar oss från definition 2.11 att

$$\operatorname{span}\{a_1, \ldots, a_n\} = \{y \in \mathbb{R}^m \mid y = Ax \text{ för något } x \in \mathbb{R}^n\} \tag{4.23}$$

dvs att kolonnrummet till A är värdemängden till L, $\mathcal{C}(A) = \mathcal{R}(L)$.

Sats 4.2 (Nollrummet är underrum) Låt $A \in \mathbb{R}^{m\times n}$. Nollrummet $\mathcal{N}(A)$ är ett underrum till $\mathbb{R}^n$.

Bevis. Antag $x, y \in \mathcal{N}(A)$, dvs $Ax = Ay = 0$. Då får vi

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay = 0 \tag{4.24}$$

dvs $\alpha x + \beta y \in \mathcal{N}(A)$, vilket visar att $\mathcal{N}(A)$ är sluten under linjär kombination. □

Sats 4.3 (Kolonnrummet är underrum) Låt $A \in \mathbb{R}^{m\times n}$. Kolonnrummet $\mathcal{C}(A)$ är ett underrum till $\mathbb{R}^m$.

Bevis. Eftersom $\mathcal{C}(A) = \text{span}\{a_1, \dots, a_n\}$, så följer påståendet av sats 4.1. □

Exempel 4.3 (Nollrum och kolonnrum) Med en (dator)beräkning får vi

$$A = \begin{bmatrix} -1 & 4 & 2 & 5 & -1 \\ 4 & -6 & 2 & 5 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 & 5 & 3 \\ 0 & 1 & 1 & \frac{5}{2} & \frac{1}{2} \end{bmatrix} \tag{4.25}$$

Nollrummet fås genom att lösa den homogena ekvationen $Ax = 0$. Vi har tre fria variabler: x_3, x_4 och x_5. Med $x_3 = s$, $x_4 = t$, $x_5 = u$ blir lösningarna på parameterform

$$x = \begin{bmatrix} -2s - 5t - 3u \\ -s - \frac{5}{2}t - \frac{1}{2}u \\ s \\ t \\ u \end{bmatrix} = s\begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -5 \\ -\frac{5}{2} \\ 0 \\ 1 \\ 0 \end{bmatrix} + u\begin{bmatrix} -3 \\ -\frac{1}{2} \\ 0 \\ 0 \\ 1 \end{bmatrix} = sv_1 + tv_2 + uv_3 \tag{4.26}$$

Det betyder att nollrummet är

$$\mathcal{N}(A) = \text{span}\{v_1, v_2, v_3\} = \text{span}\left\{ \begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ -\frac{5}{2} \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ -\frac{1}{2} \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\} \tag{4.27}$$

Mängden $\{v_1, v_2, v_3\}$ är linjärt oberoende, för vi har en etta och resten nollor på plats 3, 4, 5 i respektive vektor.

Kolonnrummet kan vi skriva ned direkt

$$\mathcal{C}(A) = \text{span}\{a_j\}_{j=1}^5 = \text{span}\left\{ \begin{bmatrix} -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ -6 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 5 \end{bmatrix}, \begin{bmatrix} -1 \\ 9 \end{bmatrix} \right\} \tag{4.28}$$

Eftersom det finns icke-triviala homogenlösningar, så förstår vi att mängden $\{a_j\}_{j=1}^5$ är linjärt beroende. Vi ska senare se att de två första vektorerna (dvs pivotkolonnerna) bildar en linjärt oberoende mängd och räcker för att spänna upp hela $\mathcal{C}(A)$, dvs $\mathcal{C}(A) = \text{span}\{a_1, a_2\}$ (se exempel 4.10).

Vi har beskrivit de två rummen $\mathcal{N}(A)$ och $\mathcal{C}(A)$ genom att ange två linjärt oberoende mängder som spänner upp dem.

Vi samlar några observationer om nollrummet och kolonnrummet. Vi erinrar oss först att lösningsmängden till $Ax = b$ har formen $x = x_\text{p} + x_\text{h}$ (om den inte är tom) enligt sats 2.7.

Ruta 4.1 (Nollrummet och kolonnrummet)
Låt $A \in \mathbb{R}^{m\times n}$.

1. $\mathcal{N}(A) \subseteq \mathbb{R}^n, \mathcal{C}(A) \subseteq \mathbb{R}^m$.
2. Lätt att testa om $x \in \mathcal{N}(A)$: beräkna Ax.
3. Svårt att hitta en $x \in \mathcal{N}(A)$: lös $Ax = 0$.
4. Lätt att hitta en $b \in \mathcal{C}(A)$: bilda linjär kombination $b = \sum_{k=1}^n c_k a_k$.
5. Svårt att testa om $b \in \mathcal{C}(A)$: lös $Ax = b$.
6. $x \in \mathcal{N}(A)$ om och endast om x är homogenlösning till $Ax = b$.
7. $b \in \mathcal{C}(A)$ om och endast om $Ax = b$ har en partikulärlösning.
8. Lösningen till $Ax = b$ är unik (om det finns någon) om och endast om $\mathcal{N}(A) = \{0\}$.
9. $Ax = b$ är lösbar för varje $b \in \mathbb{R}^m$ om och endast om $\mathcal{C}(A) = \mathbb{R}^m$.

Villkoret $\mathcal{N}(A) = \{0\}$ (trivialt nollrum) är alltså associerat med entydighet och villkoret $\mathcal{C}(A) = \mathbb{R}^m$ (trivialt kolonnrum) är associerat med lösbarhet enligt punkterna 8 och 9. Då $m = n$ är dessa villkor ekvivalenta enligt nästa sats, som är en omformulering av sats 3.9.

Sats 4.4 (Triviala kolonnrum eller nollrum är ekvivalent med inverterbarhet) Låt $A \in \mathbb{R}^{n\times n}$. Följande påståenden är ekvivalenta.

- A är inverterbar.
- $\mathcal{N}(A) = \{0\}$.
- $\mathcal{C}(A) = \mathbb{R}^n$.

Bevis. A är inverterbar om och endast om $A \sim I$. Men $A \sim I$ är ekvivalent med att A har pivotpositioner i varje rad, dvs vi har inga fria variabler i ekvationen $Ax = b$. Detta är i sin tur ekvivalent med att $Ax = b$ har unik lösning för alla $b \in \mathbb{R}^n$. Alltså: både $\mathcal{N}(A) = \{0\}$ (inga fria variabler) och $\mathcal{C}(A) = \mathbb{R}^n$ (lösbarhet för alla b) är ekvivalenta med att $A \sim I$. $\square$

Vi avslutar detta avsnitt med att definiera linjär operator i allmänna vektorrum. Då kan vi sedan presentera viktiga linjära operatorer som inte är av formen $L(x) = Ax$.

Definition 4.4 (Linjär funktion) Låt V och W vara linjära rum. En funktion $L\colon V \to W$ kallas *linjär funktion* (eller *linjär operator*) om den bevarar linjär kombination, dvs om

$$L(\alpha u + \beta v) = \alpha(L(u)) + \beta(L(v)) \tag{4.29}$$

för alla $u, v \in V$ och för alla $\alpha, \beta \in \mathbb{R}$. Nollrummet och värderummet för L är

$$\mathcal{N}(L) = \{u \in V \mid L(u) = 0\} \tag{4.30}$$

$$\mathcal{R}(L) = \{w \in W \mid w = L(u) \text{ för något } u \in V\} \tag{4.31}$$

Värderummet kan nu inte kallas kolonnrum, för det behöver inte finnas någon matris A sådan att $L(x) = Ax$. Vi säger värderummet istället för värdemängden därför att $\mathcal{R}(L)$ är ett linjärt rum, vilket vi ska strax visa i sats 4.5.

Linjära operatorer skrivs ofta utan parenteser, dvs Lu istället för $L(v)$. Till exempel får vi då $L(\alpha u + \beta v) = \alpha(Lu) + \beta(Lv)$. Vi kan också addera och multiplicera linjära operatorer:

$$(L + K)u = Lu + Ku \tag{4.32}$$

$$(\alpha L)u = \alpha(Lu) \tag{4.33}$$

Den senare identiten skrivs ofta utan parenteser: $\alpha Lu = (\alpha L)u = \alpha(Lu)$, precis som för matris-vektormultiplikation i (3.39). Mängden av linjära operatorer $V \to W$ blir då ett vektorrum och de "vanliga räknereglerna" gäller, på samma sätt som för mängden av matriser i sats 3.1. Vi går inte igenom detaljerna.

Sats 4.5 ($\mathcal{N}(L)$ och $\mathcal{R}(L)$ är underrum) Nollrummet $\mathcal{N}(L)$ och värderummet $\mathcal{R}(L)$ är underrum till V respektive W.

Bevis. Att $\mathcal{N}(L)$ är ett underrum till V visas på samma vis som sats 4.2. För att visa att värderummet är ett underrum till W, antar vi nu att $w, z \in \mathcal{R}(L)$, dvs $w = Lu$ och $z = Lv$ för några $u, v \in V$. Då får vi $\alpha w + \beta z = \alpha Lu + \beta Lv = L(\alpha u + \beta v)$, dvs $\alpha w + \beta z \in \mathcal{R}(L)$. Värderummet är alltså slutet under linjär kombination och därmed ett underrum. □

Exempel 4.4 (Differentialoperator) Låt $V = C^2(\mathbb{R})$ vara mängden av alla två gånger kontinuerligt deriverbara funktioner $u\colon \mathbb{R} \to \mathbb{R}$ och $W = C(\mathbb{R})$ vara mängden av alla kontinuerliga funktioner. Att V och W är linjära rum visas som i exempel 4.1. (Rummet V är slutet under linjär kombination, eftersom en linjär kombination av deriverbara funktioner är också deriverbar. I själva verket är $C^2(\mathbb{R})$ ett underrum till $C(\mathbb{R})$.)

Vi definierar differentialoperatorer av ordning 1 och 2:

$$\mathrm{D}\colon V \to W, \qquad \mathrm{D}\colon u \mapsto u' \tag{4.34}$$

$$\mathrm{D}^2 = \mathrm{DD}\colon V \to W, \qquad \mathrm{D}^2\colon u \mapsto u'' \tag{4.35}$$

$$L = \mathrm{D}^2 + 1\colon V \to W, \qquad L\colon u \mapsto u'' + u \tag{4.36}$$

Vi har $L\colon V \to W$ eftersom om vi deriverar $u \in C^2(\mathbb{R})$ två gånger så får vi en kontinuerlig funktion, dvs en funktion i $W = C(\mathbb{R})$. Vidare ger deriveringsreglerna att

$$L(\alpha u+\beta v) = (\alpha u+\beta v)''+(\alpha u+\beta v) = \alpha(u''+u)+\beta(v''+v) = \alpha Lu+\beta Lv \tag{4.37}$$

dvs L är linjär.

Nu ska vi undersöka ekvationen $Lu = f$ med $f \in W$. Precis som för linjära ekvationssystem i sats 2.7 kan vi visa att lösningsmängden har formen $u = u_\mathrm{p} + u_\mathrm{h}$ om den inte är tom, dvs om det finns en partikulärlösning (se sats 3.7 i bokseriens del II). Vi börjar med den homogena ekvationen $Lu = 0$, dvs differentialekvationen

$$u''(x) + u(x) = 0 \tag{4.38}$$

Vi kommer ihåg från del II att allmänna lösningen är

$$u(x) = a\cos(x) + b\sin(x), \quad a, b \in \mathbb{R} \tag{4.39}$$

Det betyder att nollrummet till L är

$$\mathcal{N}(L) = \mathrm{span}\{\cos, \sin\} \tag{4.40}$$

Vi tar sedan (för enkelhets skull) $f(x) = x^2$ och söker en partikulärlösning till $Lu = f$, dvs

$$u''(x) + u(x) = x^2 \tag{4.41}$$

Insättning av ansatsen $u_\mathrm{p}(x) = Ax^2 + Bx + C$ ger $A = 1$, $B = 0$, $C = -2$, dvs $u_\mathrm{p}(x) = x^2 - 2$. Lösningsmängden är på formen

$$u = u_\mathrm{p} + u_\mathrm{h}, \quad \text{med } u_\mathrm{p}(x) = x^2 - 2 \text{ och } u_\mathrm{h} \in \mathrm{span}\{\cos, \sin\} \tag{4.42}$$

I del II visade vi att differentialekvationen är lösbar för alla $f \in C(\mathbb{R})$. Det betyder att värderummet är $\mathcal{R}(L) = C(\mathbb{R}) = W$.

Nu kan vi studera linjära operatorekvationer, mer allmänna än linjära ekvationssystem, till exempel linjära differentialekvationer. Om $L\colon V \to W$ är en linjär operator, så kan vi ställa upp en *linjär ekvation*: givet $w \in W$ sök $v \in V$ sådan att

$$Lv = w \tag{4.43}$$

Precis som för linjärt ekvationssystem $Ax = b$ i ruta 4.1 har vi att

1. Ekvationen $Lv = w$ är lösbar för alla $w \in W$ om och endast om $\mathcal{R}(L) = W$.

2. Lösningen är unik om och endast om $\mathcal{N}(L) = \{0\}$.

Om båda dessa villkor är uppfyllda så har vi alltså för varje $w \in W$ en unik lösning $v \in V$. Detta definierar en invers funktion $L^{-1}\colon W \to V$ genom att vi tar $L^{-1}w = v$.

Vi uppepar detta med beteckningar från avsnitt 2.1 i bokseriens del I. Villkoret $\mathcal{R}(L) = W$ innebär att funktionen $L\colon V \to W$ är *surjektiv*,

$$\forall\, w \in W \;\; \exists\, v \in V \;\; (w = Lv) \tag{4.44}$$

Detta är självklart. Villkoret $\mathcal{N}(L) = \{0\}$ innebär att L är *injektiv*, dvs

$$\forall\, v_1, v_2 \in V \;\; (v_1 \neq v_2 \implies Lv_1 \neq Lv_2) \tag{4.45}$$

För att se detta använder vi att L är en linjär operator så att

$$Lv_1 = Lv_2 \iff L(v_1 - v_2) = 0 \iff v_1 - v_2 \in \mathcal{N}(L) \tag{4.46}$$

Det är nu klart att L är injektiv om och endast om $\mathcal{N}(L) = \{0\}$. Om L är *bijektiv*, dvs både injektiv och surjektiv, så är den *inverterbar*. Inversen $v = L^{-1}w$ definieras som den unika lösningen till $Lv = w$.

Sats 4.6 **(Inversen är en linjär operator)** Antag att $L\colon V \to W$ är en linjär och bijektiv operator. Då är den inversa operatorn $L^{-1}\colon W \to V$ linjär.

Bevis. Vi lämnar beviset till problem 4.1. □

I exempel 4.4 formulerade vi en linjär differentialekvation som en linjär operatorekvation. Vi tar ett exempel till.

Exempel 4.5 **(Randvärdesproblem)** I avsnitt 6.5 i del II studerade vi randvärdesproblem av typen

$$\begin{cases} -\,u''(x) = f(x), \quad 0 < x < 1 \\ u(0) = 0, \; u(1) = 0 \end{cases} \tag{4.47}$$

Vi ska formulera detta som en linjär operatorekvation.

Låt $C([0,1])$ och $C^2([0,1])$ vara mängderna av alla kontinuerliga funktioner respektive alla två gånger kontinuerligt deriverbara funktioner $u\colon [0,1] \to \mathbb{R}$. Låt $C_0^2([0,1])$ beteckna underrummet av funktioner som uppfyller de homogena randvillkoren i (4.47):

$$C_0^2([0,1]) = \{u \in C^2([0,1]) \mid u(0) = 0, \; u(1) = 0\} \tag{4.48}$$

Vi väljer $V = C_0^2([0,1])$ och $W = C([0,1])$. Att V och W är linjära rum visas som i exempel 4.1. Att $C_0^2([0,1])$ är ett underrum till $C^2([0,1])$ beror på att de homogena randvillkoren bevaras under linjär kombination.

Vidare definierar vi differentialoperatorn

$$L = -\mathrm{D}^2 \colon V \to W, \quad L \colon u \mapsto -u'' \tag{4.49}$$

och låter $f \in W$. Randvärdesproblemet (4.47) är då ekvivalent med operatorekvationsproblemet att finna $u \in V$ sådan att $Lu = f$.

Lösningar fås genom att integrera differentialekvationen i (4.47) två gånger:

$$-u''(x) = f(x), \quad u'(x) = -F(x) + C, \quad u(x) = -G(x) + Cx + D \tag{4.50}$$

där F är en primitiv funktion till f, G är en primitiv funktion till F och C, D är integrationskonstanter. Konstanterna kan bestämmas med hjälp av randvillkoren. Detta visar att det finns lösningar, dvs $\mathcal{R}(L) = W$ (surjektiv).

Den homogena ekvationen har allmänna lösningen $u(x) = Cx + D = 0$, eftersom randvillkoren ger $C = D = 0$. Detta visar att $\mathcal{N}(L) = \{0\}$ (injektiv).

Operatorn $L \colon V \to W$ är alltså inverterbar och den unika lösningen kan skrivas $u = L^{-1} f$. Obs att randvillkoren är inbyggda i operatorn L via dess definitionsmängd $V = C_0^2([0,1])$. Detta är viktigt; annars vore L inte injektiv.

4.2 Baser och komponenter

För $x \in \mathbb{R}^n$ har vi

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + x_n \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \sum_{j=1}^{n} x_j e_j \tag{4.51}$$

där e_j är den j:te basvektorn med endast nollor utom en etta i rad nummer j. Vi kallar mängden $E = \{e_1, \ldots, e_n\}$ *standardbasen* för $\mathbb{R}^n$ och $[x]_E = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ *komponenterna* för x i basen E. (Här är $[x]_E = x$ eftersom E är standardbasen.) Vi noterar att mängden E är linjärt oberoende och spänner upp $\mathbb{R}^n$. Mer allmänt har vi följande definition.

Definition 4.5 (Bas) En *bas* för vektorrummet V är en *linjärt oberoende mängd* $B = \{b_1, \dots, b_n\}$ av vektorer som *spänner upp* V, dvs $V = \operatorname{span}\{b_1, \dots, b_n\}$. Komponenterna för $v \in V$ i basen B är koefficienterna

$$[v]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}, \quad \text{där } v = \sum_{j=1}^{n} c_j b_j \tag{4.52}$$

Beteckningar: $B = \{b_1, \dots, b_n\}$ är *en bas* som består flera *basvektorer* b_j. Vektorerna kallas *inte* baser, de är basvektorer![1]

Vi gör en enkel observation.

Sats 4.7 (Komponenterna är unika) Låt $B = \{b_1, \dots, b_n\}$ vara en bas för vektorrummet V. Då kan varje $v \in V$ skrivas entydigt som $v = \sum_{j=1}^{n} c_j b_j$, dvs komponenterna $[v]_B = [c_1 \cdots c_n]^T$ är unika.

Bevis. Eftersom $V = \operatorname{span}\{b_1, \dots, b_n\}$ så kan varje $v \in V$ skrivas $v = \sum_{j=1}^{n} c_j b_j$. Antag nu att v kan skrivas som linjär kombination på två sätt,

$$v = \sum_{j=1}^{n} c_j b_j, \quad v = \sum_{j=1}^{n} d_j b_j \tag{4.53}$$

Då får vi

$$\sum_{j=1}^{n} (c_j - d_j) b_j = 0 \tag{4.54}$$

Men B är linjärt oberoende, så att alla $c_j - d_j = 0$, dvs $c_j = d_j$. Alltså: att B är linjärt oberoende medför att komponenterna är unika. $\square$

En bas B för V innehåller den minsta tillräckliga information som beskriver V: ta bort en vektor från B så kan vi inte bilda hela V; lägg till en vektor så är B inte längre linjärt oberoende.

Exempel 4.6 (En bas för $\mathbb{R}^3$) Vi betraktar mängden

$$B = \{b_1, b_2, b_3\} = \left\{ \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ 3 \end{bmatrix} \right\} \tag{4.55}$$

För att avgöra om B är en bas för $\mathbb{R}^3$ bildar vi matrisen $P_B = [b_1\ b_2\ b_3]$ och radreducerar (med datorns hjälp)

$$P_B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & 1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.56}$$

1. Detta är ett vanligt nybörjarmisstag.

Vi drar slutsatsen att P_B är inverterbar och därmed att $\mathcal{N}(P_B) = \{0\}$ och $\mathcal{C}(P_B) = \mathbb{R}^3$ enligt sats 4.4. Testet för linjärt oberoende är vektorekvationen $x_1 b_1 + x_2 b_2 + x_3 b_3 = 0$, dvs $P_B x = 0$. Men $\mathcal{N}(P_B) = \{0\}$ betyder att vi har endast triviala lösningen $x = 0$, dvs B är linjärt oberoende. Att $\operatorname{span}(B) = \mathcal{C}(P_B) = \mathbb{R}^3$ innebär att B spänner upp hela $\mathbb{R}^3$. Alltså: B är en bas för $\mathbb{R}^3$. Den är annorlunda än standardbasen $E = \{e_1, e_2, e_3\}$.

Nu kan vi uttrycka en godtycklig vektor i basen B. Till exempel $x = [1\ 1\ 1]^T$. Vi löser $P_B c = x$, dvs

$$\begin{bmatrix} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 1 \\ 1 & 1 & 3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{2} \end{bmatrix}, \quad c = \begin{bmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix} \tag{4.57}$$

Alltså: komponenterna är $[x]_B = c = [-\frac{1}{2}\ 0\ \frac{1}{2}]^T$.

Vi ska nu diskutera hur komponenterna ändras när vi byter bas i $\mathbb{R}^n$. I standardbasen $E = \{e_1, \dots, e_n\}$ har vektorn x komponenterna $[x]_E = x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$. I en annan bas $B = \{b_1, \dots, b_n\}$ har vi

$$[x]_B = c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}, \quad x = \sum_{j=1}^{n} c_j b_j = [b_1\ \dots\ b_n]c = P_B c \tag{4.58}$$

där vi infört *basvektormatrisen*

$$P_B = [b_1\ \cdots\ b_n] \tag{4.59}$$

Sambandet mellan komponenterna i standardbasen och i den nya basen är alltså $x = P_B c$, dvs

$$[x]_E = P_B\,[x]_B \tag{4.60}$$

Matrisen P_B är inverterbar (varför?), så den inversa transformationen ges av

$$[x]_B = P_B^{-1}\,[x]_E \tag{4.61}$$

Om vi har ytterligare en bas $D = \{d_1, \dots, d_n\}$, så får vi

$$[x]_E = P_D\,[x]_D, \quad [x]_D = P_D^{-1}\,[x]_E, \quad P_D = [d_1\ \cdots\ d_n] \tag{4.62}$$

Tillsammans med (4.60) ger detta

$$[x]_D = P_D^{-1}\,[x]_E = P_D^{-1}\,P_B\,[x]_B \tag{4.63}$$

Vi har bevisat följande sats.

Sats 4.8 (Basbyte i $\mathbb{R}^n$) Låt $B = \{b_1, \ldots, b_n\}$ och $D = \{d_1, \ldots, d_n\}$ vara två baser i $\mathbb{R}^n$ med basvektormatriser $P_B = [b_1 \ \ldots \ b_n]$ respektive $P_D = [d_1 \ \ldots \ d_n]$. Då gäller

$$P_D\,[x]_D = P_B\,[x]_B, \quad [x]_D = P_D^{-1}\,P_B\,[x]_B \tag{4.64}$$

I ett vektorrum V som inte är $\mathbb{R}^n$ har vi ingen standardbas och ingen basvektormatris. Men om det finns en bas $B = \{b_1, \ldots, v_n\}$, så kan vi alltid skriva entydigt $v = \sum_{j=1}^{n} c_j b_j$ och definiera en *komponentavbildning* L_B:

$$L_B \colon V \to \mathbb{R}^n \tag{4.65}$$

$$v \mapsto [v]_B \tag{4.66}$$

där $[v]_B = c = [c_1 \ \cdots \ c_n]^T$. Vi illustrerar detta med ett rum av polynom.

Exempel 4.7 (Bas för polynomrummet) Vi låter P_5 beteckna mängden av alla polynom av grad högst 5. Det är klart att en linjär kombination av polynom av grad högst 5 är också ett polynom av grad högst 5. P_5 är alltså ett linjärt rum, ett underrum till $C(\mathbb{R})$.
Om $p \in P_5$ så

$$p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 = \sum_{k=1}^{5} a_k p_k(x) \tag{4.67}$$

med $p_k(x) = x^k$, dvs

$$p_0(x) = 1,\ p_1(x) = x,\ p_2(x) = x^2,\ p_3(x) = x^3,\ p_4(x) = x^4,\ p_5(x) = x^5 \tag{4.68}$$

Vi har alltså $P_5 = \text{span}(B)$ med $B = \{p_0, p_1, p_2, p_3, p_4, p_5\}$.
Är mängden B linjärt oberoende? Vi löser den homogena vektorekvationen $\sum_{k=1}^{5} c_k p_k = 0$, dvs

$$c_0 + c_1x + c_2x^2 + c_3x^3 + c_4x^4 + c_5x^5 = 0 \tag{4.69}$$

Genom att välja $x = 0$ får vi $c_0 = 0$ och därmed $c_1x + c_2x^2 + c_3x^3 + c_4x^4 + c_5x^5 = 0$. Vi förkortar med x, väljer sedan $x = 0$ och får $c_1 = 0$. På detta vis visar vi att vi har endast den triviala lösningen där alla $c_k = 0$. Alltså: B är linjärt oberoende och $P_5 = \text{span}(B)$, dvs en bas för P_5.
Komponentavbildningen är

$$L_B \colon P_5 \to \mathbb{R}^6 \tag{4.70}$$

$$p = \sum_{k=1}^{5} a_k p_k \mapsto [p]_B = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} \tag{4.71}$$

Till exempel $p(x) = (1-x)^2 = 1 - 2x + x^2$ ger

$$[p]_B = \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{4.72}$$

Sats 4.7 betyder att komponentavbildningen är inverterbar.

Sats 4.9 (Komponentavbildningen är isomorfism) Låt $B = \{b_1, \ldots, b_n\}$ vara en bas för V. Komponentavbildningen

$$L_B \colon V \to \mathbb{R}^n \tag{4.73}$$
$$v \mapsto [v]_B \tag{4.74}$$

är linjär och inverterbar.

Bevis. Att L_B är linjär är klart. Resten följer av Sats 4.7. Att varje $v \in V$ kan skrivas $v = \sum_{j=1}^n c_j b_j$, där $[v]_B = c = [c_1 \ \cdots \ c_n]^T$, betyder att $\mathcal{R}(L_B) = \mathbb{R}^n$, dvs att L_B är surjektiv. Att komponenterna är unika betyder att L_B är injektiv. Den är alltså bijektiv och därmed inverterbar. □

En linjär och inverterbar avbildning kallas *isomorfism*. Komponentavbildningen är alltså en isomorfism mellan vektorrummen V och $\mathbb{R}^n$. Vi säger att vektorrummen är isomorfa, $V \simeq \mathbb{R}^n$. Det betyder att de har samma linjära struktur. Till exempel har vi sett att $P_5 \simeq \mathbb{R}^6$.

4.3 Dimension och rang

Vi har sett om V har en bas $B = \{b_1, \ldots, b_n\}$ så är V isomorft med $\mathbb{R}^n$. Isomorfismen ges av komponentavbildningen $L_B \colon v \mapsto [v]_B$. Det verkar troligt att alla baser för V har n element, dvs lika många som $\mathbb{R}^n$. Vi ska nu bevisa detta. Först visar vi att alla baser för V har högst n element.

Sats 4.10 (Alla baser har högst n element) Låt $B = \{b_1, \ldots, b_n\}$ vara en bas för V. Då är varje delmängd av V med fler än n element linjärt beroende.

Bevis. Låt $U \subseteq V$ vara en delmängd med fler än n element, dvs $U = \{u_k\}_{k=1}^p \subseteq V$ med $p > n$. Vi har bevisat satsen i $\mathbb{R}^n$, se sats 2.9. Alltså är mängden av komponentvektorer $L_B(U) =$

$\{[u_k]_B\}_{k=1}^p$ linjärt beroende i $\mathbb{R}^n$. Det betyder att det finns koefficienter c_k, ej alla lika med 0, så att

$$0 = c_1[u_1]_B + \cdots + c_p[u_p]_B \tag{4.75}$$
$$= c_1 L_B(u_1) + \cdots + c_p L_B(u_p) \tag{4.76}$$
$$= L_B(c_1 u_1 + \cdots + c_p u_p) \tag{4.77}$$

Här har vi använt att komponentavbildningen är linjär. Men den är också inverterbar (sats 4.9), så att

$$c_1 u_1 + \cdots + c_p u_p = L_B^{-1}(0) = 0 \tag{4.78}$$

Det bevisar att U är linjärt beroende. □

Sats 4.11 (Alla baser har lika många element) Alla baser för ett vektorrum V har lika många element.

Bevis. Låt B och D vara två baser för V. Eftersom D är linjärt oberoende och B en bas, så säger sats 4.10 att $\#D \leqslant \#B$. Genom att låta B och D byta roll får vi att $\#B \leqslant \#D$. Alltså: lika många. ($\#S$ betecknar antalet element i mängden S.) □

Antalet basvektorer är rummets dimension.

Definition 4.6 (Dimension) Om vektorrummet V spänns upp av en ändlig mängd så är V ett *ändlig-dimensionellt* rum, annars *oändlig-dimensionellt*. I ändliga fallet är *dimensionen* för V, $\dim(V) = \#B$, där B är en bas. I oändliga fallet skriver vi $\dim(V) = \infty$ och om $V = \{0\}$ så är $\dim(V) = 0$.

Vi har redan några exempel: $\dim(\mathbb{R}^n) = n$, $\dim(P_5) = 6$.

Exempel 4.8 (Oändlig-dimensionellt vektorrum) Låt V vara mängden av alla talföljder $x = (x_1, x_2, \dots) = (x_k)_{k=1}^\infty$. Vi adderar dem och multiplicerar dem med skalär termvis och får nya talföljder, se del I. V är alltså ett linjärt rum.

Rummet V har ingen (ändlig) bas. Mängden $\{e_j\}_{j=1}^p$ med

$$e_j = (0, \dots, 0, 1, 0, \dots) \quad \text{(1 i term nummer } j\text{)} \tag{4.79}$$

är linjärt oberoende; det är lätt att se. Den är en linjärt oberoende delmängd av V med p element. Enligt sats 4.10 måste varje bas $B = \{b_1, \dots, b_n\}$ för V ha minst p element, dvs $n \geqslant p$. Men p är godtyckligt, så detta är omöjligt. Alltså: $\dim(V) = \infty$.

Även $C(\mathbb{R})$ är ett oändlig-dimensionellt vektorrum. Mängden $\{p_k\}_{k=0}^m$, där $p_k(x) = x^k$, är nämligen en linjärt oberoende delmängd av $C(\mathbb{R})$ med $m+1$ element. Eftersom m är godtyckligt, så får vi $\dim(C(\mathbb{R})) = \infty$.

Exempel 4.9 (Differentialoperator) I exempel 4.4 studerade vi differentialoperatorn $L = \mathrm{D}^2 + 1$ och bestämde dess nollrum

$$\mathcal{N}(L) = \text{span}\{\cos, \sin\} \tag{4.80}$$

Är mängden $B = \{\cos, \sin\}$ linjär oberoende? Vi löser den homogena ekvationen $a\cos + b\sin = 0$, dvs

$$a\cos(x) + b\sin(x) = 0 \tag{4.81}$$

Med $x = 0$ får vi $a = 0$ och med $x = \pi/2$ får vi $b = 0$. Vi har endast den triviala lösningen och därför är B en bas för $\mathcal{N}(L)$ och $\dim(\mathcal{N}(L)) = 2$.

Nu återvänder vi till viktiga underrum i $\mathbb{R}^n$: nollrum och kolonnrum för matriser.

Exempel 4.10 (Dimension för nollrum och kolonnrum) Vi återvänder till exempel 4.3. Radreducering ger

$$A = \begin{bmatrix} -1 & 4 & 2 & 5 & -1 \\ 4 & -6 & 2 & 5 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 & 5 & 3 \\ 0 & 1 & 1 & \frac{5}{2} & \frac{1}{2} \end{bmatrix} = U \tag{4.82}$$

Vi har tre fria variabler x_1, x_2, x_3. Homogena ekvationen $Ax = 0$ har lösningarna på parameterform

$$x = s\begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -5 \\ -\frac{5}{2} \\ 0 \\ 1 \\ 0 \end{bmatrix} + u\begin{bmatrix} -3 \\ -\frac{1}{2} \\ 0 \\ 0 \\ 1 \end{bmatrix} = sv_1 + tv_2 + uv_3 \tag{4.83}$$

Det betyder att nollrummet är

$$\mathcal{N}(A) = \text{span}\{v_1, v_2, v_3\} \tag{4.84}$$

Det är klart att mängden $B = \{v_1, v_2, v_3\}$ är linjärt oberoende, vektorerna har ju varsin etta och resten nollor i raderna $3, 4, 5$. B är en bas för $\mathcal{N}(A)$ och $\dim(\mathcal{N}(A)) = 3$.

Kolonnrummet är

$$\mathcal{C}(A) = \text{span}\{a_j\}_{j=1}^5 = \text{span}\left\{\begin{bmatrix} -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ -6 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 5 \end{bmatrix}, \begin{bmatrix} -1 \\ 9 \end{bmatrix}\right\} \tag{4.85}$$

Fem vektorer i $\mathbb{R}^2$ kan inte vara linjärt oberoende (sats 2.9). Men pivotkolonnerna, kolonn 1 och 2, i den reducerade trappstegsmatrisen U är linjärt oberoende och spänner upp $\mathbb{R}^2$. Ekvationerna $Ax = 0$ och $Ux = 0$ har samma lösningar. Det betyder att *elementära radoperationer bevarar linjära samband mellan kolonnerna*. Därför är också pivotkolonnerna, kolonn 1 och 2, i A linjärt oberoende och spänner upp $\mathbb{R}^2$. Vi har alltså

$$\mathcal{C}(A) = \text{span}\{a_j\}_{j=1}^5 = \text{span}\{a_j\}_{j=1}^2 = \text{span}\left\{\begin{bmatrix} -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ -6 \end{bmatrix}\right\} \tag{4.86}$$

och en bas för $\mathcal{C}(A)$ är $D = \{a_1, a_2\}$ med $\dim(\mathcal{C}(A)) = 2$.

Definition 4.7 (Rang) *Rangen* för matrisen $A \in \mathbb{R}^{m\times n}$ är

$$\operatorname{rang}(A) = \dim(\mathcal{C}(A)) \tag{4.87}$$

Vi gör radreducering $A \sim U$. Genom att inspektera den reducerade trappstegsmatrisen U som i exempel 4.10 drar vi slutsatser om $\mathcal{C}(A)$ och $\mathcal{N}(A)$. Vi sammanfattar i följande ruta.

Ruta 4.2 (Bas och dimension för nollrum och kolonnrum)
Låt $A \in \mathbb{R}^{m\times n}$.

1. Pivotkolonnerna i A bildar en bas för kolonnrummet $\mathcal{C}(A)$.
2. $\operatorname{rang}(A) = \dim(\mathcal{C}(A))$ är lika med antalet pivotkolonner i A.
3. En bas för $\mathcal{N}(A)$ fås genom att skriva lösningarna till $Ax = 0$ på parameterform.
4. $\dim(\mathcal{N}(A))$ är lika med antalet icke-pivotkolonner i A, dvs antalet fria variabler i $Ax = 0$.
5. Om $\operatorname{rang}(A) = m$ säger vi att A har *full rang* (eller maximal rang). Detta är ekvivalent med att $Ax = b$ är lösbar för alla b.
6. Lösningen till $Ax = b$ är unik om endast om $\dim(\mathcal{N}(A)) = 0$.

Eftersom $\mathcal{C}(A) \subseteq \mathbb{R}^m$ och $\mathcal{N}(A) \subseteq \mathbb{R}^n$ har vi $\operatorname{rang}(A) \leqslant m$ och $\dim(\mathcal{N}(A)) \leqslant n$. I punkt 5 har kolonnrummet maximal dimension och i punkt 6 har nollrummet minimal dimension. Summan av dessa dimensioner är alltid lika med n enligt nästa sats.

Sats 4.12 (Rangsatsen) Om $A \in \mathbb{R}^{m\times n}$ så gäller att

$$\operatorname{rang}(A) + \dim(\mathcal{N}(A)) = n \tag{4.88}$$

Bevis. Vi har att $\#\{\text{pivotkolonner}\} + \#\{\text{icke-pivotkolonner}\} = \#\{\text{kolonner}\} = n$. Detta är detsamma som (4.88). ($\#S$ betecknar antalet element i mängden S.) □

Sats 4.4 kan nu formuleras så här. Obs att nu är $m = n$.

Sats 4.13 (Maximal rang eller minimalt nollrum är ekvivalent med inverterbarhet) Låt $A \in \mathbb{R}^{n\times n}$. Följande påståenden är ekvivalenta.

- A är inverterbar.
- $\dim(\mathcal{N}(A)) = \{0\}$.
- $\operatorname{rang}(A) = n$.

4.4 Skalärprodukt, ortogonalitet, projektion

Vi har redan använt skalärprodukt, längd och ortogonal projektion för geometriska vektorer:

$$\boldsymbol{u}\cdot\boldsymbol{v} = |\boldsymbol{u}|\,|\boldsymbol{v}|\cos(\theta) \tag{4.89}$$

$$\boldsymbol{u}\cdot\boldsymbol{u} = |\boldsymbol{u}|^2 \tag{4.90}$$

$$P_{\boldsymbol{u}}(\boldsymbol{v}) = (\boldsymbol{v}\cdot\hat{\boldsymbol{u}})\hat{\boldsymbol{u}}, \quad \hat{\boldsymbol{u}} = \frac{\boldsymbol{u}}{|\boldsymbol{u}|} \tag{4.91}$$

Vi ska nu göra detsamma för vektorer i $\mathbb{R}^n$ och i allmänna rum, till exempel funktionsrum.

Definition 4.8 (Skalärprodukt och norm i $\mathbb{R}^n$) Skalärprodukten av vektorerna $u, v \in \mathbb{R}^n$ definieras av

$$u \cdot v = \sum_{i=1}^{n} u_i v_i \tag{4.92}$$

Normen (längden) av vektorn $u \in \mathbb{R}^n$ definieras av

$$\|u\| = \sqrt{u\cdot u} = \sqrt{\sum_{i=1}^{n} u_i^2} \tag{4.93}$$

Vi noterar att dessa kan beräknas som matrisprodukter:

$$u\cdot v = u^T v = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \tag{4.94}$$

$$\|u\|^2 = u\cdot u = u^T u \tag{4.95}$$

Sats 4.14 (Räkneregler för skalärprodukt och norm) Låt $u, v, w \in \mathbb{R}^n$ och $\alpha \in \mathbb{R}$. Då gäller

$$u\cdot v = v\cdot u \qquad \text{(kommutativ lag)} \tag{4.96}$$

$$u\cdot(v+w) = u\cdot v + u\cdot w \qquad \text{(distributiv lag)} \tag{4.97}$$

$$(\alpha u)\cdot v = \alpha(u\cdot v) = u\cdot(\alpha v) \qquad \text{(distributiv lag)} \tag{4.98}$$

$$u\cdot u \geqslant 0, \quad u\cdot u = 0 \iff u = 0 \qquad \text{(positivt definit)} \tag{4.99}$$

$$\|\alpha u\| = |\alpha|\,\|u\| \tag{4.100}$$

Bevis. Beviset är enkelt. Vi utelämnar det. □

Obs att (4.99) innebär att den enda vektor som har normen lika med noll är nollvektorn. Om $u \neq 0$ kan vi *normera* den

$$\hat{u} = \frac{1}{\|u\|} u = \frac{u}{\|u\|} \tag{4.101}$$

och få en vektor som är parallell med u och med $\|\hat{u}\| = 1$ (enhetsvektor). (Tag $\alpha = 1/\|u\|$ i (4.100).) Med hjälp av normen kan vi också definiera *avståndet* mellan u och v:

$$\operatorname{dist}(u, v) = \|u - v\| \tag{4.102}$$

Definition 4.9 (Ortogonalitet) Om $u \cdot v = 0$ säger vi att u och v är *ortogonala*, $u \perp v$.

Sats 4.15 (Pythagoras sats)

$$u \perp v \iff \|u + v\|^2 = \|u\|^2 + \|v\|^2 \tag{4.103}$$

Bevis. Vi utvecklar kvadraten med hjälp av reglerna i sats 4.14:

$$\begin{aligned}
\|u + v\|^2 &= (u + v) \cdot (u + v) && (4.104)\\
&= u \cdot u + u \cdot v + v \cdot u + v \cdot v && (4.105)\\
&= \|u\|^2 + 2u \cdot v + \|v\|^2 && (4.106)
\end{aligned}$$

Vi ser att $u \cdot v = 0$ om och endast om $\|u + v\|^2 = \|u\|^2 + \|v\|^2$. □

Definition 4.10 (Ortogonal mängd) En *ortogonal mängd* är en mängd $S = \{u_1, \ldots, u_p\}$ där vektorerna är parvis ortogonala, dvs $u_i \cdot u_j = 0$ för $i \neq j$. Mängden är en *ortonormerad mängd* (ON mängd) om vektorerna är både ortogonala och normerade, dvs

$$u_i \cdot u_j = \begin{cases} 1 & \text{om } i = j \\ 0 & \text{om } i \neq j \end{cases} \tag{4.107}$$

Ortonormeringsvillkoret kan skrivas $u_i \cdot u_j = \delta_{ij}$ med Kroneckers symbol:

$$\delta_{ij} = \begin{cases} 1 & \text{om } i = j \\ 0 & \text{om } i \neq j \end{cases} \tag{4.108}$$

Exempel 4.11 (Ortogonal mängd) Låt $S = \{u_1, u_2, u_3\}$ med

$$u_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \; u_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \; u_3 = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix} \tag{4.109}$$

Vi beräknar

$$u_1 \cdot u_1 = 3^2 + 1^2 + 1^2 = 11 \tag{4.110}$$
$$u_1 \cdot u_2 = 3 \cdot (-1) + 2 \cdot 1 + 1 \cdot 1 = 0 \tag{4.111}$$
$$u_1 \cdot u_3 = \cdots = 0 \tag{4.112}$$
$$u_2 \cdot u_2 = \cdots = 6 \tag{4.113}$$
$$u_2 \cdot u_3 = \cdots = 0 \tag{4.114}$$
$$u_3 \cdot u_3 = \cdots = 66 \tag{4.115}$$

Eller på matrisform med $U = [u_1 \; u_2 \; u_3]$:

$$U^T U = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}^T \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} = \begin{bmatrix} u_1^T \\ u_2^T \\ u_3^T \end{bmatrix} \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} \tag{4.116}$$

$$= \begin{bmatrix} 3 & 1 & 1 \\ -1 & 2 & 1 \\ -1 & -4 & 7 \end{bmatrix} \begin{bmatrix} 3 & -1 & -1 \\ 1 & 2 & -4 \\ 1 & 1 & 7 \end{bmatrix} \begin{bmatrix} 11 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 66 \end{bmatrix} \tag{4.117}$$

(praktiskt vid datorberäkning). Vi ser att S är en ortogonal mängd. Efter normering

$$\hat{u}_1 = \frac{1}{\sqrt{11}} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \; \hat{u}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \; \hat{u}_3 = \frac{1}{\sqrt{66}} \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix} \tag{4.118}$$

får vi en ON mängd $\{\hat{u}_1, \hat{u}_2, \hat{u}_3\}$ med matrisen $\hat{U} = [\hat{u}_1 \; \hat{u}_2 \; \hat{u}_3]$ och $\hat{U}^T \hat{U} = I$.

Sats 4.16 (En ortogonal mängd är linjärt oberoende) Om $S = \{u_1, \ldots, u_p\}$ är en ortogonal mängd där alla $u_i \neq 0$, så är den linjär oberoende, dvs en bas för underrummet $W = \operatorname{span}\{u_1, \ldots, u_p\}$. Speciellt gäller detta om S är en ON mängd (där alla $\|u_i\| = 1 \neq 0$).

Bevis. Vi löser vektorekvationen $\sum_{j=1}^{p} c_j u_j = 0$. Multiplicera ekvationen skalärt med u_i och använd att $u_j \cdot u_i = 0$ om $i \neq j$:

$$0 = 0 \cdot u_i = \left(\sum_{j=1}^{p} c_j u_j \right) \cdot u_i = \sum_{j=1}^{p} c_j (u_j \cdot u_i) = c_i (u_i \cdot u_i) \tag{4.119}$$

Eftersom $u_i \cdot u_i = \|u_i\|^2 \neq 0$ får vi $c_i = 0$. Endast den triviala lösningen, alltså linjärt oberoende. Obs hur ortogonalitet underlättar räkningen: alla termer utom en blir noll. □

Definition 4.11 (Ortogonal bas) En *ortogonal bas* är en bas $B = \{u_1, \ldots, u_p\}$ som också är en ortogonal mängd. Basen är en *ortonormerad bas* (ON bas) om den också är en ortonormerad mängd.

Standardbasen $E = \{e_1, \ldots, e_n\}$ är en ON bas för $\mathbb{R}^n$.

Sats 4.17 (Komponenter i ortogonal bas) Låt $B = \{u_1, \ldots, u_p\}$ vara en ortogonal bas för underrummet $W = \operatorname{span}\{u_1, \ldots, u_p\}$. Om $w \in W$, dvs

$$w = \sum_{j=1} c_j u_j \tag{4.120}$$

så ges komponenterna av

$$c_i = \frac{w \cdot u_i}{u_i \cdot u_i} \tag{4.121}$$

Om $B = \{\hat{u}_1, \ldots, \hat{u}_p\}$ är en ON bas, så har vi istället

$$w = \sum_{j=1} c_j \hat{u}_j \tag{4.122}$$

med komponenterna

$$c_i = w \cdot \hat{u}_i \tag{4.123}$$

Bevis. Vi multiplicerar $w = \sum_{j=1}^{p} c_j u_j$ skalärt med u_i och använder att $u_j \cdot u_i = 0$ om $i \neq j$:

$$w \cdot u_i = \left(\sum_{j=1}^{p} c_j u_j \right) \cdot u_i = \sum_{j=1}^{p} c_j (u_j \cdot u_i) = c_i (u_i \cdot u_i) \tag{4.124}$$

Eftersom $u_i \cdot u_i = \|u_i\|^2 \neq 0$ får vi $c_i = (w \cdot u_i)/(u_i \cdot u_i)$. Här är $u_i \cdot u_i = \|u_i\|^2 = 1$ om basen är ortonormerad. □

Exempel 4.12 (Ortogonal mängd, igen) Vi återvänder till exempel 4.11. Vi vet att $B = \{u_1, u_2, u_3\}$ är en ortogonal mängd och därför en bas för $\mathbb{R}^3$ enligt sats 4.16. Vi beräknar komponenterna för vektorn $x = [6\ 1\ \ -8]^T$ i denna bas. Basvektormatrisen är $U = [u_1\ u_2\ u_3]$. Enligt metoden från sats 4.8, skulle vi lösa ekvationssystemet $U[x]_B = [x]_E$, dvs $Uc = x$, med hjälp av Gauss-elimination. Men nu ska vi dra nytta av ortogonaliteten och använda sats 4.17. Vi beräknar skalärprodukterna:

$$x \cdot u_1 = 11,\ x \cdot u_2 = -12,\ x \cdot u_3 = -66 \tag{4.125}$$

Vi har redan

$$u_1 \cdot u_1 = 11,\ u_2 \cdot u_2 = 6,\ u_3 \cdot u_3 = 66 \tag{4.126}$$

Dessa skalärprodukter kan enkelt beräknas på dator med $U^T x$ respektive $U^T U$. Vi får alltså

$$x = \frac{11}{11}u_1 + \frac{-12}{6}u_2 + \frac{-66}{66}u_3 = u_1 - 2u_2 - u_3 \tag{4.127}$$

Med den ortonormerade basen $\{\hat{u}_1, \hat{u}_2, \hat{u}_3\}$ får vi istället

$$x \cdot \hat{u}_1 = \frac{11}{\sqrt{11}},\ x \cdot \hat{u}_2 = \frac{-12}{\sqrt{6}},\ x \cdot \hat{u}_3 = \frac{-66}{\sqrt{66}} \tag{4.128}$$

och

$$x = \frac{11}{\sqrt{11}}\hat{u}_1 + \frac{-12}{\sqrt{6}}\hat{u}_2 + \frac{-66}{\sqrt{66}}\hat{u}_3 = \frac{11}{\sqrt{11}}\frac{u_1}{\sqrt{11}} + \frac{-12}{\sqrt{6}}\frac{u_2}{\sqrt{6}} + \frac{-66}{\sqrt{66}}\frac{u_3}{\sqrt{66}} \tag{4.129}$$

Den ortonormerade basen ger krångliga handräkningar men är mycket enklare att använda när vi gör teoretiska resonemang och datorberäkningar, som vi ska ofta se i fortsättningen.

Sats 4.18 (Ortonormerade kolonner) Matrisen $\hat{U} = [\hat{u}_1\ \cdots\ \hat{u}_p] \in \mathbb{R}^{n\times p}$ har ortonormerade kolonner om och endast om $\hat{U}^T\hat{U} = I \in \mathbb{R}^{p\times p}$. Formeln (4.123) kan då skrivas

$$c = \hat{U}^T w \tag{4.130}$$

Bevis. Enligt rad-kolonn-regeln får vi

$$(\hat{U}^T\hat{U})_{ij} = \hat{u}_i^T\hat{u}_j = \hat{u}_i \cdot \hat{u}_j = \delta_{ij} \tag{4.131}$$

dvs $\hat{U}^T\hat{U} = I \in \mathbb{R}^{p\times p}$. Om $w = \hat{U}c$ som i (4.122), så får vi

$$\hat{U}^T w = \hat{U}^T\hat{U}c = Ic = c \tag{4.132}$$

vilket är (4.130). □

Definition 4.12 (Ortogonalt komplement) Låt W vara ett underrum till $\mathbb{R}^n$. Vektorn $x \in \mathbb{R}^n$ är ortogonal mot W, $v \perp W$, om $x \cdot w = 0$ för alla $w \in W$. Mängden

$$W^\perp = \{x \in \mathbb{R}^n \mid x \cdot w = 0 \text{ för alla } w \in W\} \tag{4.133}$$

kallas *ortogonala komplementet* till W.

$W^\perp$ är också ett underrum: om $z_1, z_2 \in W^\perp$ så $(\alpha z_1 + \beta z_2) \cdot w = \alpha z_1 \cdot w + \beta z_2 \cdot w = 0$.

Sats 4.19 (Ortogonal projektion) Låt W vara ett underrum till $\mathbb{R}^n$. Varje vektor $x \in \mathbb{R}^n$ kan skrivas på ett *unikt* sätt som

$$x = w + z \quad \text{där } w \in W \text{ och } z \in W^\perp \tag{4.134}$$

Detta kallas en *ortogonal uppdelning* av vektorn. Om $\{u_1, \dots, u_p\}$ är en ortogonal bas för W, så ges w av

$$w = \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \cdots + \frac{x \cdot u_p}{u_p \cdot u_p} u_p = \sum_{i=1}^{p} \frac{x \cdot u_i}{u_i \cdot u_i} u_i \tag{4.135}$$

och $z = x - w$. Om $\{\hat{u}_1, \dots, \hat{u}_p\}$ är en ON bas för W, så har vi istället

$$w = (x \cdot \hat{u}_1)\hat{u}_1 + \cdots + (x \cdot \hat{u}_p)\hat{u}_p = \sum_{i=1}^{p} (x \cdot \hat{u}_i)\hat{u}_i \tag{4.136}$$

Vektorn w kallas *ortogonala projektionen* av x på underrummet W.

Bevis. Definiera w som i (4.135) och $z = x - w$. Då är $w \in W$ och

$$z \cdot u_j = (x - w) \cdot u_j = x \cdot u_j - \Big(\sum_{i=1}^{p} \frac{x \cdot u_i}{u_i \cdot u_i} u_i \Big) \cdot u_j \tag{4.137}$$

$$= x \cdot u_j - \sum_{i=1}^{p} \frac{x \cdot u_i}{u_i \cdot u_i} (u_i \cdot u_j) = x \cdot u_j - x \cdot u_j = 0 \tag{4.138}$$

Det betyder att $z \in W^\perp$, dvs vi har en ortogonal uppdelning. För att visa att den är unik antar vi att det finns även en uppdelning

$$x = \tilde{w} + \tilde{z} \quad \text{där } \tilde{w} \in W \text{ och } \tilde{z} \in W^\perp \tag{4.139}$$

Men då fås $x = w + z = \tilde{w} + \tilde{z}$, dvs $w - \tilde{w} = \tilde{z} - z$. Men $w - \tilde{w} \in W$ och $\tilde{z} - z \in W^\perp$. Vektorn $w - \tilde{w} = \tilde{z} - z$ tillhör alltså både W och $W^\perp$. Den enda vektor som gör det är nollvektorn. Alltså: $\tilde{w} = w$ och $\tilde{z} = z$. □

Definition 4.13 (Ortogonal projektor) Låt W vara ett underrum till $\mathbb{R}^n$. För varje vektor $x \in \mathbb{R}^n$ ger sats 4.19 en unik vektor $w \in W$, den ortogonala projektionen av x på W. Detta definierar en linjär operator $P_W\colon \mathbb{R}^n \to \mathbb{R}^n$ genom $w = P_W(x)$. Den kallas den *ortogonala projektorn* på W.

Om $\{u_1, \dots, u_p\}$ är en ortogonal bas för W eller $\{\hat{u}_1, \dots, \hat{u}_p\}$ en ON bas för W, så har vi

$$P_W(x) = \sum_{i=1}^{p} \frac{x \cdot u_i}{u_i \cdot u_i} u_i = \sum_{i=1}^{p} (x \cdot \hat{u}_i)\hat{u}_i \tag{4.140}$$

Det är klart från formeln i (4.140) att funktionen P_W är linjär. Det är också klart att att $P_W(w) = w$ om $w \in W$. Det betyder att värderummet är $\mathcal{R}(P_W) = W$.

Exempel 4.13 (Ortogonal projektion) Vi återvänder till exempel 4.12. Vi projicerar $x = \begin{bmatrix} 6 & 1 & -8 \end{bmatrix}^T$ på $W = \text{span}\{u_1, u_2\}$. Eftersom basvektorerna är ortogonala har vi $W^\perp = \text{span}\{u_3\}$. Sedan tidigare har vi

$$x \cdot u_1 = 11,\ x \cdot u_2 = -12,\ u_1 \cdot u_1 = 11,\ u_2 \cdot u_2 = 6 \tag{4.141}$$

och vi får

$$w = P_W(x) = \frac{11}{11} u_1 + \frac{-12}{6} u_2 = u_1 - 2u_2 = \begin{bmatrix} 5 \\ -3 \\ -1 \end{bmatrix} \tag{4.142}$$

och

$$z = x - P_W(x) = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix} - \begin{bmatrix} 5 \\ -3 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ -7 \end{bmatrix} = -u_3 \in W^\perp = \text{span}\{u_3\} \tag{4.143}$$

Med den ortonormerade basen $\{\hat{u}_1, \hat{u}_2\}$ blir basvektormatrisen

$$\hat{U} = [\hat{u}_1\ \hat{u}_2] = \begin{bmatrix} \frac{3}{\sqrt{11}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{11}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{6}} \end{bmatrix} \quad \text{med } \hat{U}^T\hat{U} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{4.144}$$

som sats 4.18 säger. Den ortogonala projektionen blir

$$w = P_W(x) = \frac{11}{\sqrt{11}} \hat{u}_1 + \frac{-12}{\sqrt{6}} \hat{u}_2 = \frac{11}{\sqrt{11}} \frac{u_1}{\sqrt{11}} + \frac{-12}{\sqrt{6}} \frac{u_2}{\sqrt{6}} \tag{4.145}$$

Vi erinrar oss från sats 2.10 att varje linjär operator $L\colon \mathbb{R}^n \to \mathbb{R}^m$ kan skrivas på matrisform $L(x) = Ax$ med $A \in \mathbb{R}^{m \times n}$. I nästa sats använder vi ON bas för att ta fram matrisformen för projektorn $P_W\colon \mathbb{R}^n \to \mathbb{R}^n$.

Sats 4.20 (Ortogonal projektor i ON bas) Om $\{\hat{u}_1, \dots, \hat{u}_p\}$ är en ON bas för W, så har vi

$$P_W(x) = \sum_{i=1}^{p} c_i \hat{u}_i \quad \text{med } c_i = x \cdot \hat{u}_i \tag{4.146}$$

Med basvektormatrisen $\hat{U} = [\hat{u}_1 \ \cdots \ \hat{u}_p] \in \mathbb{R}^{n \times p}$ har vi

$$P_W(x) = \hat{U}\hat{U}^T x \tag{4.147}$$

dvs matrisen för projektionsoperatorn är $\hat{U}\hat{U}^T \in \mathbb{R}^{n \times n}$.

Bevis. Låt $w = P_W(x)$. På matrisform blir (4.146) då

$$w = \hat{U}c \tag{4.148}$$

Vi har

$$x = w + z \quad \text{där } z \in W^\perp \tag{4.149}$$

Vi multiplicerar med $\hat{U}^T$ och använder att $\hat{U}^T z = 0$ och $\hat{U}^T\hat{U} = I$, se sats 4.18,

$$\hat{U}^T x = \hat{U}^T w + \hat{U}^T z = \hat{U}^T w + 0 = \hat{U}^T w = \hat{U}^T\hat{U}c = c \tag{4.150}$$

och

$$\hat{U}\hat{U}^T x = \hat{U}c = w \tag{4.151}$$

dvs $P_W(x) = \hat{U}\hat{U}^T x$. □

Exempel 4.14 (Ortogonal projektion, igen) Vi återvänder till exempel 4.13. Med datorns hjälp får vi

$$\hat{U}\hat{U}^T = \begin{bmatrix} \frac{3}{\sqrt{11}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{11}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \frac{3}{\sqrt{11}} & \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{11}} \\ -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} \approx \begin{bmatrix} 0.9848 & -0.0606 & 0.1061 \\ -0.0606 & 0.7576 & 0.4242 \\ 0.1061 & 0.4242 & 0.2576 \end{bmatrix} \tag{4.152}$$

och

$$P_W(x) = \hat{U}\hat{U}^T x \approx \begin{bmatrix} 0.9848 & -0.0606 & 0.1061 \\ -0.0606 & 0.7576 & 0.4242 \\ 0.1061 & 0.4242 & 0.2576 \end{bmatrix} \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix} \approx \begin{bmatrix} 5 \\ -3 \\ -1 \end{bmatrix} \tag{4.153}$$

Vi har sett att ortogonal bas, eller hellre ortonormerad bas, underlättar många beräkningar. Men hur hittar vi en ortogonal bas? Vi ska nu se hur vi ortogonaliserar en given bas med Gram–Schmidts metod.

Ruta 4.3 (Gram–Schmidt) Gram–Schmidts ortogonaliseringsmetod skapar en ortogonalbas för underrummet W. Vi utgår från en bas $\{b_1, \dots, b_p\}$ för W och skapar en ortogonalbas $\{u_1, \dots, u_p\}$ för W genom att ortogonalisera den givna basen enligt följande metod.

steg 1. $u_1 = b_1, W_1 = \text{span}\{u_1\}$

steg 2. $u_2 = b_2 - P_{W_1}(b_2) \in W_1^\perp$
$W_2 = \text{span}\{u_1, u_2\} = \text{span}\{b_1, b_2\}$

steg 3. $u_3 = b_3 - P_{W_2}(b_3) \in W_2^\perp$
$W_3 = \text{span}\{u_1, u_2, u_3\} = \text{span}\{b_1, b_2, b_3\}$

steg i. $u_i = b_i - P_{W_{i-1}}(b_i) \in W_{i-1}^\perp$
$W_i = \text{span}\{u_1, \dots, u_i\} = \text{span}\{b_1, \dots, b_i\}$.

I steg i ska vi göra beräkningen

$$u_i = b_i - \left(\frac{b_i \cdot u_1}{u_1 \cdot u_1} u_1 + \cdots + \frac{b_i \cdot u_{i-1}}{u_{i-1} \cdot u_{i-1}} u_{i-1}\right) \tag{4.154}$$

På detta vis blir u_i en linjär kombination av $\{b_1, \dots, b_{i-1}\}$ och ortogonal mot $\{u_1, \dots, u_{i-1}\}$.
Om vi vill ha en ON bas $\{\hat{u}_1, \dots, \hat{u}_p\}$, så normerar vi vektorerna successivt. På matrisform blir (4.154) då

$$u_i = b_i - \hat{U}_{i-1}\hat{U}_{i-1}^T b_i,\ \hat{u}_i = \frac{u_i}{\|u_i\|} \quad \text{med } \hat{U}_{i-1} = [\hat{u}_1 \ \cdots \ \hat{u}_{i-1}] \tag{4.155}$$

vilket är praktiskt vid datorberäkning, se datorövning 4.4.

Ruta 4.4 (Sammanfattning: bas, ortogonal bas, ortonormerad bas) Antag att vi har en mängd av vektorer $S = \{u_1, \ldots, u_p\} \subseteq \mathbb{R}^n$. Då kan vi bilda matrisen $U = [u_1 \ \cdots \ u_p] \in \mathbb{R}^{n\times p}$, underrummet $W = \text{span}\{u_1, \ldots, u_p\} \subseteq \mathbb{R}^n$. Vi antar att $w \in W$, dvs $Uc = w$ för någon vektor $c \in \mathbb{R}^n$.

Om S är en *linjärt oberoende mängd* så gäller följande:

- S är en bas för W
- $\mathcal{N}(U) = \{0\}$
- c fås som unik lösning till $Uc = w$
- $W = \mathbb{R}^n$ om $p = n$

Om S är en *ortogonal mängd* (med alla $u_i \neq 0$) så gäller följande:

- S är en ortogonal bas för W
- c ges av $c_i = \dfrac{w \cdot u_i}{u_i \cdot u_i}$
- $w = \displaystyle\sum_{i=1}^{p} \frac{w \cdot u_i}{u_i \cdot u_i} u_i$
- $P_W(x) = \displaystyle\sum_{i=1}^{p} \frac{x \cdot u_i}{u_i \cdot u_i} u_i$

Om S är en *ortonormerad mängd* så gäller följande:

- S är en ON bas för W
- $\hat{U}^T\hat{U} = I \in \mathbb{R}^{p\times p}$ (ortonormerade kolonner)
- c ges av $c_i = w \cdot \hat{u}_i, \quad c = \hat{U}^T w$
- $w = \displaystyle\sum_{i=1}^{p} (w \cdot \hat{u}_i)\hat{u}_i$
- $P_W(x) = \displaystyle\sum_{i=1}^{p} (x \cdot \hat{u}_i)\hat{u}_i$
- $P_W(x) = \hat{U}\hat{U}^T x$
- $\|w\|^2 = \displaystyle\sum_{i=1}^{p} c_i^2, \quad \|w\|^2 = \sum_{i=1}^{p} (w \cdot \hat{u}_i)^2$ (Pythagoras sats)

4.5 Minsta kvadratmetoden

Nästa sats visar att $w = P_W(x)$ är den unika vektor i W som ligger närmast x. Men andra ord: $P_W(x)$ är den *bästa approximationen* av x med element ur W.

Sats 4.21 (Satsen om bästa approximation) Låt W vara ett underrum till $\mathbb{R}^n$ och $x \in \mathbb{R}^n$. Då gäller

$$\|x - P_W(x)\| < \|x - w\| \tag{4.156}$$

för alla $w \in W$ med $w \neq P_W(x)$.

Bevis. Tag $w \in W$ med $w \neq P_W(x)$. Vi delar upp

$$x - w = (x - P_W(x)) + (P_W(x) - w) \tag{4.157}$$

Här är $x - P_W(x) \in W^\perp$ och $P_W(x) - w \in W$, dvs de är ortogonala och Pythagoras sats, sats 4.15, ger

$$\|x - w\|^2 = \|x - P_W(x)\|^2 + \|P_W(x) - w\|^2 > \|x - P_W(x)\|^2 \tag{4.158}$$

ty $\|P_W(x) - w\|^2 > 0$ eftersom $w \neq P_W(x)$. Alltså: $\|x - w\|^2 > \|x - P_W(x)\|^2$, dvs $\|x - P_W(x)\| < \|x - w\|$. □

Vi ska nu se vi hur detta kan användas för att lösa ett inkonsistent ekvationssystem approximativt. Låt $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Vi vet att ekvationen $Ax = b$ är inkonsistent om och endast om högerledet b inte är i kolonnrummet till A, dvs $b \notin \mathcal{C}(A)$. Då finns ingen lösning men vi kan göra bästa möjliga av situationen genom att söka $x \in \mathbb{R}^n$ sådan att Ax är så nära b som möjligt, dvs så att avståndet $\|b - Ax\|$ blir minimalt. Detta är det samma som att minimera

$$\|b - Ax\|^2 = \sum_{i=1}^{m} (b - Ax)_i^2 \tag{4.159}$$

därav namnet *minsta kvadratmetoden*.

Definition 4.14 (Minsta kvadratlösning) Låt $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. En vektor $\hat{x} \in \mathbb{R}^n$ kallas *minsta kvadratlösning* till ekvationen $Ax = b$ om

$$\|b - A\hat{x}\| \leqslant \|b - Ax\| \quad \text{för alla } x \in \mathbb{R}^n \tag{4.160}$$

Vi vet från satsen om bästa approximation, sats 4.21, att den vektor i $\mathcal{C}(A)$ som ligger närmast b är den ortogonala projektionen av b på $\mathcal{C}(A)$, $\hat{b} = P_{\mathcal{C}(A)}(b)$. Det återstår att finna en vektor $\hat{x}$ sådan att $A\hat{x} = \hat{b}$. Hur det går till visas i nästa sats.

Sats 4.22 (Normalekvationerna) Låt $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Mängden av minsta kvadratlösningar till $Ax = b$ är inte tom och den är lika med lösningsmängden till *normalekvationerna*

$$A^T Ax = A^T b \tag{4.161}$$

Bevis. Satsen om bästa approximation, sats 4.21, säger att $\|b - \hat{b}\| \leqslant \|b - Ax\|$ för alla $Ax \in \mathcal{C}(A)$ om och endast om $\hat{b} = P_{\mathcal{C}(A)}(b)$. Men $\hat{b} \in \mathcal{C}(A)$, så det finns $\hat{x} \in \mathbb{R}^n$ med $A\hat{x} = \hat{b}$. Alltså är $\hat{x}$ minsta kvadratlösning om och endast om $b - A\hat{x} \perp \mathcal{C}(A)$, dvs

$$0 = (Ax)^T(b - A\hat{x}) = x^T A^T (b - A\hat{x}) = x^T(A^T b - A^T A\hat{x}) \quad \text{för alla } x \in \mathbb{R}^n \tag{4.162}$$

dvs vektorn $A^T b - A^T A\hat{x}$ är ortogonal mot alla $x \in \mathbb{R}^n$. Men den enda vektor i $\mathbb{R}^n$ som är ortogonal mot alla $x \in \mathbb{R}^n$ är nollvektorn. Alltså är (4.162) ekvivalent med att $A^T b - A^T A\hat{x} = 0$, dvs $A^T Ax = A^T b$. Detta visar att $\hat{x}$ är minsta kvadratlösning om och endast om den är lösning till (4.161). □

Observera att normalekvationerna är lösbara även om $Ax = b$ är inkonsistent. Lösningen behöver inte vara unik. Vi noterar även att koefficientmatrisen $A^T A$ är symmetrisk: $(A^T A)^T = A^T (A^T)^T = A^T A$. Symmetriska linjära ekvationssystem har vissa goda egenskaper; det ska vi återkomma till i kapitel 6.

En minsta kvadratlösning till ett inkonsistent system är inte en lösning utan en *approximativ lösning i minsta kvadratmetodens mening*. Felet i högerledet, $b - A\hat{x}$, kallas *residualen* till den approximativa lösningen $\hat{x}$. Beteckningen *normalekvationer* syftar på att $b - A\hat{x}$ är en normalvektor till $\mathcal{C}(A)$, dvs ortogonal mot $\mathcal{C}(A)$.

En situation där minsta kvadratmetoden kommer till användning är *överbestämt ekvationssystem*, dvs ett system med fler ekvationer än obekanta, $m > n$. Då är

$$\dim(\mathcal{C}(A)) = \operatorname{rang}(A) = n - \dim(\mathcal{N}(A)) < m - \dim(\mathcal{N}(A)) \leqslant m \tag{4.163}$$

enligt rangsatsen, sats 4.12, dvs $\operatorname{rang}(A) < m$, så att $\mathcal{C}(A) \neq \mathbb{R}^m$. Då är det osannolikt att $b \in \mathcal{C}(A)$, dvs ett överbestämt system är ofta inkonsistent.

Exempel 4.15 (Överbestämt ekvationssystem) Vi har ekvationssystemet

$$\begin{cases} x_1 - x_2 = 0 \\ x_1 + x_2 = 2 \\ 2x_1 - x_2 = 2 \end{cases} \tag{4.164}$$

Radreduktion av totalmatrisen ger

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 2 \\ 2 & -1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.165}$$

så att systemet är inkonsistent (överbestämt). Här har vi

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 2 & -1 \end{bmatrix}, \ b = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \tag{4.166}$$

och

$$A^T A = \begin{bmatrix} 1 & 1 & 2 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 6 & -2 \\ -2 & 3 \end{bmatrix} \tag{4.167}$$

$$A^T b = \begin{bmatrix} 1 & 1 & 2 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix} \tag{4.168}$$

Normalekvationerna blir

$$\begin{bmatrix} 6 & -2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix} \quad \text{med unik lösning } \hat{x} = \begin{bmatrix} \frac{9}{7} \\ \frac{6}{7} \end{bmatrix} \tag{4.169}$$

Residualen är

$$b - A\hat{x} = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} \frac{9}{7} \\ \frac{6}{7} \end{bmatrix} = \begin{bmatrix} -\frac{3}{7} \\ -\frac{1}{7} \\ \frac{2}{7} \end{bmatrix}, \ \|b - A\hat{x}\| = \frac{\sqrt{14}}{7} \tag{4.170}$$

4.6 Skalärproduktrum

Vi avslutar kapitlet med presentera skalärprodukt i allmänt vektorrum.

Definition 4.15 (Skalärproduktrum) Låt V vara ett vektorrum. En *skalärprodukt* på V är en reellvärd funktion som till varje par $u, v \in V$ tilldelar ett tal $\langle u, v\rangle \in \mathbb{R}$, sådan att

1. $\langle u, v\rangle = \langle v, u\rangle$ (kommutativ lag) (4.171)
2. $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ (distributiv lag) (4.172)
3. $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$ (distributiv lag) (4.173)
4. $\langle u, u\rangle > 0 \quad \text{om } u \neq 0$ (positivt definit) (4.174)

Ett vektorrum med skalärprodukt kallas *skalärproduktrum* eller *Euklidiskt rum*. Eftersom skalärprodukten är positivt definit kan vi definera en en *norm* i V enligt

$$\|v\| = \sqrt{\langle v, v\rangle} \tag{4.175}$$

Det är klart att $\langle u, u\rangle = 0$ om $u = 0$ (tag $\alpha = 0$ i (4.173)). Enligt (4.174) har vi då $\langle u, u\rangle \geqslant 0$ och nollvektorn är den enda vektor för vilken $\|u\|^2 = \langle u, u\rangle = 0$. Det är klart att normen uppfyller

$$\|\alpha v\|^2 = \langle \alpha v, \alpha v\rangle = \alpha^2\langle v, v\rangle = \alpha^2\|v\|^2 \tag{4.176}$$

dvs

$$\|\alpha v\| = |\alpha|\,\|v\| \tag{4.177}$$

Vi har redan två exempel på skalärproduktrum:

- $V = \{$ geometriska vektorer $\}$, $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = |\boldsymbol{u}|\,|\boldsymbol{v}|\cos(\theta)$, $\|\boldsymbol{u}\| = |\boldsymbol{u}|$ (se sats 1.3)
- $V = \mathbb{R}^n$, $\langle u, v\rangle = u^T v$, $\|u\| = \sqrt{u^T u}$ (se sats 4.14)

Nu kan vi använda skalärprodukt även i vissa andra rum, till exempel i funktionsrum.

Exempel 4.16 (Skalärprodukt i funktionsrum) Vi tar $V = C([a, b])$ och definierar

$$\langle f, g\rangle = \int_a^b f(x)g(x)\,\mathrm{d}x \tag{4.178}$$

och

$$\|f\| = \sqrt{\int_a^b f(x)^2\,\mathrm{d}x} \tag{4.179}$$

Egenskaperna 1, 2, 3 i definition 4.15 är självklara på grund av räkneregler för integral. För att bevisa punkt 4 noterar vi först att

$$\langle f, f\rangle = \int_a^b f(x)^2\,\mathrm{d}x \geqslant 0 \quad \text{för alla } f \in V \tag{4.180}$$

Om

$$\langle f, f\rangle = \int_a^b f(x)^2\,\mathrm{d}x = 0 \tag{4.181}$$

så måste

$$f(x) = 0 \quad \text{för alla } x \in [a, b] \tag{4.182}$$

för om $f(x_0) \neq 0$ så är $f(x) \neq 0$ i något (litet) intervall kring x_0 eftersom f är kontinuerlig. Men då blir $\int_a^b f(x)^2\,\mathrm{d}x > 0$ vilket är en motsägelse till (4.181). Men (4.182) är nollfunktionen, $f = 0$. Alltså: $\langle f, f\rangle = 0$ medför $f = 0$, dvs skalärprodukten är positivt definit.

Allt som vi redan gjort som bygger på skalärprodukt gäller nu också för allmänt skalärproduktrum. Till exempel

- ortogonalitet: $u \perp v \iff \langle u, v\rangle = 0$
- Pythagoras sats, se sats 4.15:

$$u \perp v \iff \|u + v\|^2 = \|u\|^2 + \|v\|^2 \tag{4.183}$$

- ortogonal mängd, ortonormerad (ON) mängd
- ortogonal projektion på $W = \text{span}\{u_1, \dots, u_p\}$:

$$P_W(v) = \sum_{i=1}^{p} \frac{\langle v, u_i\rangle}{\langle u_i, u_i\rangle} u_i \quad \text{om ortogonal mängd} \tag{4.184}$$

$$P_W(v) = \sum_{i=1}^{p} \langle v, \hat{u}_i\rangle \hat{u}_i \quad \text{om ON mängd} \tag{4.185}$$

- bästa approximation av v med element ur W ges av $w = P_W(v)$
- Gram–Schmidts ortogonaliseringsmetod

Vi ska nu presentera två viktiga olikheter: Cauchy–Schwarz olikhet och triangelolikheten.

Att absolutbeloppet av skalärprodukten är mindre än eller lika med produkten av vektorernas längder är självklart för geometriska vektorer eftersom $|\cos(\theta)| \leqslant 1$:

$$|\boldsymbol{u} \cdot \boldsymbol{v}| = |\boldsymbol{u}|\,|\boldsymbol{v}|\,|\cos(\theta)| \leqslant |\boldsymbol{u}|\,|\boldsymbol{v}| \tag{4.186}$$

Nästa sats säger att detta gäller för skalärprodukt i allmänhet.

Sats 4.23 (Cauchy–Schwarz olikhet)

$$|\langle u, v\rangle| \leqslant \|u\|\,\|v\| \tag{4.187}$$

Bevis. Om $u = 0$ är olikheten trivialt sann (båda leden är noll). Antag att $u \neq 0$. Då gäller enligt reglerna i definition 4.15 att

$$0 \leqslant \|v - \alpha u\|^2 = \langle v - \alpha u, v - \alpha u\rangle = \langle v, v - \alpha u\rangle + \langle -\alpha u, v - \alpha u\rangle \tag{4.188}$$

$$= \langle v, v\rangle - \alpha\langle v, u\rangle - \alpha\langle u, v\rangle + \alpha^2\langle u, u\rangle = \|v\|^2 - 2\alpha\langle u, v\rangle + \alpha^2\|u\|^2 \tag{4.189}$$

Men skalären α är godtycklig och vi väljer $\alpha = \langle v, u\rangle/\|u\|^2$, vilket leder till

$$0 \leqslant \|v\|^2 - 2\frac{\langle u, v\rangle^2}{\|u\|^2} + \frac{\langle u, v\rangle^2}{\|u\|^2} \tag{4.190}$$

så att

$$\langle u, v\rangle^2 \leqslant \|u\|^2\|v\|^2 \tag{4.191}$$

vilket bevisar (4.187). □

Notera att vårt val av α innebär att

$$\alpha u = \frac{\langle v, u\rangle}{\|u\|^2} u = \langle v, \hat{u}\rangle \hat{u} = P_W(v) \tag{4.192}$$

är den ortogonala projektionen av v på $W = \operatorname{span}\{u\}$ och $v = \alpha u + (v - \alpha u)$ är en ortogonal uppdelning av v.

Sats 4.24 (Triangelolikheten)

$$\|u + v\| \leqslant \|u\| + \|v\| \tag{4.193}$$

Bevis. Vi utvecklar kvadraten som i beviset av Pythagoras sats, sats 4.15:

$$\|u+v\|^2 = \langle u+v, u+v\rangle = \langle u,u\rangle + 2\langle u,v\rangle + \langle v,v\rangle = \|u\|^2 + 2\langle u,v\rangle + \|v\|^2 \tag{4.194}$$

Men Cauchy–Schwarz olikhet, sats 4.23, ger $\langle u,v\rangle \leqslant |\langle u,v\rangle| \leqslant \|u\|\|v\|$ så att

$$\|u+v\|^2 \leqslant \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = \big(\|u\| + \|v\|\big)^2 \tag{4.195}$$

vilket bevisar (4.193). □

Exempel 4.17 (Cauchy–Schwarz olikhet för summa och integral) I $\mathbb{R}^n$ blir (4.187)

$$\Big|\sum_{i=1}^n u_i v_i\Big| \leqslant \sqrt{\sum_{i=1}^n u_i^2}\sqrt{\sum_{i=1}^n v_i^2} \tag{4.196}$$

och i $C([a,b])$

$$\Big|\int_a^b f(x)g(x)\,\mathrm{d}x\Big| \leqslant \sqrt{\int_a^b f(x)^2\,\mathrm{d}x}\sqrt{\int_a^b g(x)^2\,\mathrm{d}x} \tag{4.197}$$

Dessa olikheter är viktiga hjälpmedel i linjär algebra och matematisk analys.

Exempel 4.18 (Ortogonala funktioner) Låt $V = C([-\pi,\pi])$ och $S = \{u_1, u_2\}$ med $u_1(x) = \cos(x)$, $u_2(x) = \sin(x)$. Vi beräknar skalärprodukten

$$\langle u_1, u_2\rangle = \int_{-\pi}^{\pi} \cos(x)\sin(x)\,\mathrm{d}x = \tfrac{1}{2}\int_{-\pi}^{\pi} \sin(2x)\,\mathrm{d}x = \tfrac{1}{2}\Big[-\tfrac{1}{2}\cos(2x)\Big]_{-\pi}^{\pi} = 0 \tag{4.198}$$

(Vi kan också säga att integralen blir noll därför att integranden är en udda funktion.) S är alltså en ortogonal mängd. För att få en ON mängd normerar vi funktionerna. Vi beräknar normerna

$$\|u_1\|^2 = \int_{-\pi}^{\pi} \cos(x)^2 \,\mathrm{d}x = \tfrac{1}{2}\int_{-\pi}^{\pi} (1+\cos(2x)) \,\mathrm{d}x = \tfrac{1}{2}\Big[1 - \tfrac{1}{2}\sin(2x)\Big]_{-\pi}^{\pi} = \pi \tag{4.199}$$

$$\|u_2\|^2 = \int_{-\pi}^{\pi} \sin(x)^2 \,\mathrm{d}x = \tfrac{1}{2}\int_{-\pi}^{\pi} (1-\cos(2x)) \,\mathrm{d}x = \tfrac{1}{2}\Big[1 + \tfrac{1}{2}\sin(2x)\Big]_{-\pi}^{\pi} = \pi \tag{4.200}$$

De normerade funktionerna är

$$\hat{u}_1(x) = \frac{1}{\|u_1\|}u_1(x) = \frac{1}{\sqrt{\pi}}\cos(x), \quad \hat{u}_2(x) = \frac{1}{\|u_2\|}u_2(x) = \frac{1}{\sqrt{\pi}}\sin(x) \tag{4.201}$$

och $\{\hat{u}_1, \hat{u}_2\}$ är en ON mängd.

Exempel 4.19 (Vinkel) I analogi med (4.89) kan vi definiera vinkeln θ mellan $u, v \in V$ $(u, v \neq 0)$ genom formeln

$$\cos(\theta) = \frac{\langle u, v\rangle}{\|u\|\,\|v\|}, \quad 0 \leqslant \theta \leqslant \pi \tag{4.202}$$

Definitionen är korrekt eftersom kvoten i högerledet är mellan -1 och 1 enligt Cauchy–Schwarz olikhet.
Till exempel får vi vinkeln mellan $f(x) = x$ och $g(x) = x^2$ i $V = C([0,1])$ genom

$$\cos(\theta) = \frac{\langle f, g\rangle}{\|f\|\,\|g\|} = \frac{\int_0^1 x \cdot x^2 \,\mathrm{d}x}{\sqrt{\int_0^1 x^2 \,\mathrm{d}x}\sqrt{\int_0^1 x^4 \,\mathrm{d}x}} = \frac{\frac{1}{4}}{\sqrt{\frac{1}{3}}\sqrt{\frac{1}{5}}} = \frac{\sqrt{15}}{4} \tag{4.203}$$

$$\theta = \arccos\left(\frac{\sqrt{15}}{4}\right) \tag{4.204}$$

Exempel 4.20 (Bästa approximation med polynom) Vi bestämmer bästa approximationen av $f(x) = x^2$ med polynomen $\{b_1, b_2\} = \{x, 1-x\}$ i $C([0,1])$. Vi använder satsen om bästa approximation, sats 4.21. För att kunna projicera f på $W = \mathrm{span}\{b_1, b_2\}$ måste vi först ortogonalisera. Gram–Schmidt ger

$$u_1(x) = b_1(x) = x \tag{4.205}$$

$$u_2 = b_2 - \frac{\langle b_2, u_1\rangle}{\langle u_1, u_1\rangle}u_1 \tag{4.206}$$

$$u_2(x) = (1-x) - \frac{\int_0^1 (1-x)x \,\mathrm{d}x}{\int_0^1 x^2 \,\mathrm{d}x}x \tag{4.207}$$

$$= 1 - x - \frac{\frac{1}{6}}{\frac{1}{3}} = 1 - \tfrac{3}{2}x \tag{4.208}$$

Vi har nu en ortogonal bas för $W = \text{span}\{u_1, u_2\} = \text{span}\{b_1, b_2\}$. Den bästa approximationen ges av den ortogonala projektionen på W:

$$w = P_W(f) = \frac{\langle f, u_1\rangle}{\langle u_1, u_1\rangle} u_1 + \frac{\langle f, u_2\rangle}{\langle u_2, u_2\rangle} u_2 \tag{4.209}$$

$$w(x) = \frac{\int_0^1 x^2 \cdot x \, \mathrm{d}x}{\int_0^1 x^2 \, \mathrm{d}x} x + \frac{\int_0^1 x^2 (1 - \frac{3}{2}x) \, \mathrm{d}x}{\int_0^1 (1 - \frac{3}{2}x)^2 \, \mathrm{d}x} (1 - \tfrac{3}{2}x) \tag{4.210}$$

$$= \frac{\frac{1}{4}}{\frac{1}{3}} x + \frac{-\frac{1}{24}}{\frac{1}{4}} (1 - \tfrac{3}{2}x) = \tfrac{3}{4}x - \tfrac{1}{6}(1 - \tfrac{3}{2}x) = -\tfrac{1}{6} + \tfrac{9}{4}x \tag{4.211}$$

Approximation med polynom av detta slag är grunden för *finita elementmetoden* som vi ska introducera i del IV.

Övningar

4.1 Vektorrum och underrum

Övning 4.1 Bestäm nollrummet till L och lös ekvationen $Lu = f$ på formen $u = u_\mathrm{p} + u_\mathrm{h}$.
(a) $L = \mathrm{D}^2 + 4$, $f(x) = x^2 + 1$ (b) $L = \mathrm{D}^2 - 1$, $f(x) = x$
(c) $L = \mathrm{D}$, $f(x) = x$ (d) $L = \mathrm{D} + 1$, $f(x) = 1$

Övning 4.2 Visa att operatorn $L\colon V \to W$ är linjär.
(a) $Lu = \int_0^1 u(s)\,\mathrm{d}s$, $V = C(\mathbb{R})$, $W = \mathbb{R}$
(b) $(Lu)(x) = \int_0^x u(s)\,\mathrm{d}s$, $V = W = C(\mathbb{R})$

Övning 4.3 Är mängden $\{u_i\}_{i=1}^p$ linjärt oberoende?
(a) $u_1(x) = \exp(x),\ u_2(x) = \exp(-x)$
(b) $u_1(x) = \cos(x),\ u_2(x) = \cos(2x)$
(c) $u_1(x) = x,\ u_2(x) = (x+1)^2,\ u_3(x) = (x-1)^2$

4.2 Baser och komponenter

Övning 4.4 Bestäm komponenterna av vektorn x i basen $B = \{b_i\}_{i=1}^p$.
$$b_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix},\ b_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix},\ x = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Övning 4.5 Är mängden $B = \{b_i\}_{i=1}^p$ är en bas för $\mathbb{R}^3$? Bestäm i så fall komponenterna för vektorn x i den givna basen.
(a) $b_1 = \begin{bmatrix} 6 \\ 3 \\ 3 \end{bmatrix}, b_2 = \begin{bmatrix} -4 \\ 1 \\ -5 \end{bmatrix}, b_3 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}, x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$
(b) $b_1 = \begin{bmatrix} -2 \\ 5 \\ 4 \end{bmatrix}, b_2 = \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix}, b_3 = \begin{bmatrix} 4 \\ -7 \\ -1 \end{bmatrix}, x = \begin{bmatrix} 1 \\ \frac{5}{2} \\ 1 \end{bmatrix}$
(c) $b_1 = \begin{bmatrix} -2 \\ 5 \\ 4 \end{bmatrix}, b_2 = \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix}, b_3 = \begin{bmatrix} 4 \\ -7 \\ -1 \end{bmatrix}, b_4 = \begin{bmatrix} 3 \\ 8 \\ 1 \end{bmatrix}, x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

4.3 Dimension och rang

Övning 4.6 Bestäm en bas för nollrummet och en bas för kolonnrummet till A och bestäm rangen för A.

(a) $A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 2 & -2 \\ -1 & -1 & -3 \end{bmatrix}$ (b) $A = \begin{bmatrix} 2 & 4 & 5 & -1 \\ -6 & -7 & -5 & 4 \end{bmatrix}$

(c) $A = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 1 & 1 & 2 & 2 \\ 1 & 0 & -5 & 0 \\ 1 & 0 & 1 & 3 \end{bmatrix}$

4.4 Skalärprodukt

Övning 4.7 Beräkna $\|u\|$, $\|v\|$ och $u \cdot v$.

(a) $u = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, v = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$, (b) $u = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, v = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

Övning 4.8 u och v är ortogonala med $\|u\| = 5$, $\|v\| = 12$. Beräkna $\|u - v\|$ genom att använda Pythagoras sats.

Övning 4.9 Låt $u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$ och $W = \text{span}\{u_1, u_2\}$. Beräkna $P_W(x)$.

(a) $x = \begin{bmatrix} 5 \\ 1 \\ 5 \\ 1 \end{bmatrix}$ (b) $x = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$ (c) $x = \begin{bmatrix} -2 \\ 2 \\ -2 \\ 2 \end{bmatrix}$

Övning 4.10 Bestäm en ON bas för $\text{span}\{b_1, \ldots, b_p\}$.

(a) $b_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, b_2 = \begin{bmatrix} 1 \\ -2 \\ 1 \\ 1 \end{bmatrix}, b_3 = \begin{bmatrix} 1 \\ 3 \\ 2 \\ -4 \end{bmatrix}$

(b) $b_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, b_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ (c) $b_1 = \begin{bmatrix} -5 \\ 0 \\ 0 \end{bmatrix}, b_2 = \begin{bmatrix} 4 \\ -7 \\ 0 \end{bmatrix}$

4.5 Minsta kvadratmetoden

Övning 4.11 Lös approximativt med minsta kvadratmetoden.

(a) $\begin{cases} x_1 - x_2 = 0 \\ x_1 - 2x_2 = 2 \\ 2x_1 + x_2 = 1 \end{cases}$ (b) $\begin{cases} x_1 + x_2 = 2 \\ x_1 - x_2 = 2 \\ x_1 - 2x_2 = 1 \\ 2x_1 - x_2 = 3 \end{cases}$ (c) $\begin{cases} x_1 - x_3 + x_2 = 1 \\ 2x_1 - x_2 = 0 \\ 2x_2 + 3x_1 = 2 \\ x_1 + x_2 + x_3 = 4 \end{cases}$

(d) $\begin{cases} x_1 + x_2 = 2 \\ x_1 - x_2 = 0 \\ 2x_1 - x_2 = 1 \end{cases}$ (e) $\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 + x_2 + x_3 = 2 \end{cases}$

4.6 Skalärproduktrum

Övning 4.12 Beräkna vinkeln mellan $u, v \in C([0,1])$.
(a) $u(x) = x, v(x) = 2x - x^2$ (b) $u(x) = 1 - 2x, v(x) = x^2 - x + 1$
(c) $u(x) = x - 1, v(x) = x^2$

Övning 4.13 Beräkna normen av $f(x) = x^k$ i $C([-1,1])$.

Övning 4.14 Ortonormera polynomen $1, x, x^2$ i $C([0,1])$.

Övning 4.15 (Legendre-polynom) Ortogonalisera polynomen $1, x, x^2, x^3$ i $C([-1,1])$, dvs bestäm en ortogonal mängd av polynom $\{P_k\}_{k=0}^3$ med $\mathrm{grad}(P_k) = k$. Normera dem så att $P_k(1) = 1$ istället för $\|P_k\| = 1$.

Problem

4.1 Vektorrum och underrum

Problem 4.1 Bevisa sats 4.6.

4.4 Skalärprodukt

Problem 4.2 Visa parallellogramlagen $\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$.

Problem 4.3 Låt $S = \{u_1, \ldots, u_n\}$ vara en ON bas för $\mathbb{R}^n$. Bevisa Pythagoras sats

$$\|v\|^2 = \sum_{i=1}^{n} c_i^2, \quad \text{om } v = \sum_{i=1}^{n} c_i u_i \tag{4.212}$$

4.6 Skalärproduktrum

Problem 4.4 Bevisa $\sum_{i=1}^{n} u_i \leqslant \sqrt{n}\sqrt{\sum_{i=1}^{n} u_i^2}$.

Datorövningar

Datorövning 4.1 Ortogonalisering. MATLAB-funktionen `U = orth(A)` skapar en matris `U` vars kolonner är en ortonormerad bas för kolonnrummet för matrisen `A`. Prova med några matriser, t ex,

$$A = \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix}$$

Kolla att $U^T U = I$.

Datorövning 4.2 Ortogonal projektion. Skriv en MATLAB-funktion `P=ortoproj(A)` som skapar matrisen för den ortogonala projektionen på kolonnrummet till matrisen A. Testa med matrisen A i förra uppgiften. Projicera vektorerna

$$v = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \quad w = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad u = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 5 \end{bmatrix}$$

Datorövning 4.3 Gram–Schmidts ortogonaliseringmetod 1. Skriv en MATLAB-funktion `U=gram_schmidt1(A)` som använder Gram–Schmidts metod med (4.154) för att skapa en matris `U` vars kolonner är en ortogonal bas för kolonnrummet till matrisen `A`.

Datorövning 4.4 Gram–Schmidts ortogonaliseringmetod 2. Skriv en MATLAB-funktion `U=gram_schmidt2(A)` som använder Gram–Schmidts metod med (4.155) för att skapa en matris `U` vars kolonner är en ortonormerad bas för kolonnrummet till matrisen `A`.

5. Egenvärdesproblem

Vi studerar egenvärdesproblem, dvs ekvationer av formen $Ax = \lambda x$, där matrisen A är kvadratisk och där vi söker både vektorn x och talet λ. På detta sätt kan vi (oftast) skapa en ny bas (ibland ortogonal bas) för $\mathbb{R}^n$ i vilken matrisen A får en speciellt enkel form: diagonal matris.

5.1 Egenvärden och egenvektorer

Egenvärdesproblem $Ax = \lambda x$ studerar vi för kvadratiska matriser. Vi börjar med ett motiverande exempel.

Exempel 5.1 (Egenvärde och egenvektor) Med matrisen $A = \begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix}$ prövar vi avbildningen $v \mapsto Av$. Till exempel

$$w = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto Aw = \begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix} \tag{5.1}$$

$$v = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mapsto Av = \begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix} = 5 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{5.2}$$

Vektorn $v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ är speciell: den avbildas endast genom skalning. En sådan vektor kallas egenvektor och skalningsfaktorn kallas egenvärde till matrisen.

Definition 5.1 (Egenvärdesproblem) Låt $A \in \mathbb{R}^{n\times n}$. Om det finns ett tal $\lambda \in \mathbb{R}$ och en vektor $x \in \mathbb{R}^n$ med $x \neq 0$ sådana att $Ax = \lambda x$, så är λ ett *egenvärde* till A och x en *egenvektor* hörande till λ. Paret (λ, x) kallas ett *egenpar* till A. Problemet att söka egenvärden och egenvektorer kallas *egenvärdesproblem*. Mängden av alla egenvärden till A kallas *spektrum* till A.

Notera att $x \neq 0$, men λ kan vara 0. Om x är egenvektor till λ, så är också αx också en egenvektor för alla skalärer $\alpha \neq 0$.

Hur löser vi egenvärdesproblem? Antag att $Ax = \lambda x$, $x \neq 0$. Detta är ekvivalent med att $Ax - \lambda x = 0$, dvs att den homogena ekvationen $(A - \lambda I)x = 0$ har icke-trivial lösning. Vi vet från avsnitt 2.3 hur detta kan undersökas genom radreducering av matrisen[1] $A - \lambda I$.

Exempel 5.2 (Egenvärdesproblem) Vi återvänder till exempel 5.1. Vi har

$$A - \lambda I = \begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 4-\lambda & 1 \\ 3 & 2-\lambda \end{bmatrix} \tag{5.3}$$

Radreducering ger (obs att $4 - \lambda$ kan vara 0)

$$\begin{bmatrix} 4-\lambda & 1 \\ 3 & 2-\lambda \end{bmatrix} \sim \begin{bmatrix} 3 & 2-\lambda \\ 4-\lambda & 1 \end{bmatrix} \begin{matrix} -\frac{4-\lambda}{3} \\ + \end{matrix} \sim \begin{bmatrix} 3 & 2-\lambda \\ 0 & -\frac{(4-\lambda)(2-\lambda)}{3} + 1 \end{bmatrix} \tag{5.4}$$

Vi ser att vi har icke-trivial lösning om och endast om

$$-\tfrac{(4-\lambda)(2-\lambda)}{3} + 1 = 0, \quad \lambda^2 - 6\lambda + 5 = 0, \quad \text{dvs } \lambda = 1 \text{ eller } \lambda = 5 \tag{5.5}$$

För att bestämma egenvektorerna sätter vi in respektive egenvärde λ och löser homogena ekvationen $(A - \lambda I)x = 0$. Lösningarna skrivs på parameterform.
Med $\lambda = 1$:

$$\begin{bmatrix} 3 & 1 \\ 3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & \frac{1}{3} \\ 0 & 0 \end{bmatrix}, \quad \begin{cases} x_1 + \frac{1}{3}x_2 = 0, \\ \qquad 0 = 0, \end{cases} \quad x = t \begin{bmatrix} -\frac{1}{3} \\ 1 \end{bmatrix} = -\tfrac{1}{3}t \begin{bmatrix} 1 \\ -3 \end{bmatrix} \tag{5.6}$$

Med $\lambda = 5$:

$$\begin{bmatrix} -1 & 1 \\ 3 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \quad \begin{cases} x_1 - x_2 = 0, \\ \qquad 0 = 0, \end{cases} \quad x = t \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{5.7}$$

Vi har funnit två egenpar: $\lambda_1 = 1$, $v_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ och $\lambda_2 = 5$, $v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

1. Enhetsmatrisen behövs här: uttrycket $A - \lambda$ har ingen mening; vi kan inte addera matris och tal.

Det är inte de enskilda egenvektorerna som är av intresse utan de underrum som de spänner upp.

Definition 5.2 (Egenrum) Om λ är egenvärde till A så kallas

$$E_\lambda = \mathcal{N}(A - \lambda I) = \{x \in \mathbb{R}^n \mid Ax = \lambda x\} \tag{5.8}$$

egenrummet hörande till λ. Dess dimension, $m = \dim(E_\lambda)$, kallas *geometriska multipliciteten* för λ.

Egenrummet $E_\lambda \subseteq \mathbb{R}^n$ är ett underrum. Det är ju nollrummet till $A - \lambda I$ och spänns upp av alla homogenlösningar till $(A - \lambda I)x = 0$. Obs att nollvektorn $0 \in E_\lambda$ trots att den inte är en egenvektor.

Ruta 5.1 (Egenvärdesproblem) Egenvärdesproblemet kan nu formuleras: finn alla egenvärden till A och bestäm en bas för respektive egenrum.

I exempel 5.2 har vi $E_{\lambda_1} = \operatorname{span}\left\{\begin{bmatrix}1\\-3\end{bmatrix}\right\}$ och $E_{\lambda_2} = \operatorname{span}\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$. Båda egenvärdena har geometrisk multiplicitet 1.

Exempel 5.3 (Enhetsmatrisen) Med $A = I \in \mathbb{R}^{n\times n}$ har vi ett enda egenvärde $\lambda_1 = 1$ och alla $x \neq 0$ är egenvektorer, dvs $Ix = x$. Alltså: $E_{\lambda_1} = \mathbb{R}^n$ och geometriska multipliciteten är n.

Sats 5.1 (Egenvärdena för triangulär matris) Om $A \in \mathbb{R}^{n\times n}$ är en triangulär matris, så är dess egenvärden lika med dess diagonalelement. Alltså: $\lambda = a_{11}, \ldots, a_{nn}$.

Bevis. Om A är triangulär så är också $A - \lambda I$ triangulär. Ekvationen $(A - \lambda I)x = 0$ har icke-triviala lösningar om och endast om det finns någon fri variabel, vilket inträffar då något av diagonalelementen $a_{ii} - \lambda = 0$, dvs $\lambda = a_{ii}$. □

Exempel 5.4 (Triangulär matris) Matrisen $A = \begin{bmatrix}1&2&3\\0&2&4\\0&0&2\end{bmatrix}$ är triangulär så att $\lambda = 1,\ 2$.

$$\lambda_1 = 1: \quad A - \lambda_1 I = \begin{bmatrix}0&2&3\\0&1&4\\0&0&1\end{bmatrix} \sim \begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}, \quad x = s\begin{bmatrix}1\\0\\0\end{bmatrix} \tag{5.9}$$

$$\lambda_2 = 2: \quad A - \lambda_2 I = \begin{bmatrix}-1&2&3\\0&0&4\\0&0&0\end{bmatrix} \sim \begin{bmatrix}1&-2&0\\0&0&1\\0&0&0\end{bmatrix}, \quad x = s\begin{bmatrix}2\\1\\0\end{bmatrix} \tag{5.10}$$

Vi har

$$\lambda_1 = 1, \quad E_{\lambda_1} = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}, \quad \text{geometrisk multiplicitet } m_1 = 1 \tag{5.11}$$

$$\lambda_2 = 2, \quad E_{\lambda_2} = \text{span}\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \right\}, \quad \text{geometrisk multiplicitet } m_2 = 1 \tag{5.12}$$

Vi ser att egenvektorerna tillsammans bildar en linjärt oberoende mängd

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \right\}, \tag{5.13}$$

men den spänner ej upp hela $\mathbb{R}^3$.

För att ytterligare förstå vilka tal som är egenvärden till en matris $A \in \mathbb{R}^{n\times n}$ noterar vi att λ är egenvärde om och endast om matrisen $A - \lambda I$ är singulär, vilket i sin tur är ekvivalent med att

$$\det(A - \lambda I) = 0 \tag{5.14}$$

Denna ekvation kallas den *karakteristiska ekvationen* för A. Funktionen $p(\lambda) = \det(A - \lambda I)$ är ett polynom av grad n. Det kallas det *karakteristiska polynomet* för A. Detta visas genom utveckling av determinanten. Till exempel för $n = 3$ har vi

$$\det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix} = (a_{11} - \lambda) \begin{vmatrix} a_{22} - \lambda & a_{23} \\ a_{32} & a_{33} - \lambda \end{vmatrix} \tag{5.15}$$

$$- a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} - \lambda \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} - \lambda \\ a_{31} & a_{32} \end{vmatrix} \tag{5.16}$$

vilket är ett polynom i λ av grad 3.

Egenvärdena är alltså rötterna till det karakteristiska polynomet. Från *algebrans fundamentalsats* vet vi att vet vi att ett polynom av grad $n \geqslant 1$ har minst ett (komplext) nollställe. Tillsammans med *faktorsatsen* visar detta att vi har precis n (eventuellt komplexa) rötter, dvs

$$\det(A - \lambda I) = c(\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_p)^{n_p} \tag{5.17}$$

med $n_1 + \cdots + n_p = n$ och distinkta egenvärden $\lambda_k \in \mathbb{C}$, $k = 1, \ldots, p$. Exponenten n_k är den *algebraiska multipliciteten* för egenvärdet λ_k. Eftersom karakteristika polynomet har reella koefficienter, så kommer eventuella komplexa egenvärden i konjugerade par, $\lambda = \alpha \pm \mathrm{i}\beta$. I så fall kan motsvarande egenvektorer också vara komplexa, $x \in \mathbb{C}^n$. (Vi hänvisar till avsnitt 2.3 i del I.)

Exempel 5.5 (Karakteristisk ekvation) Vi återvänder till exempel 5.2. Vi har

$$\det(A - \lambda I) = \begin{vmatrix} 4-\lambda & 1 \\ 3 & 2-\lambda \end{vmatrix} = (4-\lambda)(2-\lambda) - 3 \tag{5.18}$$

$$= \lambda^2 - 6\lambda + 5 = (\lambda-3)^2 - 4 = \{\lambda = 3 \pm 2\} \tag{5.19}$$

$$= (\lambda-1)(\lambda-5) \tag{5.20}$$

dvs $\lambda_1 = 1$ med $n_1 = 1$ och $\lambda_2 = 5$ med $n_2 = 1$. Vi har redan sett att de geometriska multipliciteterna är $m_1 = 1$ och $m_2 = 1$.

Exempel 5.6 (Karakteristisk ekvation) Vi återvänder till exempel 5.4. Determinanten är triangulär (sats 3.11) så vi får genast

$$\det(A-\lambda I) = \begin{vmatrix} 1-\lambda & 2 & 3 \\ 0 & 2-\lambda & 4 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda)(2-\lambda) = -(\lambda-1)(\lambda-2)^2 \tag{5.21}$$

Här är $\lambda_1 = 1$ med algebraisk multiplicitet $n_1 = 1$ och $\lambda_2 = 2$ med algebraisk multiplicitet $n_2 = 2$. Vi har redan sett att de geometriska multipliciteterna är $m_1 = 1$ och $m_2 = 1$. Vi noterar att $m_2 < n_2$.

Vi frågar nu hur stort rum egenvektorerna spänner upp. Bildar de kanske en bas för $\mathbb{R}^n$?

Sats 5.2 (Egenvektorer till distinkta egenvärden är linjärt oberoende) Om $v_1, \dots, v_p$ är egenvektorer hörande till *distinkta* egenvärden $\lambda_1, \dots, \lambda_p$, så är mängden $\{v_1, \dots, v_p\}$ linjärt oberoende. Speciellt: om matrisen har n stycken distinkta egenvärden, så bildar egenvektorerna en bas för $\mathbb{R}^n$.

Bevis. Vi har $Av_j = \lambda_j v_j$, $j = 1, \dots, p$, med $\lambda_i \neq \lambda_j$ om $i \neq j$. Antag motsatsen till satsens påstående, dvs att någon av vektorerna är en linjär kombination av de övriga. Låt k vara det första index sådant att $\{v_1, \dots, v_k\}$ är linjärt oberoende, dvs

$$v_{k+1} = c_1 v_1 + \cdots + c_k v_k \tag{5.22}$$

Då har vi

$$\lambda_{k+1} v_{k+1} = c_1 \lambda_{k+1} v_1 + \cdots + c_k \lambda_{k+1} v_k \tag{5.23}$$

och, eftersom $Av_j = \lambda_j v_j$, även

$$\lambda_{k+1} v_{k+1} = A v_{k+1} = A(c_1 v_1 + \cdots + c_k v_k) = c_1 \lambda_1 v_1 + \cdots + c_k \lambda_k v_k \tag{5.24}$$

Genom att kombinera (5.23) och (5.24) får vi

$$c_1(\lambda_1 - \lambda_{k+1}) v_1 + \cdots + c_k(\lambda_k - \lambda_{k+1}) v_k = 0 \tag{5.25}$$

Men $\{v_1, \dots, v_k\}$ är linjärt oberoende, vilket medför att alla koefficienterna $c_j(\lambda_j - \lambda_{k+1}) = 0$. Här är $\lambda_j \neq \lambda_{k+1}$ (distinkta), så att alla $c_j = 0$. Men då leder (5.22) till $v_{k+1} = 0$, vilket är omöjligt då det är en egenvektor. □

För varje distinkt egenvärde λ_k i (5.17) har vi ett egenrum E_{λ_k} med geometrisk multiplicitet $m_k = \dim(E_{\lambda_k})$. Satsen säger att om vi väljer en bas av egenvektorer för varje egenrum, så bildar de tillsammans en linjärt oberoende mängd. De spänner upp ett rum med dimensionen $m_1 + \cdots + m_p \leqslant n$. Om $m_1 + \cdots + m_p = n$, så spänner egenvektorerna upp hela rummet $\mathbb{R}^n$, annars inte. Vi har då en bas för $\mathbb{R}^n$ av bestående av egenvektorer, en *egenbas* för $\mathbb{R}^n$.

Vi ska senare bevisa att $m_k \leqslant n_k$, dvs de geometriska multipliciteterna överstiger inte de algebraiska multipliciteterna (sats 5.5). Det betyder att $m_1 + \cdots + m_p \leqslant n = n_1 + \cdots + n_p$, med likhet om och endast om alla $m_k = n_k$. Om $m_k < n_k$, säger vi att motsvarande egenvärde λ_k är *defekt*; det egenvärdet har inte tillräckligt många linjärt oberoende egenvektorer för att vi ska få en egenbas för $\mathbb{R}^n$.

Exempel 5.7 (Egenbas för $\mathbb{R}^3$) Vi tar matrisen

$$A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} \tag{5.26}$$

och får (med hjälp av symbolisk datorberäkning)

$$\det(A - \lambda I) = \begin{vmatrix} 4-\lambda & -1 & 6 \\ 2 & 1-\lambda & 6 \\ 2 & -1 & 8-\lambda \end{vmatrix} \tag{5.27}$$

$$= 36 - 40\lambda + 13\lambda^2 - \lambda^3 = -(\lambda - 9)(\lambda - 2)^2 \tag{5.28}$$

Vi har två distinkta karakteristiska rötter (egenvärden): $\lambda_1 = 2$ med algebraisk multiplicitet $n_1 = 2$ och $\lambda_2 = 9$ med algebraisk multiplicitet $n_2 = 1$. Vi beräknar egenvektorer genom att lösa $(A - \lambda I)x = 0$.

Med $\lambda_1 = 2$:

$$A - \lambda_1 I = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & -\frac{1}{2} & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad x = s\begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \tag{5.29}$$

Egenrummet är $E_{\lambda_1} = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$ med geometrisk multiplicitet $m_1 = 2$.

Med $\lambda_2 = 9$:

$$A - \lambda_2 I = \begin{bmatrix} -5 & -1 & 6 \\ 2 & -8 & 6 \\ 2 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}, \quad x = s\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{5.30}$$

Egenrummet är $E_{\lambda_2} = \operatorname{span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}\right\}$ med geometrisk multiplicitet $m_2 = 1$.

Här är alltså de geometriska och algebraiska multipliciteterna lika. Enligt sats 5.2 är vektorn i E_{λ_2} ej linjär kombination av vektorerna i E_{λ_1}. Egenvektorerna bildar en bas för $\mathbb{R}^3$:

$$\left\{\begin{bmatrix}\frac{1}{2}\\1\\0\end{bmatrix}, \begin{bmatrix}-3\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right\} \quad \text{med matrisen } P = \begin{bmatrix}\frac{1}{2} & -3 & 1\\1 & 0 & 1\\0 & 1 & 1\end{bmatrix} \tag{5.31}$$

Matrisen är inverterbar ty den har linjärt oberoende kolonner. Vi får (datorberäkning)

$$P^{-1}AP = \begin{bmatrix}2 & 0 & 0\\0 & 2 & 0\\0 & 0 & 9\end{bmatrix} = D \tag{5.32}$$

där matrisen D är en *diagonal matris* med egenvärdena på diagonalen. Vi ska förklara dessa saker nu.

Definition 5.3 (Similära matriser) Två matriser A och B är *similära* om det finns en inverterbar matris P sådan att

$$A = PBP^{-1} \tag{5.33}$$

Obs att $A = PBP^{-1} \iff P^{-1}AP = B$.

Sats 5.3 (Similära matriser har samma egenvärden) Om A och B är similära matriser så har de samma karakteristiska polynom och därmed samma egenvärden.

Bevis. Vi visar att A och B har samma karakteristiska polynom. Om $A = PBP^{-1}$, så

$$A - \lambda I = PBP^{-1} - \lambda PP^{-1} = P(B - \lambda I)P^{-1} \tag{5.34}$$

och med produktregeln för determinant (sats 3.12)

$$\det(A - \lambda I) = \det(P(B - \lambda I)P^{-1}) = \det(P)\det(B - \lambda I)\det(P^{-1}) \tag{5.35}$$

$$= \det(PP^{-1})\det(B - \lambda I) = \det(I)\det(B - \lambda I) = \det(B - \lambda I) \tag{5.36}$$

vilket skulle bevisas. □

Definition 5.4 (Diagonaliserbar matris) Matrisen A är *diagonaliserbar* om den är similär med en *diagonal matris*, dvs

$$A = PDP^{-1} \quad \text{med } D = \operatorname{diag}(d_{11}, \dots, d_{nn}) = \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_{nn} \end{bmatrix} \tag{5.37}$$

Sats 5.4 (Diagonaliserbar matris) $A \in \mathbb{R}^{n\times n}$ är diagonaliserbar om och endast om den har en linjärt oberoende mängd av n egenvektorer $\{v_1, \dots, v_n\}$. I så fall är $A = PDP^{-1}$ där $P = [v_1 \ \cdots \ v_n]$ är *egenvektormatrisen* och $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, *egenvärdesmatrisen*, är diagonal med motsvarande egenvärden på diagonalen. Mängden $\{v_1, \dots, v_n\}$ är då en bas för $\mathbb{R}^n$; den kallas *egenbasen*.

Obs att $A = PDP^{-1} \iff D = P^{-1}AP \iff AP = PD$.

Bevis. Vi bildar $P = [v_1 \ \cdots \ v_n]$ och $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Då har vi

$$AP = A[v_1 \ \cdots \ v_n] = [Av_1 \ \cdots \ Av_n] \tag{5.38}$$
$$PD = [\lambda_1 v_1 \ \cdots \ \lambda_n v_n] \tag{5.39}$$

Det betyder att $AP = PD$ om och endast om $Av_k = \lambda_k v_k$ för $k = 1, \dots, n$.

Om nu $\{v_1, \dots, v_n\}$ är linjärt oberoende med egenvärden $\{\lambda_1, \dots, \lambda_n\}$, så är P är inverterbar och $AP = PD$, så att $A = PDP^{-1}$.

Å andra sidan: om $A = PDP^{-1}$, så får vi $AP = PD$, dvs $\{\lambda_1, \dots, \lambda_n\}$ är egenvärden med egenvektorer $\{v_1, \dots, v_n\}$. De är linjärt oberoende eftersom P är inverterbar. □

Sats 5.5 (Geometriska multipliciteten överstiger ej den algebraiska) Låt λ_k vara ett egenvärde till matrisen $A \in \mathbb{R}^{n\times n}$ med geometrisk multiplicitet m_k och algebraisk multiplicitet n_k. Då gäller

$$1 \leqslant m_k \leqslant n_k \leqslant n \tag{5.40}$$

Bevis. ★ Att $m_k \geqslant 1$ och $n_k \leqslant n$ är klart, eftersom vi har minst en egenvektor och $\sum_{j=1}^{p} n_j = n$. För att visa att $m_k \leqslant n_k$ tar vi en bas $\{v_1, \dots, v_{m_k}\}$ för egenrummet E_{λ_k} och kompletterar med vektorer $v_{m_k+1}, \dots, v_n$, så att $\{v_1, \dots, v_{m_k}, v_{m_k+1}, \dots, v_n\}$ blir en bas för $\mathbb{R}^n$. Matrisen

$$P = \begin{bmatrix} v_1 & \cdots & v_{m_k} & v_{m_k+1} & \cdots & v_n \end{bmatrix} \tag{5.41}$$

är inverterbar eftersom den har linjärt oberoende kolonner (full rang). På liknande sätt som i (5.38) och (5.39) har vi $AP = PD$, men nu med

$$D = \begin{bmatrix} \lambda_k & 0 & \cdots & 0 & d_{1,m_k+1} & \cdots & d_{1n} \\ 0 & \ddots & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & \lambda_k & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & \vdots & & \vdots \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & d_{n,m_k+1} & \cdots & d_{nn} \end{bmatrix} \tag{5.42}$$

där det övre vänstra hörnet är den diagonala matrisen $\text{diag}(\lambda_k, \cdots, \lambda_k) = \lambda_k I_{m_k} \in \mathbb{R}^{m_k \times m_k}$ och de högra kolonnerna är obestämbara.

Eftersom $AP = PD$ och P är inverterbar, så är A och D similära och därmed $\det(A - \lambda I) = \det(D - \lambda I)$. Men $\det(D - \lambda I)$ innehåller faktorn $(\lambda_k - \lambda)^{m_k}$. Detta betyder att den algebraiska multipliciteten är minst m_k. Alltså: $m_k \leqslant n_k$. □

När förutsättningarna i sats 5.4 är uppfyllda så är alltså de geometriska och algebraiska multipliciteterna lika. Egenvärdena upprepas då i $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$ lika många gånger som multipliciteten anger.

Vi har redan sett att egenvektorer till distinkta egenvärden bildar en linjärt oberoende mängd och om de är n stycken så får vi en bas för $\mathbb{R}^n$ (sats 5.2). I så fall är matrisen diagonaliserbar. Vi formulerar detta i nästa sats.

Sats 5.6 (En matris är diagonaliserbar om den har n distinkta egenvärden) Om $A \in \mathbb{R}^{n\times n}$ har n distinkta egenvärden, så är den diagonaliserbar.

Exempel 5.8 (Icke-diagonaliserbar matris) Vi återvänder till exempel 5.4 och exempel 5.6. Här har vi $\lambda_1 = 1$ med geometrisk multiplicitet $m_1 = 1$, algebraisk multiplicitet $n_1 = 1$ och $\lambda_2 = 2$ med $m_2 = 1$ och $n_2 = 2$. Eftersom $m_2 < n_2$, så är matrisen ej diagonaliserbar. Det "fattas" en vektor i E_{λ_2}.

Exempel 5.9 (Icke-diagonaliserbar matris) Låt

$$A = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}, \quad \det(A - \lambda I) \begin{vmatrix} 3-\lambda & 1 & 0 \\ 0 & 3-\lambda & 1 \\ 0 & 0 & 3-\lambda \end{vmatrix} = (3-\lambda)^3 \tag{5.43}$$

Vi har ett enda egenvärde $\lambda = 3$ med algebraisk multiplicitet 3. Med $\lambda = 3$ har vi

$$A - \lambda I = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{5.44}$$

dvs egenrummet är $E_\lambda = \text{span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}\right\}$ med geometrisk multiplicitet 1, vilket är mindre än den algebraiska multipliciteten. Det fattas två vektorer för att vi ska få en bas för $\mathbb{R}^3$. Matrisen är inte diagonaliserbar.

Exempel 5.10 (Matrispotens med diagonaliserbar matris) Att beräkna matrispotensen $A^k = A \cdots A$ (k faktorer) är jobbigt ($(k-1)n^2$ multiplikationer). Men antag att A är diagonaliserbar. Då får vi

$$A^2 = (PDP^{-1})^2 = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PD^2P^{-1} \tag{5.45}$$

där

$$D^2 = \text{diag}(\lambda_1^2, \dots, \lambda_n^2) \tag{5.46}$$

På samma vis

$$A^k = PD^kP^{-1}, \quad D^k = \text{diag}(\lambda_1^k, \dots, \lambda_n^k) \tag{5.47}$$

Enkelt! Men vi måste förstås beräkna P, P^{-1} och D först, vilket inte är så lätt.

Formeln i (5.47) är viktig för teori snarare än för praktisk beräkning. Till exempel om alla $|\lambda_k| < 1$, så ser vi genast att $A^k \to 0$ då $k \to \infty$, därför att alla $\lambda_j^k \to 0$.

Ruta 5.2 (Metod för egenvärdesproblem)

1. Bestäm egenvärdena $\{\lambda_k\}_{k=1}^p$ genom att lösa den karakteristiska ekvationen $\det(A - \lambda I) = 0$.
2. För varje distinkt egenvärde λ_k bestäm en bas för egenrummet E_{λ_k} med geometrisk multiplicitet $m_k = \dim(E_{\lambda_k})$. Om $\sum_{k=1}^p m_k < n$, så är A inte diagonaliserbar och vi stoppar här. Annars bildar egenvektorerna en bas $\{v_1, \dots, v_n\}$ för $\mathbb{R}^n$.
3. Bilda egenvektormatrisen $P = [v_1 \ \cdots \ v_n]$, där v_j är egenvektorerna från steg 2.
4. Bilda egenvärdesmatrisen $D = \text{diag}(\lambda_1, \dots, \lambda_n)$, där egenvärdena upprepas lika många gånger som den geometriska (=algebraiska) multipliciteten anger.
5. Då gäller $A = PDP^{-1}$.

Denna metod är viktig för teorin, för den sammanfattar hur egenvärdesproblem fungerar och detta kan användas för att bevisa saker om matriser. Metoden är opraktisk för beräkning då $n > 3$. Den fungerar inte bra ens för numeriska beräkningar då n är stort. Det som räddar oss är att vi ofta endast behöver beräkna några få egenvärden, till exempel det största eller minsta och eventuellt några fler. I nästa kapitel ska vi göra det numeriskt.

5.2 Spektralsatsen för symmetriska matriser

Vi erinrar oss att $A \in \mathbb{R}^{n \times n}$ är diagonaliserbar om den är similär med en diagonal matris, dvs

$$A = PDP^{-1} \tag{5.48}$$

I så fall är P egenvektormatrisen och D egenvärdesmatrisen. Vilka matriser är diagonaliserbara? Vi har inget enkelt svar på den frågan, men vi ska nu visa att *alla symmetriska matriser är diagonaliserbara*.

Definition 5.5 (Symmetrisk matris) En matris $A \in \mathbb{R}^{n \times n}$ är *symmetrisk* om $A^T = A$, dvs $a_{ji} = a_{ij}$.

Exempel 5.11 (Symmetriskt egenvärdesproblem) Matrisen $A = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$ är symmetrisk. Med symbolisk datorberäkning får vi det karakteristiska polynomet:

$$\det(A - \lambda I) = -\lambda^3 + 17\lambda^2 - 90\lambda + 144 = -(\lambda - 3)(\lambda - 6)(\lambda - 8) \tag{5.49}$$

Egenvärdena är $\lambda_1 = 3,\ \lambda_2 = 6,\ \lambda_3 = 8$ och egenrummen är

$$E_{\lambda_1} = \text{span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}, \quad E_{\lambda_2} = \text{span}\left\{\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}\right\}, \quad E_{\lambda_3} = \left\{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}\right\} \tag{5.50}$$

Med egenvektormatrisen $P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -2 & 0 \end{bmatrix}$ får vi $P^T P = \text{diag}(3, 6, 2)$, dvs vi har en ortogonal egenbas. Efter normering får vi en ortonormerad bas med matrisen

$$P = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 8 \end{bmatrix} \tag{5.51}$$

Eftersom P har ortonormerade kolonner, har vi nu $P^T P = I$. Vi vet att P är inverterbar (linjärt oberoende kolonner). Genom att multiplicera med P^{-1} från höger får vi

$$(P^T P)P^{-1} = IP^{-1}, \quad P^T = P^{-1} \tag{5.52}$$

dvs inversen är lika med P^T. Enkelt! Nu får (5.48) formen

$$A = PDP^T \tag{5.53}$$

dvs vi behöver inte beräkna någon invers.

Definition 5.6 (Ortogonal matris) En kvadratisk matris P kallas *ortogonal matris* om den har ortonormerade kolonner, dvs om

$$P^T P = I \tag{5.54}$$

En sådan matris borde egentligen kallas ortonormerad matris, men beteckningen ortogonal matris är vedertagen.

Ortogonala matriser har några speciella egenskaper som vi lyfter fram i nästa sats.

Sats 5.7 (Egenskaper för ortogonal matris) Antag att $P \in \mathbb{R}^{n\times n}$ är en ortogonal matris, dvs $P^T P = I$. Då gäller följande:

1. $\det(P) = \pm 1$
2. P är inverterbar med $P^{-1} = P^T$
3. $(Pu)\cdot(Pv) = u \cdot v$ för alla $u, v \in \mathbb{R}^n$
4. $\|Pu\| = \|u\|$ för alla $u \in \mathbb{R}^n$

Punkt 2 innebär att även $PP^T = I$, dvs även raderna är ortonormerade. Punkt 3 och 4 innebär att skalärprodukt och norm är invarianta (ändras ej) vid ortogonal transformation $u \mapsto Pu$.

Bevis. 1. Produktregeln och transponeringsregeln för determinant (sats 3.12) ger

$$1 = \det(I) = \det(P^T P) = \det(P^T)\det(P) = (\det(P))^2 \tag{5.55}$$

2. Enligt punkt 1 är $\det(P) \neq 0$, så att P är inverterbar. Genom att multiplicera $P^T P = I$ med P^{-1} från höger får vi $P^T = P^T P P^{-1} = P^{-1}$, dvs inversen är lika med P^T.

3. $(Pu)\cdot(Pv) = (Pu)^T(Pv) = u^T P^T P v = u^T I v = u^T v = u \cdot v$.

4. Följer av 3. $\square$

Vi återvänder nu till det symmetriska egenvärdesproblemet. I allmänhet kan egenvärden och egenvektorer vara komplexa, men inte för en symmetrisk matris som vi visar i nästa sats.

I beviset använder vi *skalärprodukt för komplexa vektorer* $u, v \in \mathbb{C}^n$:

$$u \cdot v = \bar{v}^T u = \sum_{i=1}^{n} \bar{v}_i u_i \in \mathbb{C} \tag{5.56}$$

där $\bar{z}$ är komplexkonjugat, så att $\bar{z}z = (\alpha - \mathrm{i}\beta)(\alpha + \mathrm{i}\beta) = \alpha^2 + \beta^2 = |z|^2$. Då blir

$$\|u\|^2 = u \cdot u = \bar{u}^T u = \sum_{i=1}^{n} \bar{u}_i u_i = \sum_{i=1}^{n} |u_i|^2 \geqslant 0 \tag{5.57}$$

och $u \cdot u = 0$ endast då $u = 0$. Det betyder att produkten i (5.56) är positivt definit. De övriga reglerna i sats 4.14 gäller också. Kommutativa lagen modifieras dock till $v \cdot u = \overline{u \cdot v}$.

Notera att vi konjugerar den andra faktorn, v, i (5.56). Det leder till att $(\alpha u) \cdot v = \alpha(u \cdot v)$ och $u \cdot (\alpha v) = \bar{\alpha}(u \cdot v)$.

Sats 5.8 (Symmetrisk matris har reella egenvärden) Egenvärdena till en reell symmetrisk matris är *reella*.

Bevis. Antag att (λ, v) är ett egenpar, dvs $Av = \lambda v$, där $\lambda \in \mathbb{C}$ och $v \in \mathbb{C}^n$ kan vara komplexa och $v \neq 0$. Vi ska visa att de är reella om A är reell och symmetrisk.

Vi bildar skalärprodukten $v \cdot (Av)$, dvs på matrisform enligt (5.56)

$$\overline{(Av)}^T v = (\bar{A}\bar{v})^T v = \bar{v}^T \bar{A}^T v = \bar{v}^T A v = \bar{v}^T (Av) \tag{5.58}$$

där vi använt att matrisen är reell och symmetrisk, dvs $\bar{A}^T = A$. Här är högerledet

$$\bar{v}^T (Av) = \bar{v}^T (\lambda v) = \lambda \bar{v}^T v = \lambda \|v\|^2 \tag{5.59}$$

medan vänsterledet är

$$\overline{(Av)}^T v = \overline{(\lambda v)}^T v = (\bar{\lambda}\bar{v})^T v = \bar{\lambda}\bar{v}^T v = \bar{\lambda}\|v\|^2 \tag{5.60}$$

Alltså:

$$\bar{\lambda}\|v\|^2 = \lambda\|v\|^2, \quad (\bar{\lambda} - \lambda)\|v\|^2 = 0 \tag{5.61}$$

Vi drar slutsatsen att $\bar{\lambda} = \lambda$, eftersom $\|v\| \neq 0$. Men detta betyder att $\lambda \in \mathbb{R}$, vilket skulle bevisas. Eftersom både A och λ är reella, så måste även v vara reell. $\square$

Vi erinrar oss från sats 5.2 att distinkta egenvärden i allmänhet leder till linjärt oberoende egenvektorer. Om matrisen är symmetrisk så blir det ännu bättre:

Sats 5.9 (Egenvektorer till distinkta egenvärden är ortogonala) Egenvektorer som hör till distinkta egenvärden till en symmetrisk matris är ortogonala.

Bevis. Antag att $Av_1 = \lambda_1 v_1$ och $Av_2 = \lambda_2 v_2$ med $\lambda_1 \neq \lambda_2$. Vi ska visa att $v_1^T v_2 = 0$. Vi räknar och använder $A^T = A$:

$$\lambda_1 v_1^T v_2 = (\lambda_1 v_1)^T v_2 = (Av_1)^T v_2 = v_1^T A^T v_2 = v_1^T A v_2 = v_1^T (\lambda_2 v_2) = \lambda_2 v_1^T v_2 \tag{5.62}$$

Detta leder till

$$(\lambda_1 - \lambda_2) v_1^T v_2 = 0 \tag{5.63}$$

och därmed $v_1^T v_2 = 0$ eftersom $\lambda_1 \neq \lambda_2$. $\square$

Vi tar ett exempel till.

Exempel 5.12 (Symmetriskt egenvärdesproblem) Matrisen $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$ är symmetrisk. Karakteristiska polynomet är

$$\det(A - \lambda I) = -\lambda^3 + 12\lambda^2 - 21\lambda - 98 = -(\lambda - 7)^2(\lambda + 2) \tag{5.64}$$

Egenvärdena är $\lambda_1 = 7$ med algebraisk multiplicitet $n_1 = 2$ och $\lambda_2 = -2$ med algebraisk multiplicitet $n_2 = 1$. Egenrummen är

$$E_{\lambda_1} = \text{span}\{v_1, v_2\}, \qquad v_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix}, \quad m_1 = 2 \tag{5.65}$$

$$E_{\lambda_2} = \text{span}\{v_3\}, \qquad v_3 = \begin{bmatrix} -1 \\ -\frac{1}{2} \\ 1 \end{bmatrix}, \quad m_2 = 1 \tag{5.66}$$

Vi ser att $v_3^T v_1 = 0$ och $v_3^T v_2 = 0$ (v_3 är ortogonal mot v_1 och v_2) i enlighet med sats 5.9, men $v_2^T v_1 = -\frac{1}{2} \neq 0$. Men $\{v_1, v_2\}$ är linjärt oberoende så vi kan ortogonalisera:

$$u_1 = v_1, \quad W_1 = \text{span}\{u_1\}, \quad \|u_1\| = \sqrt{2} \tag{5.67}$$

$$u_2 = v_2 - P_{W_1}(v_2) = v_2 - \frac{v_2^T u_1}{u_1^T u_1} u_1 \tag{5.68}$$

$$= \begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} - \frac{-\frac{1}{2}}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{4} \\ 1 \\ \frac{1}{4} \end{bmatrix}, \quad \|u_2\| = \frac{3\sqrt{2}}{4} \tag{5.69}$$

Då är $\{u_1, u_2\}$ en ortogonal bas för E_{λ_1} och $\{u_1, u_2, v_3\}$ en ortogonal bas för $\mathbb{R}^3$. Om vi slutligen normerar får vi en *ortonormerad egenbas* med *ortogonal egenvektormatris*

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & -\frac{1}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \end{bmatrix} \tag{5.70}$$

och

$$P^T A P = D = \text{diag}(7, 7, -2) = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix} \tag{5.71}$$

I allmänhet kan den geometriska multipliciteten till ett egenvärde vara mindre än den algebraiska multipliciteten, vilket betyder att det fattas egenvektorer för att bilda en bas för hela $\mathbb{R}^n$. Detta händer inte för symmetriska matriser.

Sats 5.10 (Geometrisk multiplicitet är lika med algebraisk för symmetrisk matris) Om λ är ett egenvärde med algebraisk multiplicitet m till en symmetrisk matris, så finns en ortonormerad mängd $\{v_1, \dots, v_m\}$ av m egenvektorer som hör till λ.

Bevis. Vi utelämnar beviset. Det svåra är att visa att den geometriska multipliciteten är lika med den algebraiska. Sedan tar vi en bas för egenrummet och ortonormerar den med Gram–Schmidts metod. □

Vi kan nu sammanfatta dessa resultat.

Sats 5.11 (Spektralsatsen för symmetriska matriser) En symmetrisk matris $A \in \mathbb{R}^{n\times n}$ har *reella egenvärden* $\lambda_1, \dots \lambda_n$, där varje egenvärde förekommer lika många gånger som dess algebraiska multiplicitet anger, och en *ortonormerad bas* $\{v_1, \dots, v_n\}$ av motsvarande egenvektorer.

Alternativt kan vi säga: om $A \in \mathbb{R}^{n\times n}$ är en symmetrisk matris så finns en *ortogonal matris* $P \in \mathbb{R}^{n\times n}$ och en *reell diagonal matris* $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ sådana att $P^T AP = D$.

Bevis. Satsen följer direkt av föregående satser. □

Exempel 5.13 (Symmetrisk differentialoperator) Transponatet av en matris uppfyller relationen

$$(Au)\cdot v = (Au)^T v = u^T(A^T v) = u \cdot (A^T v) \quad \text{för alla } u, v \in \mathbb{R}^n \tag{5.72}$$

Om $L\colon V \to V$ är en linjär operator i ett allmänt skalärproduktrum kan vi imitera (5.72) och definiera transponatet L^T av L genom

$$\langle Lu, v\rangle = \langle u, L^T v\rangle \quad \text{för alla } u, v \in V \tag{5.73}$$

Man kan visa att det alltid finns en operator $L^T\colon V \to V$ som uppfyller (5.73). Vi utelämnar beviset. Istället ska vi visa det för ett specifikt exempel, nämligen för differentialoperatorn $L\colon V \to W$ med

$$L = -\mathrm{D}^2 \tag{5.74}$$
$$V = C_0^2([0,1]) = \{u \in C^2([0,1]) \mid u(0) = u(1) = 0\} \tag{5.75}$$
$$W = C([0,1]) \tag{5.76}$$

se exempel 4.5. Funktioner i rummet $C_0^2([0,1])$ uppfyller *randvillkor* i ändpunkterna av intervallet $[0,1]$. Då är ekvationen $Lu = f$ ett *randvärdesproblem* (se avsnitt 6.5 i del II)

$$\begin{cases} -u''(x) = f(x), & x \in (0,1) \\ u(0) = 0,\ u(1) = 0 \end{cases} \tag{5.77}$$

För att beräkna transponatet av operatorn L tar vi $u, v \in V$, bildar $\langle Lu, v\rangle$, integrerar partiellt två gånger med användning av randvillkoren $u(0) = u(1) = 0, v(0) = v(1) = 0$:

$$\langle Lu, v\rangle = \langle -\mathrm{D}^2 u, v\rangle = \int_0^1 -u''(x)\, v(x)\, \mathrm{d}x \tag{5.78}$$

$$= \Big[-u'(x)v(x)\Big]_0^1 + \int_0^1 u'(x)\, v'(x)\, \mathrm{d}x \tag{5.79}$$

$$= \Big[-u'(x)v(x)\Big]_0^1 + \Big[u(x)v'(x)\Big]_0^1 - \int_0^1 u(x)\, v''(x)\, \mathrm{d}x \tag{5.80}$$

$$= \int_0^1 u(x)\,(-v''(x))\, \mathrm{d}x = \langle u, -\mathrm{D}^2 u\rangle = \langle u, L^T v\rangle \tag{5.81}$$

Detta visar att $L^T = -\mathrm{D}^2 = L$, dvs operatorn är symmetrisk. I allmänna skalärproduktrum kallas oftast den transponerade operatorn den *adjungerade operatorn* till L och tecknas L^*. Om $L^* = L$, så är L *självadjungerad.*

Differentialoperatorn $-\mathrm{D}^2\colon C_0^2([0,1]) \to C_0^2([0,1])$ är alltså självadjungerad. Vi ska studera randvärdesproblem mera utförligt i del IV.

5.3 Kvadratiska former

Funktioner av typen

$$f(x) = Ax_1^2 + Bx_2^2 + Cx_3^2 + Dx_1x_2 + Ex_1x_3 + Fx_2x_3 \tag{5.82}$$

förekommer ofta i teknik och vetenskap. Ofta vill vi minimera eller maximera dem. Notera att f innehåller endast kvadratiska termer. Detta är typiskt för kvantiteter som energi, kostnad och så vidare.

Exempel 5.14 (Kvadratisk form) Vi söker minimum och maximum av den kvadratiska funktionen

$$f(x) = x_1^2 + 3x_1x_2 + 4x_2^2 + x_3^2 \tag{5.83}$$

Vi kompletterar kvadraten:

$$x_1^2 + 3x_1x_2 = (x_1 + \tfrac{3}{2}x_2)^2 - \tfrac{9}{4}x_2^2 \tag{5.84}$$

så att

$$f(x) = (x_1 + \tfrac{3}{2}x_2)^2 + \tfrac{7}{4}x_2^2 + x_3^2 \tag{5.85}$$

Vi ser att $f(x) \geqslant 0$ och $f(0) = 0$ så att $\min(f) = 0$. Det finns inget maximum eftersom alla termer är $\geqslant 0$ så att $f(x) \to \infty$ då $x_k \to \infty$.

Men om vi antar *bivillkoret* $x_1^2 + x_2^2 + x_3^2 = 1$, dvs vi minimerar och maximerar över sfären istället för hela rummet $\mathbb{R}^3$, så får vi att min $f = \frac{5-3\sqrt{2}}{2}$, max $f = \frac{5+3\sqrt{2}}{2}$. Vi ska bevisa detta med hjälp av ett symmetriskt egenvärdesproblem.

Vi skriver funktionen på matrisform:

$$f(x) = x^T A x, \quad A = \begin{bmatrix} 1 & \frac{3}{2} & 0 \\ \frac{3}{2} & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.86}$$

Vi har nämligen

$$x^T A x = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & \frac{3}{2} & 0 \\ \frac{3}{2} & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{5.87}$$

$$= x_1^2 + \tfrac{3}{2} x_1 x_2 + \tfrac{3}{2} x_2 x_1 + 4x_2^2 + x_3^2 = f(x) \tag{5.88}$$

Observera att vi har delat upp termen $3x_1x_2 = \frac{3}{2}x_1x_2 + \frac{3}{2}x_2x_1$ i två lika delar för att få en symmetrisk matris.

I allmänhet har vi

$$f(x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n \tag{5.89}$$

$$+ a_{22}x_2^2 + 2a_{23}x_2x_3 + \cdots + 2a_{2n}x_2x_n \tag{5.90}$$

$$+ a_{33}x_3^2 + 2a_{34}x_3x_4 + \cdots + 2a_{3n}x_3x_n \tag{5.91}$$

$$+ \cdots + a_{nn}x_n^2 \tag{5.92}$$

$$= x^T A x \tag{5.93}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{1n} & \cdots & \cdots & a_{nn} \end{bmatrix} \tag{5.94}$$

Observera att matrisen är symmetrisk. Formeln innehåller alla kvadratiska termer $x_i x_j$ som går att bilda och utnyttjar att $x_i x_j = x_j x_i$.

Definition 5.7 (Kvadratisk form) En *kvadratisk form* är en funktion $Q\colon \mathbb{R}^n \to \mathbb{R}$ som ges av $Q(x) = x^T A x$ med en symmetrisk matris $A \in \mathbb{R}^{n\times n}$.

En kvadratisk form är alltså en reellvärd funktion i flera variabler som endast innehåller kvadratiska termer.

Det underlättar att vi kan byta variabler $x = Py$ så att den kvadratiska formen får en enklare form utan blandade produkter $x_i x_j$ med $i \neq j$.

Sats 5.12 (Satsen om principalaxlar) Låt $A \in \mathbb{R}^{n\times n}$ vara symmetrisk. Då finns en ortogonal matris P sådan att variabeltransformationen $x = Py$ leder till $x^T Ax = y^T Dy$ med en diagonal matris D. Matrisen $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ är egenvärdesmatrisen till A och P är en motsvarande egenvektormatris. Vi har

$$x^T Ax = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \tag{5.95}$$

Bevis. Spektralsatsen, sats 5.11, säger att matrisen A är ortogonalt diagonaliserbar:

$$P^T P = I,\ D = \mathrm{diag}(\lambda_1, \dots, \lambda_n),\ A = PDP^T,\ D = P^T AP \tag{5.96}$$

Med $x = Py$ får vi

$$x^T Ax = (Py)^T A(Py) = y^T (P^T AP)y = y^T Dy = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \tag{5.97}$$

□

Exempel 5.15 (Kvadratisk form) Tillbaka till exempel 5.14. Vi har

$$f(x) = x_1^2 + 3x_1x_2 + 4x_2^2 + x_3^2 = x^T Ax, \quad A = \begin{bmatrix} 1 & \frac{3}{2} & 0 \\ \frac{3}{2} & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.98}$$

Vi löser egenvärdesproblemet för A. Karakteristiska ekvationen är

$$\begin{vmatrix} 1-\lambda & \frac{3}{2} & 0 \\ \frac{3}{2} & 4-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} 1-\lambda & \frac{3}{2} \\ \frac{3}{2} & 4-\lambda \end{vmatrix} \tag{5.99}$$

$$= (1-\lambda)((1-\lambda)(4-\lambda) - \tfrac{9}{4}) \tag{5.100}$$

$$= (1-\lambda)(\lambda^2 - 5\lambda + \tfrac{7}{4}) \tag{5.101}$$

$$= (1-\lambda)((\lambda - \tfrac{5}{2})^2 - \tfrac{25}{4} + \tfrac{7}{4}) \tag{5.102}$$

$$= (1-\lambda)((\lambda - \tfrac{5}{2})^2 - \tfrac{18}{4}) \tag{5.103}$$

$$= -(\lambda - 1)(\lambda - \tfrac{5}{2} + \tfrac{3\sqrt{2}}{2})(\lambda - \tfrac{5}{2} - \tfrac{3\sqrt{2}}{2}) \tag{5.104}$$

Egenvärdena är (i storleksordning)

$$\lambda_1 = \tfrac{5}{2} - \tfrac{3\sqrt{2}}{2} \approx 0.38,\ \lambda_2 = 1,\ \lambda_3 = \tfrac{5}{2} + \tfrac{3\sqrt{2}}{2} \approx 4.62 \tag{5.105}$$

Med den ortogonala egenvektormatrisen P och $x = Py$ får vi

$$f(x) = x^T Ax = y^T Dy = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2 \tag{5.106}$$

$$= (\tfrac{5}{2} - \tfrac{3\sqrt{2}}{2})y_1^2 + y_2^2 + (\tfrac{5}{2} + \tfrac{3\sqrt{2}}{2})y_3^2 \tag{5.107}$$

Vi vill bestämma minimum och maximum av f över sfären $\|x\| = 1$. Eftersom P är ortogonal har vi $\|x\| = \|Py\| = \|y\|$, så att

$$\min_{\|x\|=1} f(x) = \min_{\|x\|=1} x^T Ax = \min_{\|y\|=1} y^T Dy = \tfrac{5}{2} - \tfrac{3\sqrt{2}}{2} \quad \text{fås då } y = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{5.108}$$

På samma vis

$$\max_{\|x\|=1} f(x) = \max_{\|x\|=1} x^T Ax = \max_{\|y\|=1} y^T Dy = \tfrac{5}{2} + \tfrac{3\sqrt{2}}{2} \quad \text{fås då } y = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{5.109}$$

Minimum respektive maximum över sfären ges alltså av minsta respektive största egenvärdet, λ_{min} respektive λ_{max}.

Axlarna y_1, y_2, y_3 kallas *principalaxlar* eller *huvudaxlar* för A. De bildar ett ortogonalt koordinatsystem i $\mathbb{R}^3$.

Exempel 5.16 (Max/min problem) Vi bestämmer minimum och maximum för funktionen

$$f(x) = \frac{5x_1^2 + 5x_2^2 + 8x_3^2 + 8x_1x_2 - 4x_1x_3 + 4x_2x_3}{x_1^2 + x_2^2 + x_3^2} \tag{5.110}$$

Täljaren kan skrivas $x^T Ax$ med $A = \begin{bmatrix} 5 & 4 & -2 \\ 4 & 5 & 2 \\ -2 & 2 & 8 \end{bmatrix}$. Egenvärdena är $\lambda = 0,\ 9,\ 9$, se övning 5.1 (d). Med motsvarande ortogonala egenvektormatris P sätter vi $x = Py$ och får $x^T Ax = y^T Dy$, $x^T x = y^T y$, så att

$$f(x) = \frac{x^T Ax}{x^T x} = \frac{y^T Dy}{y^T y} = \frac{0y_1^2 + 9y_2^2 + 9y_3^2}{y_1^2 + y_2^2 + y_3^2} = \frac{9y_2^2 + 9y_3^2}{y_1^2 + y_2^2 + y_3^2} \tag{5.111}$$

Minimum fås då $y_2 = y_3 = 0$:

$$\min f = \frac{0}{y_1^2} = 0 \tag{5.112}$$

Maximum fås då $y_1 = 0$:

$$\max f = \frac{9y_2^2 + 9y_3^2}{y_2^2 + y_3^2} = 9 \tag{5.113}$$

Definition 5.8 (Definit kvadratisk form) En kvadratisk form $Q(x) = x^T Ax$ är

1. *positivt definit* om $Q(x) > 0$ för alla $x \neq 0$
2. *negativt definit* om $Q(x) < 0$ för alla $x \neq 0$
3. *indefinit* om $Q(x)$ tar både positiva och negativa värden

Om $Q(x) \geqslant 0$ respektive $Q(x) \leqslant 0$ för alla x, så är Q *positivt semidefinit* respektive *negativt semidefinit*. Vi säger även att matrisen A är positivt definit etc om dess kvadratiska form är det.

Obs att $Q(0) = 0$.
I exempel 5.15 är matrisen positivt definit eftersom (5.106) ger

$$x^T Ax = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2 \geqslant \lambda_1 y_1^2 > 0 \quad \text{för alla } y \neq 0 \tag{5.114}$$

ty $\lambda_1 = \frac{5}{2} - \frac{3\sqrt{2}}{2} > 0$. Detta resonemang gäller allmänt enligt nästa sats.

Sats 5.13 (Definit kvadratisk form) Låt $A \in \mathbb{R}^{n\times n}$ vara symmetrisk med egenvärdena $\lambda_1, \ldots, \lambda_n$. Då är $Q(x) = x^T Ax$

1. *positivt definit* om och endast om alla $\lambda_i > 0$
2. *negativt definit* om och endast om alla $\lambda_i < 0$
3. *indefinit* om och endast om A har både positiva och negativa egenvärden (något $\lambda_i > 0$ och något $\lambda_i < 0$)

Obs att strängt tecken krävs. Om till exempel alla $\lambda_i \geqslant 0$ så är A endast positivt semidefinit, $x^T Ax \geqslant 0$.

Bevis. Enligt sats 5.12 har vi

$$x^T Ax = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \tag{5.115}$$

Eftersom alla $y_i^2 > 0$ då $x = Py \neq 0$, kontrolleras tecknet på $x^T Ax$ av tecknen på egenvärdena så som satsen säger. $\square$

Exempel 5.17 (Positivt definit differentialoperator) Från exempel 5.13 vet vi att differentialoperatorn

$$L = -\mathrm{D}^2 \tag{5.116}$$

$$V = C_0^2([0,1]) = \{u \in C^2([0,1]) \mid u(0) = u(1) = 0\} \tag{5.117}$$

$$W = C([0,1]) \tag{5.118}$$

är symmetrisk (självadjungerad), $L^T = -\mathrm{D}^2 = L$, dvs $\langle Lu, v\rangle = \langle u, Lv\rangle$. I analogi med $x^T Ax$ för symmetrisk matris kan vi då definiera en kvadratisk form $\langle Lu, u\rangle = \langle u, Lu\rangle$. Vi visar att den är positivt definit. Vi integrerar partiellt och använder randvillkoren:

$$\langle Lu, u\rangle = \int_0^1 -u''(x)\, u(x)\, \mathrm{d}x \tag{5.119}$$

$$= \Big[-u'(x)\, u(x)\Big]_0^1 + \int_0^1 u'(x)\, u'(x)\, \mathrm{d}x = \int_0^1 u'(x)^2\, \mathrm{d}x \geqslant 0 \tag{5.120}$$

dvs $\langle Lu, u\rangle$ är positivt semidefinit. Men om $\langle Lu, u\rangle = \int_0^1 u'(x)^2\, \mathrm{d}x = 0$, så måste $u'(x) = 0$ för alla $x \in [0, 1]$, ty u' är kontinuerlig. Men då är $u(x) = c$ för någon konstant c. Randvillkoret $u(0) = 0$ ger $c = 0$. Alltså: $\langle Lu, u\rangle = 0$ medför att $u = 0$, dvs den är positivt definit.

Nu ser vi att minustecknet är naturligt i definitionen $L = -\mathrm{D}^2$. Med plustecken, dvs D^2, får vi ju en negativt definit operator.

Övningar

5.1 Egenvärden och egenvektorer

Övning 5.1 Lös egenvärdesproblemet för matrisen A. Ange egenvärdena och deras respektive egenrum samt algebraisk och geometrisk multiplicitet.

(a) $A = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$ (b) $A = \begin{bmatrix} 3 & -5 \\ 2 & -3 \end{bmatrix}$

(c) $A = \begin{bmatrix} 1 & 0 & 0 \\ 3 & -1 & 0 \\ 5 & 4 & 0 \end{bmatrix}$ (d) $A = \begin{bmatrix} 5 & 4 & -2 \\ 4 & 5 & 2 \\ -2 & 2 & 8 \end{bmatrix}$

Övning 5.2 Avgör om matrisen är diagonaliserbar och bestäm i så fall matriserna P och D.

(a) $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$ (c) $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ (d) $A = \begin{bmatrix} 1 & 0 & 0 \\ 3 & -1 & 0 \\ 5 & 4 & 0 \end{bmatrix}$

(e) $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Övning 5.3 Beräkna A^{10} då A är matrisen i övning 5.2 (b).

5.2 Spektralsatsen för symmetriska matriser

Övning 5.4 Bestäm en ortogonal matris P sådan att $P^T AP$ blir en diagonal matris.

(a) $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ (b) $A = \begin{bmatrix} 6 & 2 \\ 2 & 9 \end{bmatrix}$ (c) $A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$ (d) $A = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$

Övning 5.5 Bestäm en ortogonal matris P sådan att $P^T AP$ blir en diagonal matris.

(a) $A = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 5 & 4 \\ -4 & 4 & 3 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & -1 & 2 \\ -2 & 2 & 0 \end{bmatrix}$ (c) $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 3 \end{bmatrix}$

Övning 5.6 Beräkna A^k, k positivt heltal.

(a) A från övning 5.4 (a) (b) A från övning 5.4 (b)

(c) A från övning 5.4 (c) (d) A från övning 5.5 (b)

5.3 Kvadratiska former

Övning 5.7 Bestäm karaktären (positivt definit etc) hos den kvadratiska formen.
(a) $5x_1^2 + 3x_2^2 + x_3^2 - 4x_1x_2 - 2x_2x_3$
(b) $x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_1x_3 + x_2x_3$
(c) $x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3$
(d) $x_1^2 + 5x_2^2 + 3x_3^2 - 2x_1x_2 + 4x_1x_3 - 8x_2x_3$

Övning 5.8 Bestäm värdemängden för funktionen

$$f(x) = \frac{7x_1^2 + 6x_2^2 + 5x_3^2 - 4x_1x_2 - 4x_2x_3}{x_1^2 + x_2^2 + x_3^2}.$$

Övning 5.9 Bestäm minsta värdet av funktionen $f(x) = x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3$ på sfären $x_1^2 + x_2^2 + x_3^2 = 8$ och ange för vilka x det inträffar.

Problem

5.1 Egenvärden och egenvektorer

Problem 5.1 Antag att $A, B \in \mathbb{R}^{n\times n}$ med B inverterbar. Visa att AB och BA har samma egenvärden.

Problem 5.2 Visa att A är inverterbar om och endast om $\lambda = 0$ inte är ett egenvärde till A.

Problem 5.3 Visa att om v är egenvektor till A för egenvärdet λ, så är v egenvektor till A^k för egenvärdet λ^k.

Problem 5.4 En kvadratisk matris A kallas *idempotent* om $A^2 = A$. Visa att en idempotent matris bara kan ha egenvärdena 0 eller 1.

Datorövning 5.1 Använd funktionen `eig` i Matlab för att lösa egenvärdesproblemen i föregående övningar.

6. Numerisk lösning av linjära ekvationer

Vi presenterar några numeriska algoritmer för de ekvationer vi har studerat: $Ax = b$ och $Ax = \lambda x$.

Låt $A \in \mathbb{R}^{n\times n}$ vara inverterbar och $b \in \mathbb{R}^n$. Vi vet att den unika lösningen är $x = A^{-1}b$ och den kan beräknas med hjälp av transformation till reducerad trappstegsform, $[A\ b] \sim [U\ d]$. Vi är nu ute efter mer effektiva metoder för beräkning av lösningen, speciellt för mycket stora ekvationssystem, dvs då n är mycket stort. Vi vill alltså undvika att bilda A^{-1} eller att göra en fullständig radreduktion.

6.1 LU-faktorisering

Låt $A \in \mathbb{R}^{n\times n}$ vara inverterbar. Vi vill faktorisera matrisen, dvs skriva den som en produkt

$$A = LU \tag{6.1}$$

av två matriser $L, U \in \mathbb{R}^{n\times n}$. I denna sektion ska vi försöka göra det så att L är nedåt triangulär och U är uppåt triangulär[1]. I så fall kan $Ax = b$ skrivas $LUx = b$ och lösas effektivt i två steg. Sätt $y = Ux$ så fås

$$Ly = b \tag{6.2}$$
$$Ux = y \tag{6.3}$$

Vi löser först $Ly = b$. Eftersom L är nedåt triangulär kan det göras snabbt genom framåtsubstitution. Vi har nu y och löser $Ux = y$ snabbt med bakåtsubstitution. Vi visar detta i ett exempel.

1. Från engelskan: L = lower triangular, U = upper triangular.

Exempel 6.1 (LU-faktorisering) Vi transformerar matrisen till trappstegsform:

$$A = \begin{bmatrix} 2 & 4 & -1 \\ -4 & -5 & 3 \\ 2 & 10 & 4 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 \\ 0 & 3 & 1 \\ 0 & 6 & 5 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix} = U \tag{6.4}$$

Vi har alltså

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & -1 \\ -4 & -5 & 3 \\ 2 & 10 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix} \tag{6.5}$$

eller

$$E_3 E_2 E_1 A = U \tag{6.6}$$

Vi fortsätter inte till reducerad trappstegsform. Vi har endast använt radoperationer av typ 1, dvs inga permutationer och skalningar. Radoperationerna är inverterbara och vi får

$$A = (E_3 E_2 E_1)^{-1} U = E_1^{-1} E_2^{-1} E_3^{-1} U \tag{6.7}$$

där

$$E_1^{-1} E_2^{-1} E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} = L \tag{6.8}$$

Obs hur E_j^{-1} fås genom byta tecken på det "aktiva" matriselementet i E_j och hur detta placeras in på samma plats i L. Det beror på att E_1^{-1}, E_2^{-1} bara ändrar kolonn 1 i E_3^{-1}, där det är nollor.

Vi har nu

$$A = LU \quad \text{med} \quad L = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 2 & 4 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix} \tag{6.9}$$

Obs att L är *nedåt triangulär med 1 på diagonalen* och att U är *uppåt triangulär*. Med $b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ blir ekvationen $Ax = b$ först $Ly = b$, dvs

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{6.10}$$

med lösningen (framåtsubstitution)

$$y_1 = 1 \tag{6.11}$$

$$y_2 = 1 + 2y_1 = 3 \tag{6.12}$$

$$y_3 = 1 - y_1 - 2y_2 = -6 \tag{6.13}$$

Sedan $Ux = y$, dvs

$$\begin{bmatrix} 2 & 4 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ -6 \end{bmatrix} \tag{6.14}$$

med lösningen (bakåtsubstitution)

$$x_3 = -6/3 = -2 \tag{6.15}$$

$$x_2 = (3 - x_3)/3 = 5/3 \tag{6.16}$$

$$x_1 = (1 - 4x_2 + x_3)/2 = -23/6 \tag{6.17}$$

Allt detta ska naturligtvis göras med datorberäkning.

Definition 6.1 (LU-faktorisering) Låt $A \in \mathbb{R}^{n\times n}$. En LU-faktorisering av A ges av

$$A = LU \tag{6.18}$$

där $L \in \mathbb{R}^{n\times n}$ är *nedåt triangulär med 1 på diagonalen* och $U \in \mathbb{R}^{n\times n}$ är *uppåt triangulär*.

Följande metod visar att LU-faktorisering finns om det går att transformera matrisen till trappstegsform utan permutationer och skalningar.

Ruta 6.1 (LU-faktorisering utan pivotering)

1. Transformera till trappstegsmatris genom elementära radoperationer utan permutation (om det går):
$$U = E_p \cdots E_1 A \tag{6.19}$$
2. Bilda
$$L = E_1^{-1} \cdots E_p^{-1} \tag{6.20}$$
Detta går att göra genom att byta tecken på de "aktiva" matriselementen i E_j och placera in dem i L.

Ibland måste vi göra permutationer för att kunna fortsätta radoperationerna. Då måste vi hålla reda på permutationerna och L blir inte triangulär utan en produkt av triangulära matriser och permutationsmatriser. Detta kallas LU-faktorisering med *pivotering*. Då samlar vi information om vilka radbyten (permutationer) som gjorts i en *permutationsmatris* P. Denna har precis en 1 i varje rad och varje kolonn: $p_{ij} = 1$ betyder att rad j flyttats och blivit rad i. Då har vi att PA, den permuterade matrisen, kan LU-faktoriseras utan pivotering, dvs $PA = LU$. Vi löser sedan det permuterade systemet $PAx = Pb$ med $Ly = Pb$, $Ux = y$.

Exempel 6.2 (LU-faktorisering med pivotering) Vi återvänder till exempel 6.1. Vi använder MATLAB. Kommandoraden `[L U]=lu(A)` ger

$$L = \begin{bmatrix} -0.5 & 0.2 & 1.0 \\ 1.0 & 0 & 0 \\ -0.5 & 1.0 & 0 \end{bmatrix}, \quad U = \begin{bmatrix} -4.0 & -5.0 & 3.0 \\ 0 & 7.5 & 5.5 \\ 0 & 0 & -0.6 \end{bmatrix} \tag{6.21}$$

Matrisen L är inte triangulär. Tydligen har MATLAB funnit det lämpligt att pivotera trots att vi vet att det inte behövs. Kommandoraden `[L U P]=lu(A)` ger

$$L = \begin{bmatrix} 1.0 & 0 & 0 \\ -0.5 & 1.0 & 0 \\ -0.5 & 0.2 & 1.0 \end{bmatrix}, U = \begin{bmatrix} -4.0 & -5.0 & 3.0 \\ 0 & 7.5 & 5.5 \\ 0 & 0 & -0.6 \end{bmatrix}, P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \tag{6.22}$$

Permutationsmatrisen visar följande radbyten: $2 \to 1, 1 \to 3, 3 \to 2$. Den permuterade matrisen PA är

$$PA = \begin{bmatrix} -4 & -5 & 3 \\ 2 & 10 & 4 \\ 2 & 4 & -1 \end{bmatrix} \tag{6.23}$$

Kommadoraden `[L U]=lu(P*A)` ger samma L, U som i (6.22).

LU-faktorisering är effektiv speciellt då vi vill lösa $Ax = b$ med många högerled b men samma A. I stället för att bilda A^{-1} och $x = A^{-1}b$, vilket är ineffektivt, så bildar vi L och U och löser i två steg som beskrivits ovan.

6.2 Iterativa lösningsmetoder

Iterativa lösningsmetoder bygger på att göra en matris-vektor multiplikation i varje steg. Detta är speciellt effektivt om matrisen är *stor och gles*, dvs då de flesta elementen är noll. Då kan vi spara matrisen i glest format, dvs spara endast de nollskiljda elementen, och implementera snabb matris-vektor multiplikation.

Låt $A \in \mathbb{R}^{n\times n}$ vara inverterbar och $b \in \mathbb{R}^n$. Vi vill lösa ekvationen $Ax = b$ med en fixpunktsiteration, se avsnitt 6.3 i del I. Ett sätt att skriva ekvationen på fixpunktsform är att dela upp matrisen enligt $A = L + K$, där matrisen L bör vara lätt att invertera, dvs det ska vara lätt att lösa $Lx = y$. Då får vi

$$Lx = b - Kx \tag{6.24}$$

Om vi multiplicerar med L^{-1} får vi en fixpunktsekvation:

$$x = L^{-1}b - L^{-1}Kx \tag{6.25}$$

dvs

$$x = g(x), \quad g(x) = L^{-1}b - L^{-1}Kx \tag{6.26}$$

och fixpunktsiterationen blir

$$x^{k+1} = L^{-1}b - L^{-1}Kx^k \tag{6.27}$$

Iterationen konvergerar om $g\colon \mathbb{R}^n \to \mathbb{R}^n$ är en kontraktion (Banachs fixpunktssats):

$$\|g(x) - g(y)\| = \|(L^{-1}b - L^{-1}Kx) - (L^{-1}b - L^{-1}Ky)\| \tag{6.28}$$
$$= \|L^{-1}K(x-y)\| \leqslant \gamma\|x-y\| \quad \text{för alla } x, y \in \mathbb{R}^n \tag{6.29}$$

med $\gamma < 1$. Detta är ekvivalent med

$$\|L^{-1}Kx\| \leqslant \gamma\|x\| \quad \text{för alla } x \in \mathbb{R}^n \text{ med } \gamma < 1 \tag{6.30}$$

Vi vill inte bilda inversa matrisen L^{-1} utan skriver iterationen i (6.27) på formen

$$Lx^{k+1} = b - Kx^k \tag{6.31}$$

där vi alltså ska lösa en ekvation av typen $Lx = y$ i varje steg.

Jacobis metod. En naturlig uppdelning $A = L + K$ är att ta $L = \text{diag}(A)$, dvs en diagonal matris med samma diagonal som A. Matrisen K blir då matrisen av de icke-diagonala elementen i A, dvs

$$L = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix}, \quad K = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ a_{n1} & \cdots & a_{n,n-1} & 0 \end{bmatrix} \tag{6.32}$$

Utskriven elementvis blir då iterationen (6.31)

$$a_{ii}x_i^{k+1} = b_i - \sum_{j\neq i} a_{ij}x_j^k, \quad i = 1, \ldots, n \tag{6.33}$$

vilket löses lätt

$$x_i^{k+1} = \frac{1}{a_{ii}}\Big(b_i - \sum_{j\neq i} a_{ij}x_j^k\Big), \quad i = 1, \ldots, n \tag{6.34}$$

Detta kallas Jacobis metod. Vi ska inte genomföra beviset att metoden konvergerar via konvergensvillkoret (6.30). Men vi kan nog ana att det betyder att matrisen L ska vara "stor" i förhållande till matrisen K, så att $L^{-1}K$ blir "liten". I det här fallet betyder det att diagonalelementen ska vara stora i förhållande till icke-diagonalelementen. Det precisa villkoret är följande.

Definition 6.2 (Strängt diagonaldominant matris) En matris $A \in \mathbb{R}^{n\times n}$ är *strängt diagonaldominant* om

$$|a_{ii}| > \sum_{j\neq i} |a_{ij}| \quad \text{för alla } i = 1, \ldots, n \tag{6.35}$$

I algoritmisk form blir Jacobis metod: beräkna

$$y_i \leftarrow \frac{1}{a_{ii}}\Big(b_i - \sum_{j\neq i} a_{ij}x_j\Big), \quad i = 1, \ldots, n \tag{6.36}$$

och uppdatera $x \leftarrow y$. Se algoritm 1.

Algoritm 1 Jacobis metod

Input: $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, $x_0 \in \mathbb{R}^n$, tolerans $\mathrm{TOL} \in \mathbb{R}$
Output: $\hat{x} \approx x = A^{-1}b$
Antaganden: A strängt diagonaldominant, $\mathrm{TOL} > 0$

$x \leftarrow x_0$
$\delta \leftarrow 2 \cdot \mathrm{TOL}$
while $\delta > \mathrm{TOL}$ **do**
    $\delta \leftarrow 0$
    for $i = 1, 2, \ldots, n$ **do**
        $s \leftarrow b_i$
        for $j = 1, 2, \ldots, n$ **do**
            if $i \neq j$ **then**
                $s \leftarrow s - a_{ij}x_j$
            end if
        end for
        $y_i \leftarrow s/a_{ii}$
        $\delta \leftarrow \max(\delta, |y_i - x_i|)$
    end for
    $x \leftarrow y$
end while
$\hat{x} \leftarrow x$

Gauss–Seidels metod. Om vi i stället för (6.36) gör

$$x_i \leftarrow \frac{1}{a_{ii}}\Big(b_i - \sum_{j\neq i} a_{ij}x_j\Big), \quad i = 1, \ldots, n \tag{6.37}$$

se algoritm 2, så slipper vi använda en temporär variabel y. Dessutom använder vi de mest aktuella värdena av x_j. För att se det tydligare skriver vi ut (6.37) med iterationsindex:

$$x_i^{k+1} = \frac{1}{a_{ii}}\Big(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{k+1} - \sum_{j=i+1}^{n} a_{ij}x_j^{k}\Big), \quad i = 1, \ldots, n \tag{6.38}$$

I den första summan använder vi elementen x_j^{k+1} som redan har blivit beräknade. Om vi förlänger med a_{ii} och flyttar första summan till vänsterledet får vi

$$\sum_{j=1}^{i} a_{ij}x_j^{k+1} = b_i - \sum_{j=i+1}^{n} a_{ij}x_j^{k}, \quad i = 1, \ldots, n \tag{6.39}$$

där vi lagt in termen $a_{ii}x_i^{k+1}$ i den första summan. Ekvationen i (6.39) har formen

$$Lx^{k+1} = b - Kx^k \tag{6.40}$$

dvs vi har uppdelningen $A = L + K$ med

$$L = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{12} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{n1} & \cdots & a_{n,n-1} & a_{nn} \end{bmatrix}, \quad K = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & \cdots & 0 & 0 \end{bmatrix} \tag{6.41}$$

Algoritm 2 Gauss–Seidels metod

Input: $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, $x_0 \in \mathbb{R}^n$, tolerans TOL $\in \mathbb{R}$
Output: $\hat{x} \approx x = A^{-1}b$
Antaganden: A strängt diagonaldominant, TOL > 0

$x \leftarrow x_0$
$\delta \leftarrow 2 \cdot \mathrm{TOL}$
while $\delta > \mathrm{TOL}$ **do**
    $\delta \leftarrow 0$
    for $i = 1, 2, \ldots, n$ **do**
        $s \leftarrow b_i$
        for $j = 1, 2, \ldots, n$ **do**
            if $i \neq j$ **then**
                $s \leftarrow s - a_{ij}x_j$
            end if
        end for
        $y \leftarrow s/a_{ii}$
        $\delta \leftarrow \max(\delta, |y - x_i|)$
        $x_i \leftarrow y$
    end for
end while
$\hat{x} \leftarrow x$

Algoritm 3, successiv överrelaxation (SOR), är en modifikation av Gauss–Seidels metod med en “relaxationsparameter” ω som vi kan experimentera med för att förbättra konvergensen.

Konjugerade gradient-metoden. För symmetrisk och positivt definit matris A finns en annan typ av iteration, den konjugerade gradient-metoden. Den minimerar funktionen $\|Ax - b\|^2$ (minimum då $x = A^{-1}b$). Den bygger på ortogonalitet med avseende på skalärprodukten $\langle x, y\rangle = x^T Ax$. (Denna uppfyller villkoren för skalärprodukt i definition 4.15 om A är symmetrisk och positivt definit.) Genom att iteraten blir ortogonala, når vi exakta lösningen inom n steg. Men då n är mycket stort når vi acceptabel noggrannhet mycket tidigare än så och iterationen stoppar. Vi går inte in på hur den fungerar utan presenterar den som en algoritm,

Algoritm 3 Successiv överrelaxation (SOR)

Input: $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$, $x_0 \in \mathbb{R}^n$, tolerans TOL > 0, relaxationsparameter $\omega > 0$
Output: $\hat{x} \approx x = A^{-1}b$
Antaganden: A strängt diagonaldominant, TOL > 0

$x \leftarrow x_0$
$\delta \leftarrow 2 \cdot \mathrm{TOL}$
while $\delta > \mathrm{TOL}$ **do**
    $\delta \leftarrow 0$
    for $i = 1, 2, \ldots, n$ **do**
        $s \leftarrow b_i$
        for $j = 1, 2, \ldots, n$ **do**
            if $i \neq j$ **then**
                $s \leftarrow s - a_{ij} x_j$
            end if
        end for
        $y \leftarrow s / a_{ii}$
        $y \leftarrow \omega y + (1 - \omega) x_i$
        $\delta \leftarrow \max(\delta, |y - x_i|)$
        $x_i \leftarrow y$
    end for
end while
$\hat{x} \leftarrow x$

algoritm 4. Den är mycket snabbare än Jacobi och Gauss–Seidel om matrisen är symmetrisk och positivt definit.

Algoritm 4 Konjugerade gradient-metoden (CG)

Input: $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, $x_0 \in \mathbb{R}^n$, tolerans TOL $\in \mathbb{R}$
Output: $\hat{x} \approx x = A^{-1}b$
Antaganden: A symmetrisk och positivt definit, TOL > 0

```
r ← b − Ax_0
p ← r
while ‖r‖ > TOL do
    y ← Ap
    α ← ‖r‖^2/(y · p)
    x ← x + αp
    s ← r − αy
    β ← ‖s‖^2/‖r‖^2
    p ← s + βp
    r ← s
end while
x̂ ← x
```

6.3 Numerisk lösning av egenvärdesproblem

Metoden i ruta 5.2 är opraktisk, rentav omöjlig, att använda för lösning av egenvärdesproblem då n är större än 3 eller 4 för att inte tala om riktigt stora n. Vi måste ju lösa en polynomekvation av grad n (svårt) och sedan många linjära ekvationssystem (också svårt).

I stället söker vi en approximativ lösning via en iterativ metod. Vi beräknar en följd av vektorer $(x^1, x^2, \dots) = (x^k)_{k=1}^{\infty}$ som konvergerar mot en egenvektor v med $Av = \lambda v$. Det vill säga att $x^k \to v$ då $k \to \infty$ och vi får sedan en följd $(\lambda^k)_{k=1}^{\infty}$ som approximerar egenvärdet genom *Rayleigh-kvoten*

$$\lambda^k = \frac{(x^k)^T A x^k}{(x^k)^T x^k} \to \frac{v^T A v}{v^T v} = \frac{v^T(\lambda v)}{v^T v} = \lambda \frac{v^T v}{v^T v} = \lambda \quad \text{då } k \to \infty \tag{6.42}$$

Rayleigh-kvoten av $A \in \mathbb{R}^{n\times n}$ och $y \in \mathbb{R}^n$ är alltså

$$\frac{y^T A y}{y^T y} \tag{6.43}$$

Obs att $y^T y = \|y\|^2$.

Vi antar inte att A är symmetrisk (som i avsnitt 5.3), endast att A är diagonaliserbar:

$$P^{-1}AP = D, \quad P = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}, \quad D = \text{diag}(\lambda_1, \dots, \lambda_n) \tag{6.44}$$

Den enklaste algoritmen är *potensmetoden*. Vi tar en godtycklig startvektor x^0 och beräknar

$$x^k = Ax^{k-1}, \quad \text{dvs } x^k = A^k x^0 \tag{6.45}$$

En kommentar om beteckningen: x^k med övre index betecknar iterationsvektor nummer k och betyder inte "x upphöjt till k", vilket saknar mening. Egenvektor nummer j betecknas v_j med nedre index. På komponentform blir dessa

$$x^k = \begin{bmatrix} x_1^k \\ \vdots \\ x_n^k \end{bmatrix}, \quad v_j = \begin{bmatrix} v_{1j} \\ \vdots \\ v_{nj} \end{bmatrix} \tag{6.46}$$

Eftersom A är diagonaliserbar har vi en bas av egenvektorer $\{v_j\}_{j=1}^n$. I egenvektorbasen får vi

$$x^0 = c_1 v_1 + \cdots + c_n v_n \quad \left(x^0 = Pc \right) \tag{6.47}$$

och (se exempel 5.10)

$$x^k = \lambda_1^k c_1 v_1 + \cdots + \lambda_n^k c_n v_n \quad \left(x^k = PD^k P^{-1} x^0 = PD^k c \right) \tag{6.48}$$

Följden x^k konvergerar normalt inte: faktorerna λ_j^k kan gå mot oändligheten eller mot noll beroende på om $|\lambda_j|$ är större eller mindre än 1. Men om λ_1 är ett *dominant egenvärde*, dvs

$$|\lambda_1| > |\lambda_2| \geqslant \cdots \geqslant |\lambda_n| \tag{6.49}$$

så kan vi skala med $\frac{1}{\lambda_1^k}$ och få

$$\frac{1}{\lambda_1^k} x^k = c_1 v_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^k c_2 v_2 + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)^k c_n v_n \to c_1 v_1 \in E_{\lambda_1} = \operatorname{span}\{v_1\} \tag{6.50}$$

då $k \to \infty$. (Här antar vi att $c_1 \neq 0$, annars tar vi ny startvektor x^0.)

Alltså: $y^k = \frac{1}{\lambda_1^k} x^k$ konvergerar mot $c_1 v_1$, en egenvektor till λ_1. Den konvergenta iterationen är

$$y^1 = \frac{1}{\lambda_1} A x^0, \quad y^k = \frac{1}{\lambda_1} A y^{k-1} \tag{6.51}$$

dvs multiplicera med A och skala med faktorn $\frac{1}{\lambda_1}$. Men λ_1 är okänd så vi måste skala på något annat sätt. Till exempel: normera

$$z^k = A y^{k-1} \tag{6.52}$$

$$y^k = \pm \frac{1}{\|z^k\|} z^k \tag{6.53}$$

där tecknet väljs så att y^k och y^{k-1} pekar åt (ungefär) samma håll (bildar spetsig vinkel). (Om λ_1 är negativ så växlar y^k riktning; då får vi ingen konvergens. Vi skulle också kunna skala så att största komponenten i y^k blir 1 för att undvika att följden växer obegränsat eller går mot noll och att y^k växlar riktning.) Rayleigh-kvoten (6.42) ger sedan en approximerande följd $\lambda_1^k \to \lambda_1$.

Villkoret (6.49), som det är formulerat, utesluter att λ_1 är ett multipelt egenvärde. Men det är inte hela sanningen: λ_1 ska vara dominant som "distinkt egenvärde". Om till exempel λ_1

är ett dominant dubbelt egenvärde, dvs med algebraisk och geometrisk multiplicitet 2, så är $\lambda_1 = \lambda_2$ ett distinkt egenvärde och (6.49) ersätts av

$$|\lambda_1| = |\lambda_2| > |\lambda_3| \geqslant \cdots \geqslant |\lambda_n| \tag{6.54}$$

Då har vi $E_{\lambda_1} = \text{span}\{v_1, v_2\}$ och (6.48) blir

$$x^k = \lambda_1^k(c_1v_1 + c_2v_2) + \lambda_3^k c_3v_3 + \cdots + \lambda_n^k c_nv_n \tag{6.55}$$

och vi får $y^k = \frac{1}{\lambda_1^k} x^k \to c_1v_1 + c_2v_2 \in E_{\lambda_1}$, där c_1, c_2 ej båda får vara 0. Rayleigh-kvoten ger sedan en approximation till $\lambda_1 = \lambda_2$, men vi får experimentera med valet av x^0 för att hitta två linjärt oberoende egenvektorer.

Hur beräknar vi ett egenvärde som inte är dominant? Svaret är att *skifta och invertera*. Det betyder

$$B = A - \alpha I \qquad \text{(skifta med } \alpha \in \mathbb{R}) \tag{6.56}$$

$$B^{-1} = (A - \alpha I)^{-1} \qquad \text{(invertera)} \tag{6.57}$$

Matriserna B och B^{-1} har samma egenvektorer som A, ty

$$Av = \lambda v \tag{6.58}$$

$$Bv = Av - \alpha v = (\lambda - \alpha)v \tag{6.59}$$

$$v = (\lambda - \alpha)B^{-1}v \tag{6.60}$$

$$B^{-1}v = (\lambda - \alpha)^{-1}v \tag{6.61}$$

Alltså: $(A - \alpha I)^{-1}$ har egenvärdet $(\lambda - \alpha)^{-1}$ med samma egenvektor v. Om nu α är nära λ så blir $(\lambda - \alpha)^{-1}$ ett mycket dominant egenvärde till $(A - \alpha I)^{-1}$.

Detta leder till den *inversa potensmetoden*:

1. Gissa en approximation $\alpha \approx \lambda$.
2. Tillämpa potensmetoden på $(A - \alpha I)^{-1}$ och beräkna en approximativ egenvektor $\hat{v}$.
3. Approximativt egenpar $(\hat{\lambda},\ \hat{v})$ till A fås som $\hat{\lambda} = (\hat{v}^T A\hat{v})/(\hat{v}^T\hat{v})$ och $\hat{v}$.

När vi tillämpar potensmetoden multiplicerar vi med $(A - \alpha I)^{-1}$ i varje steg, dvs (6.52) blir

$$z^k = (A - \alpha I)^{-1}y^{k-1} \tag{6.62}$$

Här bildar vi inte den inversa matrisen utan vi löser ekvationen

$$(A - \alpha I)z^k = y^{k-1} \tag{6.63}$$

Detta är ett bra exempel på när LU-faktorisering är effektiv: vi löser en ekvation många gånger med samma matris men med olika högerled. Vi faktoriserar $A - \alpha I = LU$ en gång för alla och löser i varje steg först $Lv^k = y^{k-1}$, sedan $Uz^k = v^k$.

En nackdel med denna metod är att den beräknar endast ett egenpar åt gången. Idén kan förfinas på många sätt. MATLAB har till exempel funktionen `eig` som beräknar alla egenvärden och en tillhörande egenvektormatris på en gång: `[P D]=eig(A)`.

Exempel 6.3 (Inversa potensmetoden) Vi experimenterar med inversa potensmetoden. Detta lämpar sig inte för handräkning, så vi använder MATLAB. Vi börjar med att skapa en diagonaliserbar matris med kända egenvärden och egenvektorer, dvs vi hittar på P och D och bildar $A = PDP^{-1}$.

```
P = [1 1 1; 1 2 -3; 4 5  6];
D = [5 0 0; 0 -5 0; 0 0 -5];
A = P*D/P
[V, E] = eig(A)
```

Egenvärdena är $\lambda_1 = \lambda_2 = -5,\ \lambda_3 = 5$. Egenvärdet -5 är dubbelt och vi har inget dominant egenvärde eftersom $|\lambda_1| = |\lambda_2| = |\lambda_3|$. Vi förbereder iterationen genom att skapa en slumpmässig startvektor, välja skiftparametern α nära -5 och LU-faktorisera den skiftade matrisen.

```
x = rand(3, 1);
alpha = -4.5;
I = eye(size(A));
[L, U] = lu(A-alpha*I);
```

Sedan itererar vi genom att upprepa följande rader till konvergens.

```
y = L\x;
x = U\y;
x = x/norm(x)
```

Vi avbryter när vi ser att ett lagom antal decimaler har fixerats. Observera att x växlar riktning eftersom vi beräknar ett negativt egenvärde. Därför konvergerar inte följden i strikt mening. Men Rayleigh-kvoten ger ändå rätt egenvärde:

```
lambda = x'*A*x
```

(Här är nämnaren $x^T x = \|x\|^2 = 1$ eftersom vi har normerat.) Om vi vill avbryta iterationen med ett stoppvillkor måste vi spara förra vektorn och se till att nästa vektor inte växlar riktning:

```
y = L\x;
z = U\y;
z = z/norm(z);
z = sign(x'*z)*z;
e = norm(z-x)
x = z
```

Här ser vi till att vinkeln mellan x och z är spetsig genom att byta riktning om vinkeln blir trubbig, $x^T z < 0$. Vi stoppar då e blir tillräckligt liten.

Algoritm 5 Potensmetoden

Input: $A \in \mathbb{R}^{n \times n}$, $x_0 \in \mathbb{R}^n$, tolerans $\text{TOL} \in \mathbb{R}$
Output: Approximativt egenvärde $\hat{\lambda} \approx \lambda$ och en motsvarande approximativ egenvektor $\hat{x}$
Antaganden: A diagonaliserbar, $\lambda \in \mathbb{R}$ dominant egenvärde, $x_0 \neq 0$, $\text{TOL} > 0$

$x \leftarrow x_0$
$\delta \leftarrow 2 \cdot \text{TOL}$
while $\delta > \text{TOL}$ **do**
    $y \leftarrow Ax$
    $y \leftarrow y/\|y\|$
    $y \leftarrow \text{sign}(x^T y) y$
    $\delta \leftarrow \|x - y\|$
    $x \leftarrow y$
end while
$\hat{\lambda} \leftarrow x^T A x$
$\hat{x} \leftarrow x$

Algoritm 6 Inversa potensmetoden

Input: $A \in \mathbb{R}^{n \times n}$, $\alpha \in \mathbb{R}$, $x_0 \in \mathbb{R}^n$, tolerans $\text{TOL} \in \mathbb{R}$
Output: Approximativt egenvärde $\hat{\lambda} \approx \lambda$ och en motsvarande approximativ egenvektor $\hat{x}$
Antaganden: A diagonaliserbar, $\lambda \in \mathbb{R}$ egenvärde nära α, $x_0 \neq 0$, $\text{TOL} > 0$

$LU = A - \alpha I$
$x \leftarrow x_0$
$\delta \leftarrow 2 \cdot \text{TOL}$
while $\delta > \text{TOL}$ **do**
    $y \leftarrow L^{-1} x$
    $y \leftarrow U^{-1} y$
    $y \leftarrow y/\|y\|$
    $y \leftarrow \text{sign}(x^T y) y$
    $\delta \leftarrow \|x - y\|$
    $x \leftarrow y$
end while
$\hat{\lambda} \leftarrow x^T A x$
$\hat{x} \leftarrow x$

Övningar

6.1 LU-faktorisering

Övning 6.1 Bestäm en LU-faktorisering till A och använd denna för att lösa $Ax = b$ (se övning 2.1).

(a) $A = \begin{bmatrix} 3 & 2 \\ 6 & 5 \end{bmatrix}, b = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 3 \\ 3 & 1 & 2 \end{bmatrix}, b = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$

(c) $A = \begin{bmatrix} 2 & 3 & -4 \\ -6 & -5 & 0 \\ 1 & 2 & 4 \end{bmatrix}, b = \begin{bmatrix} -2 \\ 4 \\ 1 \end{bmatrix}$ (d) $A = \begin{bmatrix} 6 & 3 & 12 \\ -2 & 1 & -20 \\ 3 & 2 & -10 \end{bmatrix}, b = \begin{bmatrix} -3 \\ 3 \\ 2 \end{bmatrix}$

6.1 LU-faktorisering

Problem 6.1 Antag att $A = LU$ är en LU-faktorisering. Visa att $\det(A) = u_{11} \cdots u_{nn}$.

6.2 Iterativa lösningsmetoder

Problem 6.2 Visa att (6.25) kan skrivas $x = b + L^{-1}(b - Ax)$. Detta är ett alternativt sätt att skriva $Ax = b$ som fixpunktsekvation: välj en matris L som är lätt att invertera och som approximerar A i någon mening. Ställ sedan upp ekvationen $x = b + L^{-1}(b - Ax)$. Vad blir ekvationen om $L = A$?

Problem 6.3 Visa att $\langle x, y \rangle = x^T Ay$ uppfyller villkoren för skalärprodukt i definition 4.15 om A är symmetrisk och positivt definit.

Datorövningar

Datorövning 6.1 Gör övning 6.1 med hjälp av MATLAB.

Datorövning 6.2 Skriv ett datorprogram som implementerar Jacobis metod.

Datorövning 6.3 Skriv ett datorprogram som implementerar Gauss–Seidels metod.

Datorövning 6.4 Experimentera med inversa potensmetoden som i exempel 6.3.

Datorövning 6.5 Skriv ett datorprogram som implementerar inversa potensmetoden.

7. Tillämpningar

Vi presenterar några tillämpningar av linjär algebra.

7.1 Kurvanpassning

En matematisk modell av verkligheten består ofta av ett funktionellt samband mellan fysiska variabler

$$y = f(x) \tag{7.1}$$

Vi antar i denna diskussion att vi har endast en oberoende variabel x och en beroende variabel y. Funktionen f beskrivs med hjälp av antal parametrar, som vi bestämmer genom att göra observationer i ett kontrollerat experiment. Sambandet (7.1) kan sedan användas för att göra förutsägelser om hur systemet beter sig i icke observerade situationer.

Den enklaste modellen har formen

$$y = \beta_0 + \beta_1 x \tag{7.2}$$

dvs ett linjärt samband[1] mellan x och y med två parametrar β_0 och β_1. Parametrarna ska bestämmas från experimentella mätdata (x_i, y_i), $i = 1, \ldots, N$. Det betyder att vi vill anpassa en rät linje på formen (7.2) till datapunkterna (x_i, y_i). Därför talar vi om *kurvanpassning* och mer specifikt om *linjär regression* i detta fall. Om vi sätter in mätvärdena i ekvationen får vi

1. Sambandet är egentligen *affint*, dvs konstant plus linjär funktion.

ett linjärt ekvationssystem:

$$\begin{cases} \beta_0 + \beta_1 x_1 = y_1 \\ \beta_0 + \beta_1 x_2 = y_2 \\ \vdots \\ \beta_0 + \beta_1 x_N = y_N \end{cases} \tag{7.3}$$

Detta är överbestämt och saknar normalt lösning. Vi bestämmer en approximativ lösning med minsta kvadratmetoden. Vi skriver på matrisform

$$\begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \tag{7.4}$$

dvs

$$X\beta = y \tag{7.5}$$

där

$$X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \tag{7.6}$$

kallas *designmatris*, *parametervektor* och *observationsvektor*. Värdena i X är de värden som vi väljer att använda när vi designar vårt experiment och värdena i y är de värden som vi sedan observerar (mäter).

Minsta kvadratlösningen fås genom att lösa normalekvationerna:

$$(X^T X)\beta = X^T y \tag{7.7}$$

Metoden är inte begränsad till linjära modeller utan kan även användas för mer allmänna modeller

$$y = \beta_0 f_0(x) + \cdots + \beta_p f_p(x) \tag{7.8}$$

Det viktiga är att modellen beror linjärt på parametrarna β_j. Till exempel kan vi ha en kvadratisk modell:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 \tag{7.9}$$

Då får vi

$$\begin{bmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_N & x_N^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \tag{7.10}$$

Här anpassar vi en kvadratisk kurva till mätpunkterna. Detta kallas *kvadratisk regression*.

Exempel 7.1 (Kvadratisk regression) Höjden z för en boll som kastas vertikalt ges av begynnelsevärdesproblemet

$$m\ddot{z}(t) = -mg,\ t > 0; \quad z(0) = z_0,\ \dot{z}(0) = v_0 \tag{7.11}$$

med lösningen

$$z(t) = z_0 + v_0 t - \tfrac{1}{2} g t^2 \tag{7.12}$$

Detta är en kvadratisk modell med parametrarna $\beta_0 = z_0$, $\beta_1 = v_0$, $\beta_2 = -\frac{1}{2}g$. Vi bestämmer parametrarna från observationer z_i vid tidpunkterna t_i, $i = 1, \dots, N$.

Vi har följande data:

i	1	2	3	4	5
t_i [s]	1.0	2.0	3.0	4.0	5.0
z_i [m]	25.0	40.0	45.0	41.0	27.0

Designmatrisen och observationsvektorn blir

$$X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{bmatrix}, \quad z = \begin{bmatrix} 25 \\ 40 \\ 45 \\ 41 \\ 27 \end{bmatrix} \tag{7.13}$$

Med MATLAB beräknar vi:

```
t = [1 2 3 4 5]';
z = [25 40 45 41 27]';
X = [ones(5,1) t t.^2];
rref([X'*X X'*z])
beta = X\z
```

Vi får $\beta = [0.6000\ 29.2143\ -4.7857]^T$, vilket svarar mot $z_0 \approx 0.6000\,\mathrm{m}$, $v_0 \approx 29.2143\,\mathrm{m\,s^{-1}}$ och $g \approx 9.5714\,\mathrm{m\,s^{-2}}$. (Exemplet skapades genom att beräkna $z(t_i)$ enligt (7.12) med $z_0 = 0$, $v_0 = 30$ och $g = 9.81$ och sedan lägga till "mätfel" i z_i genom att stryka decimalerna.)

Vi illustrerar kurvanpassningen genom att plotta kurvan tillsammans med datapunkterna:

```
f=@(t) beta(1)+beta(2).*t+beta(3).*t.^2;
clf, hold on
fplot(f,[0,6])
plot(t,z,'*')
```

7.2 System av linjära ordinära differentialekvationer

Vi ska studera begynnelsvärdesproblem för ett homogent linjärt system av ordinära differentialekvationer:

$$x'(t) = Ax(t),\ t > 0; \quad x(0) = b \tag{7.14}$$

(Systemet är homogent eftersom det är på formen $x'(t) - Ax(t) = 0$. Motsvarande inhomogena system är $x'(t) - Ax(t) = f(t)$.) Här är $x(t) \in \mathbb{R}^n$, $b \in \mathbb{R}^n$ och vi antar att koefficientmatrisen $A \in \mathbb{R}^{n\times n}$ är *diagonaliserbar*, dvs $A = PDP^{-1}$.

Vi använder egenvektorbasen $P = [v_1 \ \cdots \ v_n]$:

$$x = Py, \quad y = P^{-1}x, \quad b = Pc, \quad c = P^{-1}b \tag{7.15}$$

Då får vi

$$y' = \big(P^{-1}x\big)' = P^{-1}x' = P^{-1}Ax = P^{-1}APy = Dy \tag{7.16}$$

och $y(0) = P^{-1}x(0) = P^{-1}b = c$. Vi har ett diagonalt system

$$y'(t) = Dy(t),\ t > 0; \quad y(0) = c \tag{7.17}$$

Eftersom $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ blir detta

$$\begin{cases} y_1'(t) = \lambda_1 y_1(t),\ t > 0; \quad y_1(0) = c_1 \\ \qquad \vdots \\ y_n'(t) = \lambda_n y_n(t),\ t > 0; \quad y_n(0) = c_n \end{cases} \tag{7.18}$$

Vi erinrar oss från del II att det skalära begynnelsevärdesproblemet

$$y_k'(t) = \lambda_k y_k(t),\ t > 0; \quad y_k(0) = c_k \tag{7.19}$$

har unik lösning

$$y_k(t) = c_k \mathrm{e}^{\lambda_k t} \tag{7.20}$$

Vi transformerar tillbaka

$$x(t) = Py(t) = \sum_{k=1}^{n} c_k \mathrm{e}^{\lambda_k t}\, v_k = c_1 \mathrm{e}^{\lambda_1 t}\, v_1 + \cdots + c_n \mathrm{e}^{\lambda_n t}\, v_n \tag{7.21}$$

I egenvektorbasen får vi alltså ett okopplat system av differentialekvationer, där ekvationerna kan lösas individuellt. I det ursprungliga systemet är ekvationerna kopplade till varandra och måste lösas samtidigt. Resultatet är att vi får en lösningsformel

$$x(t) = \sum_{k=1}^{n} c_k \mathrm{e}^{\lambda_k t}\, v_k \tag{7.22}$$

Lösningen är en summa av *egenmoder*

$$c_k \mathrm{e}^{\lambda_k t}\, v_k = c_k \mathrm{e}^{\lambda_k t} \begin{bmatrix} v_{1k} \\ \vdots \\ v_{nk} \end{bmatrix} \tag{7.23}$$

Detta ska inte ses som en beräkningsmetod. Om vi vill beräkna lösningen använder vi hellre numeriska metoder från del II. Det viktiga med lösningsformeln är att den ger kvalitativ information om lösningens beteende. Detta bestäms i sin tur av exponentialfunktionens beteende.

1. Reellt egenvärde λ_k:

$$\mathrm{e}^{\lambda_k t} \begin{cases} \text{växande om } \lambda_k > 0 \\ \text{konstant om } \lambda_k = 0 \\ \text{avtagande om } \lambda_k < 0 \end{cases} \tag{7.24}$$

2. Komplext egenvärde $\lambda_k = \alpha_k \pm \mathrm{i}\beta_k$:

$$\mathrm{e}^{\lambda_k t} = \mathrm{e}^{(\alpha_k \pm \mathrm{i}\beta_k)t} = \mathrm{e}^{\alpha_k t}\mathrm{e}^{\pm \mathrm{i}\beta_k t} \tag{7.25}$$

$$= \mathrm{e}^{\alpha_k t}\Big(\cos(\beta_k t) \pm \mathrm{i}\sin(\beta_k t)\Big) \begin{cases} \text{växande svängning om } \alpha_k > 0 \\ \text{konstant amplitud om } \alpha_k = 0 \\ \text{avtagande svängning om } \alpha_k < 0 \end{cases} \tag{7.26}$$

Vi säger att systemet är *stabilt* om inga egenmoder växer då $t \to \infty$. Detta inträffar om och endast om alla $\mathrm{Re}(\lambda_k) = \alpha_k \leqslant 0$. Annars är systemet *instabilt*.[2]

Exempel 7.2 (Linjärisering) Linjära system av ordinära differentialekvationer $x'(t) = Ax(t)$ uppkommer ofta som ekvationen för störningar av ett ickelinjärt system $u'(t) = f(u(t))$ nära en jämviktspunkt $\bar{u}$, dvs en punkt sådan att $f(\bar{u}) = 0$. Med en störning $x(t) = u(t) - \bar{u}$ menar vi en avvikelse från jämviktsläget. Jämviktsläget $\bar{u}$ är stabilt om alla störningar som är små vid $t = 0$ förblir små då $t \to \infty$, så att $u(t) = \bar{u} + x(t) \approx \bar{u}$, dvs lösningen förblir nära jämviktsläget.

I del IV ska vi lära oss att derivera flervariabelfunktioner och se att derivatan av f i $\bar{u}$ är en matris $A = f'(\bar{u})$. Analogt med avsnitt 4.6 i del I kan vi då approximera $f(u(t))$ med linjäriseringen av f runt $\bar{u}$:

$$f(u(t)) = f(\bar{u} + x(t)) \approx f(\bar{u}) + f'(\bar{u})x(t) = Ax(t) \quad \text{om } x(t) \text{ är liten} \tag{7.27}$$

eftersom $f(\bar{u}) = 0$ och $f'(\bar{u}) = A$. För störningen $x(t) = u(t) - \bar{u}$ får vi

$$x'(t) = (u(t) - \bar{u})' = u'(t) = f(u(t)) \approx Ax(t) \tag{7.28}$$

dvs (approximativt) $x'(t) = Ax(t)$ om $x(t)$ är liten.

2. Detta förutsätter att A är diagonaliserbar, annars blir det mer komplicerat.

Den linjära ekvationen $x'(t) = Ax(t)$ fås alltså genom att linjärisera den ickelinjära ekvationen $u'(t) = f(u(t))$ kring en jämviktspunkt $\bar{u}$. Stabiliteten hos jämviktsläget $\bar{u}$ avgörs då av stabiliteten hos det linjäriserade systemet $x'(t) = Ax(t)$ enligt ovan.

Beteendena i punkt 1 och 2 kan kombineras på många sätt i summan (7.22). Vi demonstrerar några fall som kan inträffa då $n = 2$:

$$x(t) = y_1(t)\, v_1 + y_2(t)\, v_2 = c_1 \mathrm{e}^{\lambda_1 t}\, v_1 + c_2 \mathrm{e}^{\lambda_2 t}\, v_2 \tag{7.29}$$

Exempel 7.3 (Reella egenvärden)

1. Båda $\lambda_1 > 0$ och $\lambda_2 > 0$. Till exempel: $\lambda_1 = 1$, $\lambda_2 = 2$.

$$x(t) = c_1 \mathrm{e}^{t}\, v_1 + c_2 \mathrm{e}^{2t}\, v_2 \tag{7.30}$$

I y_1, y_2-planet har vi då

$$y_2(t) = c_2 \mathrm{e}^{2t} = \frac{c_2}{c_1^2}\left(c_1 \mathrm{e}^{t}\right)^2 = \frac{c_2}{c_1^2}(y_1(t))^2 \quad (c_1 \neq 0) \tag{7.31}$$

dvs kurvor av typen $y_2 = c y_1^2$ ($c \in \mathbb{R}$). Alla egenmoder går mot oändligheten (utom de som startar i origo). **Instabilt system.**

2. Båda $\lambda_1 < 0$ och $\lambda_2 < 0$. Till exempel: $\lambda_1 = -1$, $\lambda_2 = -2$.

$$x(t) = c_1 \mathrm{e}^{-t}\, v_1 + c_2 \mathrm{e}^{-2t}\, v_2 \tag{7.32}$$

Igen får vi kurvor av typen $y_2 = c y_1^2$ ($c \in \mathbb{R}$) men nu genomlöps de i motsatt riktning: in mot origo. **Stabilt system.**

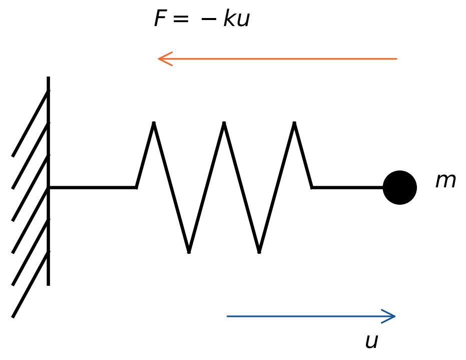
3. Ett $\lambda_1 > 0$ och ett $\lambda_2 < 0$. Till exempel: $\lambda_1 = 1$, $\lambda_2 = -2$.

$$x(t) = c_1 \mathrm{e}^{t}\, v_1 + c_2 \mathrm{e}^{-2t}\, v_2 \tag{7.33}$$

Nu får vi

$$y_2(t) = c_2 \mathrm{e}^{-2t} = c_2 c_1^2 \left(\frac{1}{c_1 \mathrm{e}^{t}}\right)^2 = c_2 c_1^2 \frac{1}{(y_1(t))^2} \quad (c_1 \neq 0) \tag{7.34}$$

dvs kurvor av typen $y_2 = c/y_1^2$ ($c \in \mathbb{R}$). (De kallas *hyperbler* och systemet kallas *hyperboliskt*.) Alla egenmoder går mot oändligheten utom de som startar på y_2-axeln. **Instabilt system.**



Figur 7.1: Massa och fjäder.

Exempel 7.4 (Imaginära egenvärden) Vi studerar begynnelsevärdesproblemet

$$u''(t) + \omega^2 u(t) = 0,\ t > 0; \quad u(0) = u_0,\ u'(0) = u_1 \quad (\omega > 0) \tag{7.35}$$

Denna ekvation beskriver diverse svängningsfenomen, till exempel rörelsen av en kropp med massan m [kg] som är fäst vid en fjäder med fjäderkonstanten k [N m^{-1}]. Fjäderns förlängning (från jämviktsläget) vid tiden t [s] är $u(t)$ [m] och fjäderkraften är då $F = -ku$ [N] (se figur 7.1). Newtons andra lag ger $mu''(t) = -ku(t)$ och vi får (7.35) med $\omega = \sqrt{k/m}$ [s^{-1}][a] och med begynnelseläge u_0 [m] och begynnelsehastighet u_1 [m s^{-1}].

Differentialekvationen är av andra ordningen. Vi lärde oss i del II att vi kan skriva den som ett system av ekvationer av första ordningen genom att sätta $x_1 = u$, $x_2 = u'$. Vi får då

$$\begin{cases} x_1'(t) = x_2(t) \\ x_2'(t) = -\omega^2 x_1(t) \end{cases} \tag{7.36}$$

dvs på matrisform

$$x'(t) = Ax(t),\ t > 0; \quad x(0) = x_0 \tag{7.37}$$

med

$$A = \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix}, \quad x_0 = \begin{bmatrix} u_0 \\ u_1 \end{bmatrix} \tag{7.38}$$

Vi löser egenvärdesproblemet. Karakteristiska polynomet är

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ -\omega^2 & -\lambda \end{vmatrix} = \lambda^2 + \omega^2 = (\lambda - \mathrm{i}\omega)(\lambda + \mathrm{i}\omega) \tag{7.39}$$

med två imaginära rötter $\lambda = \pm \mathrm{i}\omega$. Vi löser egenvektorekvationen $(A - \lambda I)v = 0$ med $\lambda = \mathrm{i}\omega$:

$$\begin{bmatrix} -\mathrm{i}\omega & 1 \\ -\omega^2 & -\mathrm{i}\omega \end{bmatrix} \begin{matrix} \mathrm{i}\omega \\ + \end{matrix} \sim \begin{bmatrix} -\mathrm{i}\omega & 1 \\ 0 & 0 \end{bmatrix} \begin{matrix} \mid \mathrm{i}/\omega \\ \end{matrix} \sim \begin{bmatrix} 1 & \mathrm{i}/\omega \\ 0 & 0 \end{bmatrix} \tag{7.40}$$

Vi har en fri variabel och vi får

$$v = s \begin{bmatrix} -\mathrm{i}/\omega \\ 1 \end{bmatrix} = \frac{-\mathrm{i}}{\omega} s \begin{bmatrix} 1 \\ \mathrm{i}\omega \end{bmatrix}, \quad \text{välj } v_1 = \begin{bmatrix} 1 \\ \mathrm{i}\omega \end{bmatrix} \tag{7.41}$$

Med $\lambda = -\mathrm{i}\omega$ får vi på samma sätt: $v_2 = \begin{bmatrix} 1 \\ -\mathrm{i}\omega \end{bmatrix}$. Egenvektormatrisen och egenvärdesmatrisen är

$$P = \begin{bmatrix} 1 & 1 \\ \mathrm{i}\omega & -\mathrm{i}\omega \end{bmatrix}, \quad P^{-1} = \frac{1}{-2\mathrm{i}\omega} \begin{bmatrix} -\mathrm{i}\omega & -1 \\ -\mathrm{i}\omega & 1 \end{bmatrix}, \quad D = \begin{bmatrix} \mathrm{i}\omega & 0 \\ 0 & -\mathrm{i}\omega \end{bmatrix} \tag{7.42}$$

Lösningsformeln blir

$$x(t) = c_1 \mathrm{e}^{\lambda_1 t} v_1 + c_2 \mathrm{e}^{\lambda_2 t} v_2 \tag{7.43}$$

$$= c_1 \mathrm{e}^{\mathrm{i}\omega t} \begin{bmatrix} 1 \\ \mathrm{i}\omega \end{bmatrix} + c_2 \mathrm{e}^{-\mathrm{i}\omega t} \begin{bmatrix} 1 \\ -\mathrm{i}\omega \end{bmatrix} \tag{7.44}$$

$$= c_1 \big(\cos(\omega t) + \mathrm{i} \sin(\omega t)\big) \begin{bmatrix} 1 \\ \mathrm{i}\omega \end{bmatrix} + c_2 \big(\cos(\omega t) - \mathrm{i} \sin(\omega t)\big) \begin{bmatrix} 1 \\ -\mathrm{i}\omega \end{bmatrix} \tag{7.45}$$

Här är båda $\mathrm{Re}(\lambda_k) = 0$: **stabilt system.** Vi tittar på komponenterna separat och samlar cos och sin:

$$x_1(t) = (c_1 + c_2) \cos(\omega t) + (c_1 - c_2) \mathrm{i} \sin(\omega t) \tag{7.46}$$

$$= a \cos(\omega t) + b \sin(\omega t) \tag{7.47}$$

$$x_2(t) = -(c_1 + c_2)\omega \sin(\omega t) + (c_1 - c_2)\mathrm{i}\omega \cos(\omega t) \tag{7.48}$$

$$= -a\omega \sin(\omega t) + b\omega \cos(\omega t) \tag{7.49}$$

där $a = c_1 + c_2$, $b = \mathrm{i}(c_1 - c_2)$ ska vara reella tal. Obs att $x_2(t) = x_1'(t)$. Detta är en svängning (oscillation). Tag till exempel $a = 1$, $b = 0$, så fås

$$x(t) = \begin{bmatrix} \cos(\omega t) \\ -\omega \sin(\omega t) \end{bmatrix} \tag{7.50}$$

se *lösningskurvor* i figur 7.2. Lösningen kan också illustreras genom ett *fasporträtt* genom att eliminera t med hjälp av trigonometriska ettan:

$$x_1^2 + \frac{1}{\omega^2} x_2^2 = 1 \tag{7.51}$$

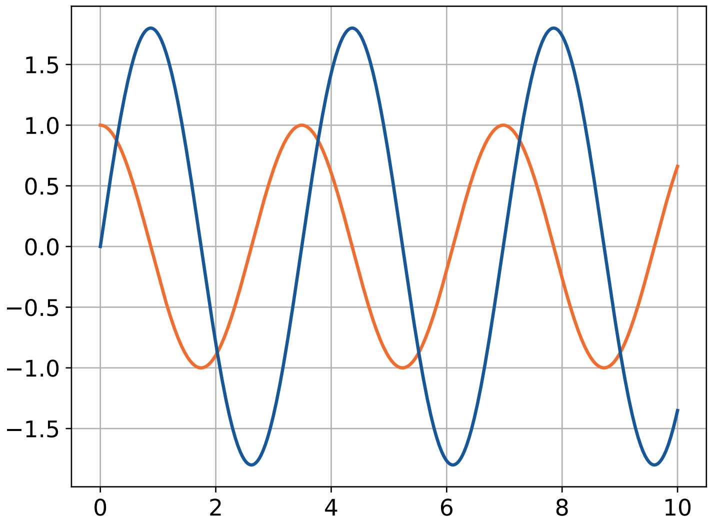
Punkten $x(t)$ startar i punkten $(1, 0)$ och rör sig medurs på en ellips i x_1, x_2-planet (se figur 7.3).

Lösningen till begynnelsevärdesproblemet (7.35) blir nu (med $a = u_0$, $b = u_1/\omega$)

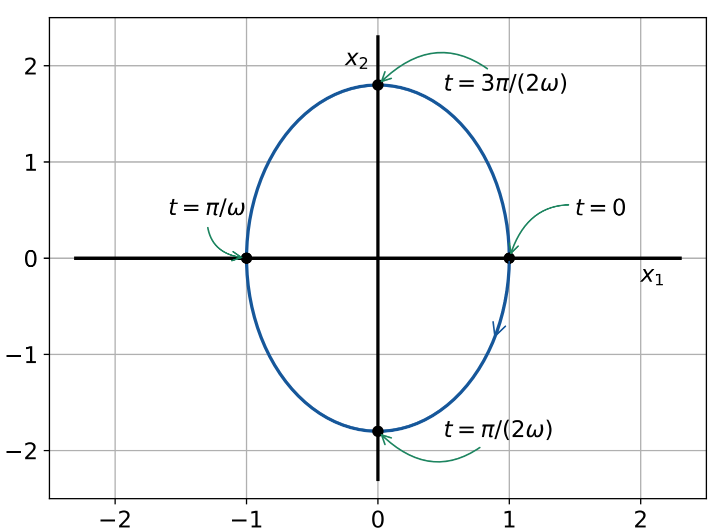
$$u(t) = u_0 \cos(\omega t) + \frac{u_1}{\omega} \sin(\omega t) \tag{7.52}$$

---

a. N=kg m/s^2 ger att $\omega^2 = k/m$ [s^{-2}].



Figur 7.2: Lösningskurvor enligt (7.50). Röd: x_1, blå: $x_2, \omega = 1.8$.



Figur 7.3: Fasporträtt enligt (7.51) med $\omega = 1.8$.

Övningar

Övning 7.1 Bestäm lösningsformeln för $x'(t) = Ax(t)$. (Se övning 5.1.)

(a) $A = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$ (b) $A = \begin{bmatrix} 3 & -5 \\ 2 & -3 \end{bmatrix}$

(c) $A = \begin{bmatrix} 1 & 0 & 0 \\ 3 & -1 & 0 \\ 5 & 4 & 0 \end{bmatrix}$ (d) $A = \begin{bmatrix} 5 & 4 & -2 \\ 4 & 5 & 2 \\ -2 & 2 & 8 \end{bmatrix}$

Problem

Problem 7.1 [Kalibrering av materialmodell]
I hållfasthetsläran kallas sambandet mellan spänning och töjning materialmodell eller konstitutiv modell. Modellen väljs så att den stämmer överens med experimentella data. Metaller kryper vid hög temperatur. Med detta menas att om vi håller spänningen konstant så ökar deformationen (töjningen) med tiden. Detta är en viskoelastisk effekt. Experiment har visat att spännings-töjningsförhållandet är ickelinjärt. En ofta använd modell för krypning hos metaller vid förhöjd temperatur är *Nortons modell*:

$$\dot{\varepsilon} = \frac{1}{\tau}\left(\frac{\sigma}{\sigma_c}\right)^n \tag{7.53}$$

där $\dot{\varepsilon}$ är kryptöjningshastigheten, σ spänningen och σ_c är en referensspänning. Tidskonstanten τ och exponenten n är parametrar som bestäms ur experiment. Detta kallas *kalibrering*. Parametrarna i Nortons modell skall anpassas till experimentella krypdata för koppar vid $T = 250$ °C. Vi har följande experimentella värden:

i	1	2	3	4	5
σ [MPa]	49.1	75.6	105.1	130.5	150.4
$\dot{\varepsilon}$ [s^{-1}]	1.4×10^{-8}	8.7×10^{-8}	1.7×10^{-6}	3.4×10^{-5}	2.3×10^{-3}

Ur tabellen ser vi att vid spänningsnivån $\sigma = 49.1$ MPa är töjningshastigheten $1.4 \times 10^{-8}\ s^{-1}$, dvs det tar $t = 0.01/1.4 \times 10^{-8}\ s^{-1} \approx 7.1 \times 10^5$ s ≈ 200 h för materialet att krypa 1 %.
Metoden i sektion 7.1 kan inte användas direkt eftersom modellen inte beror linjärt på parametrarna. Vi skriver därför om Nortons modell (7.53) genom att förlänga med en referenstid τ_c och logaritmera båda sidor:

$$\ln\left(\tau_c\dot{\varepsilon}\right) = -\ln\left(\frac{\tau}{\tau_c}\right) + n\ln\left(\frac{\sigma}{\sigma_c}\right) \tag{7.54}$$

(Obs att $\tau_c\dot{\varepsilon}$, τ/τ_c, σ/σ_c är dimensionslösa så att logaritmering är tillåten.) Vi väljer referenstiden $\tau_c = 1$ s och referensspänningen $\sigma_c = 1$ MPa. Varje mätpunkt ger då en linjär ekvation för de obekanta $\beta_0 = -\ln(\tau/\tau_c)$ och $\beta_1 = n$.

(a) Skriv detta ekvationssystem på formen $X\beta = b$.

(b) Bestäm minsta kvadratlösningen, dels genom att lösa normalekvationerna $(X^T X)\beta = X^T b$ med hjälp av `rref` och dels med `beta = X\b`.

(c) När parametrarna n och τ har bestämts, rita grafer till sambanden (7.53) och (7.54) på intervallet [49, 151] MPa. De experimentellt bestämda punkterna skall också läggas in i figurerna.

Problem 7.2 [Lotka–Volterra]

I det här problemet får du tillämpa dina kunskaper om egenvärdesproblem för att studera stabiliteten hos jämviktspunkter till ett ickelinjärt begynnelsevärdesproblem.

I avsnitt 7.2 i del II studerade vi Lotka–Volterras ekvation som beskriver dynamiken hos en population av bytes- och rovdjur:

$$\begin{cases} \dot{u}_1(t) = au_1(t) - bu_1(t)u_2(t) \\ \dot{u}_2(t) = -cu_2(t) + du_1(t)u_2(t) \end{cases} \tag{7.55}$$

Här är $u_1(t)$ antalet bytesdjur (kaniner) och $u_2(t)$ antalet rovdjur (rävar) vid tiden t. Med matrisbeteckningar får vi

$$\dot{u}(t) = f(u(t)) \quad \text{med } f(u) = \begin{bmatrix} au_1 - bu_1u_2 \\ -cu_2 + du_1u_2 \end{bmatrix} \tag{7.56}$$

Ekvationen $f(u) = 0$ har två lösningar, dvs två jämviktspunkter, $\bar{u} = (0, 0)$ och $\bar{u} = (\frac{c}{d}, \frac{a}{b})$. Vi linjäriserar runt jämviktspunkterna som i exempel 7.2. Derivatan av f är matrisen

$$f'(u) = \begin{bmatrix} \dfrac{\partial f_1}{\partial u_1} & \dfrac{\partial f_1}{\partial u_2} \\ \dfrac{\partial f_2}{\partial u_1} & \dfrac{\partial f_2}{\partial u_2} \end{bmatrix} = \begin{bmatrix} a - bu_2 & -bu_1 \\ du_2 & -c + du_1 \end{bmatrix} \tag{7.57}$$

(Vi deriverar komponenterna f_1 och f_2 med avseende på variablerna u_1 och u_2 i tur och ordning.) Vi får linjäriserade ekvationer $\dot{x} = Ax$ med matriserna $A = f'(0, 0)$ och $A = f'(\frac{c}{d}, \frac{a}{b})$. Dessa undersöks som i exempel 7.3 och exempel 7.4. Tag för enkelhets skull $a = 1, c = 1$.

(a) Undersök stabiliteten hos jämviktspunkten $\bar{u} = (0, 0)$ genom att lösa egenvärdesproblemet för $A = f'(0, 0)$ symboliskt. Rita lösningskurvor och fasporträtt för $u(t) = \bar{u} + x(t)$.

(b) Samma uppgift för jämviktspunkten $\bar{u} = (\frac{c}{d}, \frac{a}{b})$ och matrisen $A = f'(\frac{c}{d}, \frac{a}{b})$.

(c) Lös egenvärdesproblemen numeriskt. Lös $\dot{x} = Ax$ numeriskt och plotta lösningskurvor och fasporträtt för $u(t) = \bar{u} + x(t)$. (Tag till exempel $b = 0.01$, $d = 0.02$.)

(d) Lös den ickelinjära ekvationen $\dot{u} = f(u)$ numeriskt och plotta lösningskurvor och fasporträtt med startpunkter både nära och långt från jämviktspunkterna. Jämför med plottarna från de linjäriserade ekvationerna.

Lilla grekiska alfabetet

α	β	γ	δ	ϵ	ζ
alfa	beta	gamma	delta	epsilon	zeta
η	θ	ι	κ	λ	μ
eta	theta	iota	kappa	lambda	mu
ν	ξ	o	π	ρ	σ
nu	xi	omicron	pi	rho	sigma
τ	υ	ϕ	χ	ψ	ω
tau	ypsilon	phi	chi	psi	omega

Bokstäverna ϵ, θ, σ och ϕ finns också i varianterna ε, ϑ, ς och φ.

Stora grekiska alfabetet

A	B	Γ	Δ	E	Z
Alfa	Beta	Gamma	Delta	Epsilon	Zeta
H	Θ	I	K	Λ	M
Eta	Theta	Iota	Kappa	Lambda	Mu
N	Ξ	O	$\prod$	P	Σ
Nu	Xi	Omicron	Pi	Rho	Sigma
T	Y	Φ	X	Ψ	Ω
Tau	Ypsilon	Phi	Chi	Psi	Omega

Programmering

Bokens datorövningar kan lösas i MATLAB eller valfritt programmeringsspråk (Fortran, C, C++, C#, Java, Python, Julia, Haskell, …). Här sammanfattas några vanliga kommandon som kan vara användbara för att lösa bokens datorövningar i Python eller MATLAB. Python-modulen `pylab`[1] ger tillgång till funktionalitet liknande den i MATLAB. Detta gör att många av de kommandon som krävs för att lösa bokens datorövningar är identiska i MATLAB och Python.

Uppstart

Python	*MATLAB*
`from pylab import *`	
Placeras först i skriptet	*Kräver ingen import av extra funktionalitet*

Visa plottar

Python	*MATLAB*
`show()`	
Placeras sist i skriptet	*Kräver inget kommando; plottar visas direkt*

Iterera ett givet antal gånger

Python	*MATLAB*

`for n in range(10):`
`  ...`
	`for n=1:10`
	`  ...`
	`end`

Itererar över $n = 0, 1, \ldots, 9$	*Itererar över* $n = 1, 2, \ldots, 10$

Upprepa så länge ett villkor är uppfyllt

Python	*MATLAB*

`while condition:`
`  ...`
	`while condition`
	`  ...`
	`end`

Upprepar så länge `condition` *är uppfyllt*	*Upprepar så länge* `condition` *är uppfyllt*

1. Python-modulen `pylab` kombinerar de två Python-modulerna `numpy` och `matplotlib`, vilka ger tillgång till vektorer och matriser (arrayer) och plottning.

Definiera villkorssatser

Python	*MATLAB*

`if condition0:`
`  ...`
`elif condition1:`
`  ...`
`else:`
`  ...`
	`if condition1`
	`  ...`
	`elseif condition2`
	`  ...`
	`else`
	`  ...`
	`end`

Satserna ... utförs om villkoren är uppfyllda	*Satserna ... utförs om villkoren är uppfyllda*

Definiera funktioner

Python	*MATLAB*

`def foo(x):`
`  ...`
`  y = ...`
`  return y`
	`function y = foo(x)`
	`  ...`
	`  y = ...;`
	`end`

Kan läggas ovanför i samma fil	*Läggs i separat fil med namn* `foo.m`

Skapa linjär vektor (array) med x-värden

Python	*MATLAB*
`x = linspace(a, b, 101)`	`x = linspace(a, b, 101);`
Ger 101 punkter och 100 intervall	*Ger 101 punkter och 100 intervall*

Skapa logaritmisk vektor (array) med x-värden

Python	*MATLAB*
`x = logspace(-16, 16, 33)`	`x = logspace(-16, 16, 33);`
Ger 33 punkter: 1×10^{-16}, 1×10^{-15}, ...	*Ger 33 punkter:* 1×10^{-16}, 1×10^{-15}, ...

Plotta en funktion

Python	*MATLAB*
`plot(x, y)`	`plot(x, y)`
Plottar talparen från vektorerna	*Plottar talparen från vektorerna*

Sätta axelmarkörer och titlar

Python	*MATLAB*

`title('y = f(x)')`
`xlabel('x')`
`ylabel('y')`
	`title('y = f(x)')`
	`xlabel('x')`
	`ylabel('y')`

Använd `'$x$'` *för LaTeX-notation*	

Anpassa linjärt polynom till punkter

Python	*MATLAB*
`p = polyfit(x, y, 1)`	`p = polyfit(x, y, 1);`
Polynomet ges av `p[0]*x+p[1]`	*Polynomet ges av* `p(1)*x+p(2)`

Evaluera polynom i punkter

Python	*MATLAB*
`y = polyval(p, x)`	`y = polyval(p, x);`
Beräknar polynomet `p` *i punkterna* `x`	*Beräknar polynomet* `p` *i punkterna* `x`

Spara plot till fil

Python	*MATLAB*
`savefig('foo.png')` `savefig('foo.pdf')`	`print('foo', '-dpng')` `print('foo', '-pdf')`
Sparar till PNG- eller PDF-format	*Sparar till PNG- eller PDF-format*

Facit

Kapitel 1

Övningar

Ö1.1 Rita figurer.

Ö1.2 Rita figurer.

Ö1.3 Vi utvecklar kvadraten

$$\begin{aligned} |\boldsymbol{u} - 5\boldsymbol{v}|^2 &= (\boldsymbol{u} - 5\boldsymbol{v}) \cdot (\boldsymbol{u} - 5\boldsymbol{v}) = |\boldsymbol{u}|^2 - 10\boldsymbol{u} \cdot \boldsymbol{v} + 25|\boldsymbol{v}|^2 \\ &= |\boldsymbol{u}|^2 - 10|\boldsymbol{u}||\boldsymbol{v}| \cos(\theta) + 25|\boldsymbol{v}|^2 = 4 - 60\cos(\theta) + 225 \end{aligned}$$

(a) $\cos(\pi/6) = \frac{\sqrt{3}}{2}$, $\sqrt{229 - 30\sqrt{3}}$ (b) $\cos(\pi/2) = 0$, $\sqrt{229}$
(c) $\cos(\pi/4) = \frac{1}{\sqrt{2}}$, $\sqrt{229 - \frac{60}{\sqrt{2}}}$ (d) $\cos(\pi) = -1$, $\sqrt{289}$

Ö1.4 Vi har skalära projektionen $\boldsymbol{v} \cdot \hat{\boldsymbol{u}} = \frac{\boldsymbol{v} \cdot \boldsymbol{u}}{|\boldsymbol{u}|} = \frac{|\boldsymbol{v}||\boldsymbol{u}| \cos(\theta)}{|\boldsymbol{u}|} = 3\cos(\theta)$ och vektorprojektionen $P_{\boldsymbol{u}}(\boldsymbol{v}) = (\boldsymbol{v} \cdot \hat{\boldsymbol{u}})\hat{\boldsymbol{u}} = \frac{\boldsymbol{v} \cdot \boldsymbol{u}}{|\boldsymbol{u}|^2}\boldsymbol{u} = \frac{3}{2}\cos(\theta)\boldsymbol{u}$.
(a) $\cos(\pi/6) = \frac{\sqrt{3}}{2}$, $\boldsymbol{v} \cdot \hat{\boldsymbol{u}} = \frac{3\sqrt{3}}{2}$, $P_{\boldsymbol{u}}(\boldsymbol{v}) = \frac{3\sqrt{3}}{4}\boldsymbol{u}$
(b) $\cos(\pi/2) = 0$, $\boldsymbol{v} \cdot \hat{\boldsymbol{u}} = 0$, $P_{\boldsymbol{u}}(\boldsymbol{v}) = \boldsymbol{0}$
(c) $\cos(\pi/4) = \frac{1}{\sqrt{2}}$, $\boldsymbol{v} \cdot \hat{\boldsymbol{u}} = \frac{3}{\sqrt{2}}$, $P_{\boldsymbol{u}}(\boldsymbol{v}) = \frac{3}{2\sqrt{2}}\boldsymbol{u}$
(d) $\cos(\pi) = -1$, $\boldsymbol{v} \cdot \hat{\boldsymbol{u}} = -3$, $P_{\boldsymbol{u}}(\boldsymbol{v}) = -\frac{3}{2}\boldsymbol{u}$

Ö1.5 Vi har $|\boldsymbol{u} + \boldsymbol{v}|^2 = |\boldsymbol{u}|^2 + 2|\boldsymbol{u}||\boldsymbol{v}| \cos(\theta) + |\boldsymbol{v}|^2$ så att

$$\begin{aligned} \cos(\theta) &= \frac{|\boldsymbol{u} + \boldsymbol{v}|^2 - |\boldsymbol{u}|^2 - |\boldsymbol{v}|^2}{2|\boldsymbol{u}||\boldsymbol{v}|} = \frac{4 - 9 - 16}{2 \cdot 3 \cdot 4} = \frac{-21}{24} = \frac{-7}{8} \\ \theta &= \arccos(-7/8) \approx -2.6362 \end{aligned}$$

Ö1.6 (a) $\boldsymbol{u} + 2\boldsymbol{v} = (3, 4, -1)$ (b) $\boldsymbol{u} + \boldsymbol{v} + \boldsymbol{w} = (5, 8, 7)$
(c) $\boldsymbol{u} - \boldsymbol{v} = (-3, 1, 5)$ (d) $\hat{\boldsymbol{v}} = (\frac{2}{3}, \frac{1}{3}, -\frac{2}{3})$

Ö1.7 $\arccos(-41/\sqrt{29 \cdot 153})$

Ö1.8 $t = 0$ och $t = 1/5$.

Ö1.9 (a) $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$ (b) Nollvektorn kan ej normeras.
(c) $(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}})$ (d) $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$

Ö1.10 (a) $|\boldsymbol{u}| = \sqrt{14}$, $|\boldsymbol{v}| = 3$ (b) $\boldsymbol{u} \cdot \boldsymbol{v} = -6$ (c) $\theta = \arccos(-2/\sqrt{14}) \approx 2.1347 \approx 122.3°$
(d) $P_{\boldsymbol{u}}(\boldsymbol{v}) = (\boldsymbol{v} \cdot \hat{\boldsymbol{u}})\hat{\boldsymbol{u}} = \frac{-6}{14}(-1, 2, 3) = (\frac{3}{7}, -\frac{6}{7}, -\frac{9}{7})$

Ö1.11 Se exempel 1.5, $\boldsymbol{u} = \boldsymbol{u}_1 + \boldsymbol{u}_2 = P_{\boldsymbol{v}}(\boldsymbol{u}) + (\boldsymbol{u} - P_{\boldsymbol{v}}(\boldsymbol{u}))$.
(a) $\boldsymbol{u}_1 = (\frac{2}{3}, \frac{2}{3}, \frac{2}{3})$, $\boldsymbol{u}_2 = (-\frac{2}{3}, \frac{1}{3}, \frac{1}{3})$ (b) $\boldsymbol{u}_1 = (\frac{2}{3}, \frac{2}{3}, \frac{2}{3})$, $\boldsymbol{u}_2 = (\frac{1}{3}, \frac{1}{3}, -\frac{2}{3})$

Ö1.12 $\arccos(4/\sqrt{21}) \approx 0.5097 \approx 29.2°$

Ö1.13 (a) spetsig (b) rät (c) trubbig (d) spetsig om $t < 3$, rät om $t = 3$, trubbig om $t > 3$

Ö1.14 (a) 0 (b) 6 (c) 12 (d) 0

Ö1.15 (a) $9/2$ (b) $3\sqrt{11}/2$

Ö1.16 1

Ö1.17 De ortogonala vektorerna ges av $\boldsymbol{N} = \pm(\boldsymbol{u} \times \boldsymbol{v})$. Efter normering får vi $\hat{\boldsymbol{N}} = \pm\dfrac{\boldsymbol{u} \times \boldsymbol{v}}{|\boldsymbol{u} \times \boldsymbol{v}|}$.
(a) $\hat{\boldsymbol{N}} = \pm(1, 1, 1)/\sqrt{3}$ (b) $\hat{\boldsymbol{N}} = \pm(1, 1, -1)/\sqrt{3}$

Ö1.18 $\boldsymbol{M} = (13, 2, 18)$ $\texttt{[Nm]}$

Ö1.19

$$\text{(a)} \begin{cases} x = 1 + t, \\ y = 3 + 4t, \\ z = 2 + t, \end{cases} \quad -\infty < t < \infty; \quad \frac{x-1}{1} = \frac{y-3}{4} = \frac{z-2}{1}$$

$$\text{(b)} \begin{cases} x = 1 + t, \\ y = 2 - 5t, \\ z = 3 - 2t, \end{cases} \quad -\infty < t < \infty; \quad \frac{x-1}{1} = \frac{y-2}{-5} = \frac{z-3}{-2}$$

$$\text{(c)} \begin{cases} x = 2, \\ y = -3 + 5t, \\ z = 1 + 2t, \end{cases} \quad -\infty < t < \infty; \quad x - 2 = 0, \ \frac{y+3}{5} = \frac{z-1}{2}$$

Ö1.20 Vi normerar riktningsvektorn: $\boldsymbol{r} = \boldsymbol{r}_0 + t\hat{\boldsymbol{v}}$ och tar $t = \pm 1$. Vi får $\boldsymbol{r} = \boldsymbol{r}_0 + (\pm 1)\hat{\boldsymbol{v}} = \boldsymbol{r}_0 \pm \boldsymbol{v}/|\boldsymbol{v}|$.
(a) $(1 \pm \frac{1}{\sqrt{18}}, 3 \pm \frac{4}{\sqrt{18}}, 2 \pm \frac{1}{\sqrt{18}})$
(b) $(1 \pm \frac{1}{\sqrt{30}}, 2 \pm \frac{-5}{\sqrt{30}}, 3 \pm \frac{-2}{\sqrt{30}})$
(c) $(2, 2 \pm \frac{5}{\sqrt{29}}, 3 \pm \frac{2}{\sqrt{29}})$

Ö1.21

$$\begin{cases} x = 0, \\ y = 0, \\ z = t, \end{cases} \quad -\infty < t < \infty \quad \text{eller} \quad \boldsymbol{r} = t\boldsymbol{e}_z, \ -\infty < t < \infty$$

Ö1.22 (a) $\boldsymbol{v} = (3, 1, -2)$, $P = (1, 2, -3)$ (b) $\boldsymbol{v} = (0, -5, 4)$, $P = (2, 2, -3)$
(c) Dividera först med 6. $\boldsymbol{v} = (3, 6, -2)$, $P = (0, -1, 1)$

Ö1.23 Vi har $P = (1, 3, 2)$ och läser av $\boldsymbol{v} = (1, 2, 3)$, $P_0 = (1, 5, 2)$. Vi får $\overrightarrow{P_0P} = (0, -2, 0)$ och

$$\boldsymbol{v} \times \overrightarrow{P_0P} = \begin{vmatrix} \boldsymbol{e}_x & \boldsymbol{e}_y & \boldsymbol{e}_z \\ 1 & 2 & 3 \\ 0 & -2 & 0 \end{vmatrix} = (6, 0, -2),\ |\boldsymbol{v} \times \overrightarrow{P_0P}| = \sqrt{40},\ |\boldsymbol{v}| = \sqrt{14}$$

Avståndet blir $d = \sqrt{40}/\sqrt{14} = 2\sqrt{5/7}$.

Ö1.24 (a) $(-1, 3, 0)$, $\boldsymbol{N} = (5, -2, 1)$ (b) $(6, 0, 0)$, $\boldsymbol{N} = (1, 2, 3)$
(c) $(0, 0, 0)$, $\boldsymbol{N} = (\frac{1}{3}, -\frac{1}{4}, 1)$ (d) $(0, 0, -6)$, $\boldsymbol{N} = (11, -4, -1)$

Ö1.25 $11x - 4y - z = 6$

Ö1.26 $(1, -1, 1)$

Problem

P1.1 Acceleration $\boldsymbol{a}$ med $a = |\boldsymbol{a}|$ $[\mathtt{m/s}^2]$. Elektriskt fält $\boldsymbol{E}$ `[V/m]`. Vridmoment $\boldsymbol{M} = \boldsymbol{r} \times \boldsymbol{F}$ `[Nm]`.

P1.12 (a) 8 `[J]` (b) 0 `[J]` (c) -3 `[J]`

P1.15 Dela upp vektorn $\overrightarrow{P_0P}$ i ortogonala komposanter $\overrightarrow{P_0P} = \boldsymbol{w}_1 + \boldsymbol{w}_1$, där där $\boldsymbol{w}_1$ är parallell/antiparallell med $\boldsymbol{v}$ och $\boldsymbol{w}_2$ är ortogonal mot $\boldsymbol{v}$, dvs ortogonal mot linjen, se exempel 1.5. Det sökta avståndet är $d = |\boldsymbol{w}_2| = |\overrightarrow{P_0P}| \sin(\theta)$ där θ är vinkeln mellan $\overrightarrow{P_0P}$ och $\boldsymbol{v}$. Men $|\overrightarrow{P_0P} \times \boldsymbol{v}| = |\boldsymbol{v}||\overrightarrow{P_0P}| \sin(\theta)$ så att $d = |\overrightarrow{P_0P} \times \boldsymbol{v}|/|\boldsymbol{v}|$.

P1.16 $\pi/3$

P1.17 $(-2, -2, 1)$

Datorövningar

D1.1

MATLAB

```
u = [-1;2;3]
v = [2;1;-2]
normu = norm(u)
normv = norm(v)
hatu = u/norm(u)          % normerad vektor
hatv = v/norm(v)          % normerad vektor
s = u'*v                  % skalärprodukt
theta = acos(hatu'*hatv)  % vinkel
projuv = (v'*hatu)*hatu   % vektorprojektion av v på u
```

D1.2

MATLAB

```
function s = skalarprodukt(u,v)
  s = u'*v;
end
```

MATLAB

```
skalarprodukt = @(u,v) u'*v
```

D1.3

MATLAB

```
function s = sproj(u,v)
  % skalär projektion av v på u
  hatu = u/norm(u);
  s = v'*hatu;
end
```

D1.4

MATLAB

```
function w = proj(u,v)
  % vektorprojektion av v på u
  hatu = u/norm(u);
  w = (v'*hatu)*hatu;
end
```

D1.5

MATLAB

```
function w = kryss(u,v)
  w = zeros(3,1);                 % skapa kolonnvektor
  w(1) = u(2)*v(3) - u(3)*v(2);
  w(2) = u(3)*v(1) - u(1)*v(3);
  w(3) = u(1)*v(2) - u(2)*v(1);
end
```

D1.6

MATLAB

```
function a = ortotest(u,v)
  a = (u'*v == 0);
end
```

D1.7

MATLAB

```
function a = paratest(u,v)
  a = (norm(kryss(u,v)) == 0);
end
```

D1.8

MATLAB

```
v = [1;-2;4]
lv = length(v)
sv = size(v)
av = abs(v)
mv = v'*v
kv = norm(v)^2
nv = norm(v)
```

D1.9

MATLAB

```
function [v1,v2] = ortodekomposition(u,v)
  % ortogonal uppdelning av v
  hatu = u/norm(u);
  v1 = (v'*hatu)*hatu;  % vektorprojektion av v på u
  v2 = v - v1;
end
```

D1.10

MATLAB

```
function a = triangelarea(A,B,C)
  % arean av triangeln med hörn i A, B, C
  u = B - A;  % vektorn AB
  v = C - A;  % vektorn AC
  a = 0.5*norm(kryss(u,v));
end
```

Kapitel 2

Övningar

Ö2.1 (a) $x = \begin{bmatrix} 6 \\ -7 \end{bmatrix}$ (b) $x = \begin{bmatrix} \frac{3}{4} \\ \frac{3}{4} \\ \frac{1}{2} \end{bmatrix}$ (c) $x = \begin{bmatrix} -1 \\ 0.4 \\ 0.3 \end{bmatrix}$ (d) $x = \begin{bmatrix} \frac{1}{2} \\ -1 \\ -\frac{1}{4} \end{bmatrix}$

Ö2.2 (a) $x = \begin{bmatrix} -0.25 \\ -1.25 \\ 2.25 \end{bmatrix}$ (b) $x = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$ (c) $x = \begin{bmatrix} -\frac{13}{30} \\ -\frac{5}{6} \\ \frac{1}{30} \end{bmatrix}$ (d) $x = \begin{bmatrix} 13.2 \\ 7 \\ -\frac{32}{15} \\ 1 \end{bmatrix}$

Ö2.3 (a) Lösning saknas. (b) $x = \begin{bmatrix} 9+17t \\ -4-7t \\ t \end{bmatrix}$ (c) Lösning saknas. (d) Lösning saknas.

Ö2.4 (a) $x = \begin{bmatrix} 1+t \\ -2t \\ t \end{bmatrix}$ (b) $x = \begin{bmatrix} -\frac{19}{27} - \frac{2}{27}t \\ -\frac{7}{9} + \frac{4}{9}t \\ t \end{bmatrix}$ (c) $x = \begin{bmatrix} -\frac{41}{12} + \frac{5}{2}t \\ t \\ \frac{4}{3} \end{bmatrix}$
(d) Lösning saknas

Ö2.5 (a) $x = \begin{bmatrix} -\frac{4}{5} + \frac{3}{5}t \\ \frac{26}{15} - \frac{7}{15}t \\ t \end{bmatrix}$ (b) $x = \begin{bmatrix} -2 \\ \frac{49}{10} \\ 4 \end{bmatrix}$ (c) $x = \begin{bmatrix} -\frac{1}{3} \\ -\frac{26}{3} \\ 15 \end{bmatrix}$ (d) Lösning saknas.

Ö2.6 (a) $x = \begin{bmatrix} \frac{83}{12} \\ \frac{7}{4} \\ -\frac{11}{12} \\ \frac{19}{6} \end{bmatrix}$ (b) $x = \begin{bmatrix} -\frac{51}{2} \\ 85 \\ 30 \\ -4 \end{bmatrix}$ (c) $x = \begin{bmatrix} -\frac{4}{99} \\ \frac{4}{9} \\ -\frac{2}{9} \\ -\frac{1}{9} \end{bmatrix}$ (d) $x = \begin{bmatrix} \frac{17}{63} \\ \frac{1}{54} \\ -\frac{1}{27} \\ 0 \end{bmatrix}$

Ö2.7 (a) $\begin{bmatrix} 1 \\ -6 \\ 6 \end{bmatrix}$ (b) $\begin{bmatrix} 1 \\ \frac{1}{2} \\ -2 \end{bmatrix}$ (c) $\begin{bmatrix} 12 \\ -7 \\ -8 \end{bmatrix}$ (d) $\begin{bmatrix} 1 \\ 7 \\ -10 \end{bmatrix}$

Ö2.8 (a) $\begin{bmatrix} 14 \\ 15 \end{bmatrix}$ (b) $\begin{bmatrix} 5 \\ -32 \\ 20 \\ 13 \end{bmatrix}$ (c) $\begin{bmatrix} -13 \\ -23 \\ 11 \\ 0 \end{bmatrix}$ (d) $\begin{bmatrix} 45 \\ 30 \\ 31 \\ 44 \end{bmatrix}$

Ö2.9 (a) $\begin{bmatrix} 14 \\ 20 \\ -18 \\ 18 \end{bmatrix}$ (b) $\begin{bmatrix} -12 \\ 44 \\ 26 \\ -28 \\ 14 \end{bmatrix}$ (c) $\begin{bmatrix} -90 \\ 25 \end{bmatrix}$ (d) $\begin{bmatrix} \frac{49}{2} \\ -\frac{19}{2} \\ \frac{119}{5} \end{bmatrix}$

Ö2.10 (a) $x = \begin{bmatrix} -\frac{1}{2} + \frac{3}{2}t \\ 1+t \\ t \end{bmatrix}$ (b) $x = \begin{bmatrix} -\frac{1}{2} \\ 1 \end{bmatrix}$ (c) $x = \begin{bmatrix} -\frac{4}{3} + \frac{5}{6}t \\ -\frac{19}{3} + \frac{13}{3}t \\ -2+t \\ t \end{bmatrix}$
(d) Lösning saknas.

Ö2.11 (a) Nej, $b \notin \operatorname{span}\{a_1, a_2\}$ (b) Ja, $b = 2a_1 - a_2$ (c) Ja, $b = -\frac{3}{2}a_1 + \frac{1}{2}a_2$
(d) Ja, $b = 3a_2$

Ö2.12 (a) $x = \begin{bmatrix} \frac{1}{7}t \\ \frac{19}{14}t \\ t \end{bmatrix}$ (b) $x = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ (c) $x = \begin{bmatrix} \frac{3}{2}t \\ -\frac{2}{3}t \\ t \end{bmatrix}$ (d) $x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Ö2.13 (a) $x_{\mathrm{h}} = t \begin{bmatrix} \frac{1}{7} \\ \frac{19}{14} \\ 1 \end{bmatrix}$, $x_{\mathrm{p}} = \begin{bmatrix} \frac{2}{7} \\ \frac{5}{7} \\ 0 \end{bmatrix}$ (b) $x_{\mathrm{h}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, $x_{\mathrm{p}} = \begin{bmatrix} 0 \\ -\frac{1}{4} \\ -\frac{3}{4} \end{bmatrix}$
(c) $x_{\mathrm{h}} = t \begin{bmatrix} \frac{3}{2} \\ -\frac{2}{3} \\ 1 \end{bmatrix}$, $x_{\mathrm{p}} = \begin{bmatrix} \frac{3}{2} \\ \frac{2}{9} \\ 0 \end{bmatrix}$ (d) $x_{\mathrm{h}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $x_{\mathrm{p}} = \begin{bmatrix} -\frac{1}{2} \\ 0 \end{bmatrix}$

Ö2.14 (a) $x_\mathrm{h} = t\begin{bmatrix} \frac{29}{5} \\ \frac{23}{5} \\ -2 \\ 1 \end{bmatrix}, x_\mathrm{p} = \begin{bmatrix} -8 \\ -5 \\ 3 \\ 0 \end{bmatrix}$ (b) $x_\mathrm{h} = t\begin{bmatrix} \frac{3}{4} \\ \frac{2}{3} \\ \frac{5}{2} \\ 1 \end{bmatrix}, x_\mathrm{p} = \begin{bmatrix} -\frac{1}{4} \\ \frac{14}{9} \\ -\frac{19}{6} \\ 0 \end{bmatrix}$

(c) $x_\mathrm{h} = t\begin{bmatrix} \frac{10}{13} \\ -\frac{1}{13} \\ 1 \end{bmatrix}, x_\mathrm{p} = \begin{bmatrix} \frac{46}{39} \\ -\frac{41}{39} \\ 0 \end{bmatrix}$ (d) $x_\mathrm{h} = t\begin{bmatrix} \frac{29}{3} \\ -7 \\ 1 \end{bmatrix}, x_\mathrm{p} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$

Ö2.15 (a) $x = s\begin{bmatrix} \frac{3}{2} \\ -2 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} \frac{9}{10} \\ -\frac{1}{5} \\ 0 \\ 1 \end{bmatrix}$ (b) $x = s\begin{bmatrix} \frac{30}{23} \\ -\frac{19}{23} \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} \frac{10}{23} \\ -\frac{14}{23} \\ 0 \\ 1 \end{bmatrix}$

(c) $x = s\begin{bmatrix} \frac{7}{2} \\ 5 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} \frac{3}{2} \\ 2 \\ 0 \\ 1 \end{bmatrix}$ (d) $x = s\begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -5 \\ -\frac{5}{2} \\ 0 \\ 1 \\ 0 \end{bmatrix} + u\begin{bmatrix} -3 \\ -\frac{1}{2} \\ 0 \\ 0 \\ 1 \end{bmatrix}$

Ö2.16 (a) Linjärt oberoende (b) Linjärt beroende (c) Linjärt beroende
(d) Linjärt oberoende

Ö2.17 (a) Linjärt oberoende (b) Linjärt beroende (c) Linjärt oberoende
(d) Linjärt beroende

Ö2.18 (a) Nej (b) Nej (c) Nej (d) Nej

Ö2.19 (a) Ja (b) Nej (c) Nej (d) Ja

Ö2.20 (a) $c_1 = 1, c_2 = \frac{1}{2}$ (b) $c_1 = -1, c_2 = 2$ (c) $c_1 = 0, c_2 = \frac{4}{3}$ (d) $c_1 = 4, c_2 = 3$

Ö2.21 (a) $A = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \end{bmatrix}$
(c) $A = \begin{bmatrix} 1 & 0 & -5 \end{bmatrix}$ (d) $m = n$ och $A = I \in \mathbb{R}^{n\times n}$

Ö2.22 (a) $f(0) \neq 0$ (b) $f(2x) = 4f(x) \neq 2f(x)$ om $x \neq 0$
(c) $f\left((-1)\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = f\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) \neq (-1)f\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$

Ö2.23 $\frac{1}{5}\begin{bmatrix} -1 & 3 \\ 2 & -6 \\ -3 & 9 \end{bmatrix}$

Kapitel 3

Övningar

Ö3.1 (a) $\begin{bmatrix} 4 & 8 & 15 \\ 4 & 8 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 2 & 1 & -1 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = O_{3\times 2}$ (d) $\begin{bmatrix} 3 & 4 \\ 4 & 5 \\ 8 & -1 \end{bmatrix}$

Ö3.2 $X = (B^T - \frac{1}{2}C)^T = B - \frac{1}{2}C^T = \begin{bmatrix} -1 & 4 & 6 \\ -4 & 1 & 3 \end{bmatrix}$

Ö3.3 (a) $u^Tv = 32,\ uv^T = \begin{bmatrix} 4 & 5 & 6 \\ 8 & 10 & 12 \\ 12 & 15 & 18 \end{bmatrix}$

(b) $u^Tv = 0,\ uv^T = \begin{bmatrix} 2 & 2 & 2 & 2 & 2 & 2 \\ -2 & -2 & -2 & -2 & -2 & -2 \\ 2 & 2 & 2 & 2 & 2 & 2 \\ -2 & -2 & -2 & -2 & -2 & -2 \\ 2 & 2 & 2 & 2 & 2 & 2 \\ -2 & -2 & -2 & -2 & -2 & -2 \end{bmatrix}$

(c) $u^Tv = -3,\ uv^T = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ -1 & -2 & -3 & -4 & -5 & -6 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ -1 & -2 & -3 & -4 & -5 & -6 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ -1 & -2 & -3 & -2 & -5 & -6 \end{bmatrix}$

Ö3.4 (a) $AB = \begin{bmatrix} 7 & 8 \\ 30 & 38 \end{bmatrix}$, $BA = \begin{bmatrix} 17 & 18 \\ 25 & 28 \end{bmatrix}$

(b) $AB = \begin{bmatrix} 9 & 10 & 11 \\ 33 & 41 & 49 \end{bmatrix}$, BA ej definierad

Ö3.5 (a) $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ (b) $\begin{bmatrix} a & -2c \\ c & a+2c \end{bmatrix}$ (c) $\begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$ (d) $\begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 3e-3a-c & e-3b-d & e \end{bmatrix}$

Ö3.6 (a) $\begin{bmatrix} -\frac{1}{2} & \frac{3}{2} \\ 1 & -2 \end{bmatrix}$ (b) A är singulär (c) $A^{-1} = A$ (d) $\frac{1}{3}\begin{bmatrix} 3 & -2 \\ 0 & 1 \end{bmatrix}$

Ö3.7 (a) $\begin{bmatrix} 1 & -2 & 4 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$ (b) $\frac{1}{11}\begin{bmatrix} -2 & -1 & 6 \\ 5 & -3 & -4 \\ -1 & 5 & 3 \end{bmatrix}$

(c) A är singulär, se problem 3.9. (d) $\frac{1}{4}\begin{bmatrix} -3 & -2 & -1 \\ -2 & -4 & -2 \\ -1 & -2 & -3 \end{bmatrix}$

Ö3.8 $A^{-1} = \begin{bmatrix} -1 & -2 & 5 \\ 1 & 1 & -3 \\ -1 & -2 & 4 \end{bmatrix}$, $x = \begin{bmatrix} 10 \\ -6 \\ 7 \end{bmatrix}$

Ö3.9 (a) $X = A^{-1}CB^{-1} = \begin{bmatrix} 1.4 & 0.2 & -2.4 \\ -2.7 & -0.1 & 5.7 \end{bmatrix}$

(b) $X = (A-I)^{-1}CB^{-1} = \begin{bmatrix} 0.75 & 0.25 & -0.75 \\ -0.65 & 0.05 & 1.65 \end{bmatrix}$

Ö3.10 (a) 12 (b) 1 (c) 30 (d) 0 Radreducering ger rad med endast nollor.

Ö3.11 (a) 24 (b) Ej definierad, matrisen ej kvadratisk. (c) -4

Ö3.12 $\det(A^3B^7) = \det(A)^3 \det(B)^7 = 10^3 \cdot (-1)^7 = -1000$

Problem

P3.2 (a) $A^k = \begin{bmatrix} 1 & ka \\ 0 & 1 \end{bmatrix}$

Induktionsbevis: Påståendet är självklart sant då $k = 1$. Antag sant för $k-1$. Då fås $A^{k-1}A = \begin{bmatrix} 1 & (k-1)a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & ka \\ 0 & 1 \end{bmatrix}$

(b) $A^k = \begin{bmatrix} \cos(k\varphi) & -\sin(k\varphi) \\ \sin(k\varphi) & \cos(k\varphi) \end{bmatrix}$

Bevis: A är matrisen för rotation med vinkeln φ och A^k blir då matrisen för rotation med vinkeln $k\varphi$, se exempel 2.13. Vi kan också göra induktionsbevis med trigonometriska formler: $\cos((k-1)\varphi)\cos(\varphi) - \sin((k-1)\varphi)\sin(\varphi) = \cos(k\varphi)$ och så vidare.

P3.3 $A = [a_1 \ \cdots \ a_n]$ med $a_j = v_j u$, dvs $A = [v_1 u \ \cdots v_j u \ \cdots \ v_n u]$. Ax är linjär kombination av $\{a_j\}_{j=1}^n = \{v_j u\}_{j=1}^n$, dvs $Ax = \sum_{j=1}^n x_j a_j = \sum_{j=1}^n x_j v_j u = \alpha u \in \text{span}\{u\}$ med $\alpha = \sum_{j=1}^n x_j v_j$. Alternativt: $Ax = (uv^T)x = u(v^T x) = (v^T x)u = \alpha u$ med samma $\alpha = v^T x$.

P3.4 $A^T = (uu^T)^T = (u^T)^T u^T = uu^T = A$

P3.6 Vi har $a_{ji} = -a_{ij}$. Vi tar $j = i$ och får $a_{ii} = -a_{ii}$ med enda lösningen $a_{ii} = 0$.

P3.7 Multiplicera $BA = I$ med C från höger: $BAC = IC = C$. Här är $AC = I$, så att $BAC = BI = B$. Alltså: $B = C$.

P3.8 —

P3.9 Radreducering ger en rad med enbart nollor. Den raden har ingen pivotposition och matrisen är singulär enligt sats 3.8. Se övning 3.7 (c).

P3.10 —

P3.11 (a) —

(b) Produktregeln (3.105) tillämpad på $O = A^k$ ger $0 = \det(A^k) = \det(A)^k$, vilket leder till $\det(A) = 0$.

P3.12 Produktregeln (3.105) tillämpad på $AB = I$ ger $\det(AB) = \det(A)\det(B) = 1$, vilket visar att $\det(A) \neq 0$, så att A är inverterbar och A^{-1} finns. Multiplicera då $I = AB$ med A^{-1} från vänster: $A^{-1} = A^{-1}I = A^{-1}AB = IB = B$, dvs $A^{-1} = B$.

P3.13 Radreducering med skalning.

Kapitel 4

Övningar

Ö4.1 (a) $u''(x) + 4u(x) = x^2 + 1$, $u_\mathrm{h} \in \mathcal{N}(L) = \operatorname{span}\{\cos(2x), \sin(2x)\}$, $u_\mathrm{p}(x) = \frac{1}{4}x^2 + \frac{1}{8}$
(b) $u''(x) - u(x) = x$, $u_\mathrm{h} \in \mathcal{N}(L) = \operatorname{span}\{\exp(x), \exp(-x)\}$, $u_\mathrm{p}(x) = -x$
(c) $u'(x) = x$, $u_\mathrm{h} \in \mathcal{N}(L) = \mathbb{R}$, $u_\mathrm{p}(x) = \frac{1}{2}x^2$, $u(x) = \frac{1}{2}x^2 + c$
(c) $u'(x) + u(x) = 1$, $u_\mathrm{h} \in \mathcal{N}(L) = \operatorname{span}\{\exp(-x)\}$, $u_\mathrm{p}(x) = 1$

Ö4.2 (a) $\displaystyle\int_0^1 (au(s) + bv(s))\,\mathrm{d}s = a\int_0^1 u(s)\,\mathrm{d}s + b\int_0^1 v(s)\,\mathrm{d}s$
(b) $\displaystyle\int_0^x (au(s) + bv(s))\,\mathrm{d}s = a\int_0^x u(s)\,\mathrm{d}s + b\int_0^x v(s)\,\mathrm{d}s$

Ö4.3 (a) Ja. $c_1\exp(x) + c_2\exp(-x) = 0$, tag $x = 0$ och $x = 1$, $c_1 + c_2 = 0$, $ec_1 + e^{-1}c_2 = 0$, endast triviala lösningen $c_1 = c_2 = 0$.
(b) Ja. $c_1\cos(x) + c_2\cos(2x) = 0$, tag $x = \pi/2$ och $x = 0$, $-c_2 = 0$, $c_1 + c_2 = 0$, endast triviala lösningen.
(c) Nej. $u_2 - u_3 = 4u_1$, tag $x = 0,\ 1,\ -1$.

Ö4.4 $[x]_B = \begin{bmatrix} -5 \\ 4 \end{bmatrix}$

Ö4.5 (a) Nej. Gör radreducering, $P_B c = 0$ har icke-trivial lösning.
(b) Lös $P_B c = x$, $[x]_B = c = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$ (c) Nej, för många.

Ö4.6 (a) $A \sim \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, x_2 fri, $\mathcal{N}(A) = \operatorname{span}\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right\}$
$\mathcal{C}(A) = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix}\right\}$ (pivotkolonnerna), $\operatorname{rang}(A) = 2$
(b) $A \sim \begin{bmatrix} 1 & 0 & -1.5 & -0.9 \\ 0 & 1 & 2 & 0.2 \end{bmatrix}$, x_3, x_4 fria, $\mathcal{N}(A) = \operatorname{span}\left\{\begin{bmatrix} 1.5 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.9 \\ -0.2 \\ 0 \\ 1 \end{bmatrix}\right\}$

$\mathcal{C}(A) = \text{span}\left\{\begin{bmatrix}2\\-6\end{bmatrix}, \begin{bmatrix}4\\-7\end{bmatrix}\right\}$ (pivotkolonnerna), $\text{rang}(A) = 2$

(c) $A \sim \begin{bmatrix}1 & 0 & 0 & 2.5\\0 & 1 & 0 & -1.5\\0 & 0 & 1 & 0.5\\0 & 0 & 0 & 0\end{bmatrix}, \mathcal{N}(A) = \text{span}\left\{\begin{bmatrix}-\frac{5}{2}\\\frac{3}{2}\\-\frac{1}{2}\\1\end{bmatrix}\right\}$

$\mathcal{C}(A) = \text{span}\left\{\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}2\\1\\0\\0\end{bmatrix}, \begin{bmatrix}3\\2\\-5\\1\end{bmatrix}\right\}$ (pivotkolonnerna), $\text{rang}(A) = 3$

Ö4.7 (a) $\|u\| = \|v\| = 2, u \cdot v = 0$ (b) $\|u\| = \sqrt{14}, \|v\| = \sqrt{3}, u \cdot v = 2$

Ö4.8 13

Ö4.9 u_1 och u_2 är ortogonala. Efter normering har vi en ON bas: $\{\hat{u}_1, \hat{u}_2\} = \{\frac{1}{2}u_1, \frac{1}{2}u_2\}$. $P_W(x) = (x \cdot \hat{u}_1)\hat{u}_1 + (x \cdot \hat{u}_2)\hat{u}_2$. Alternativt: med basvektormatrisen $U = [\hat{u}_1\ \hat{u}_2] = \begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\\frac{1}{2} & -\frac{1}{2}\\\frac{1}{2} & \frac{1}{2}\\\frac{1}{2} & -\frac{1}{2}\end{bmatrix}$ blir $P = UU^T = \begin{bmatrix}\frac{1}{2} & 0 & \frac{1}{2} & 0\\0 & \frac{1}{2} & 0 & \frac{1}{2}\\\frac{1}{2} & 0 & \frac{1}{2} & 0\\0 & \frac{1}{2} & 0 & \frac{1}{2}\end{bmatrix}$ och $P_W(x) = Px$.

$x = \begin{bmatrix}5\\1\\5\\1\end{bmatrix}$ (b) $x = \begin{bmatrix}2\\3\\2\\3\end{bmatrix}$ (c) $x = \begin{bmatrix}-2\\2\\-2\\2\end{bmatrix}$

Ö4.10 (a) Gram–Schmidt ger $u_1 = b_1, u_2 = b_2 - \frac{b_2 \cdot u_1}{u_1 \cdot u_1}u_1 = b_2 - 0u_1 = b_2$

$$u_3 = b_3 - \frac{b_3 \cdot u_1}{u_1 \cdot u_1}u_1 - \frac{b_3 \cdot u_2}{u_2 \cdot u_2}u_2 = \begin{bmatrix}1\\3\\2\\-4\end{bmatrix} - \frac{6}{3}\begin{bmatrix}1\\1\\1\\0\end{bmatrix} - \frac{-7}{7}\begin{bmatrix}1\\-2\\1\\1\end{bmatrix} = \begin{bmatrix}0\\-1\\1\\-3\end{bmatrix}$$

Ortogonal bas: $\left\{\begin{bmatrix}1\\1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\-2\\1\\1\end{bmatrix}, \begin{bmatrix}0\\-1\\1\\-3\end{bmatrix}\right\}$

ON bas: $\left\{\frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\\0\end{bmatrix}, \frac{1}{\sqrt{7}}\begin{bmatrix}1\\-2\\1\\1\end{bmatrix}, \frac{1}{\sqrt{11}}\begin{bmatrix}0\\-1\\1\\-3\end{bmatrix}\right\}$

(b) $\left\{\frac{1}{\sqrt{6}}\begin{bmatrix}2\\1\\1\end{bmatrix}, \frac{1}{\sqrt{66}}\begin{bmatrix}-4\\1\\7\end{bmatrix}\right\}$ (c) $\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right\}$

Ö4.11 (a) $\hat{x} = \begin{bmatrix} \frac{3}{5} \\ -\frac{2}{5} \end{bmatrix}$ (b) $\hat{x} = \begin{bmatrix} \frac{19}{11} \\ \frac{3}{11} \end{bmatrix}$ (c) $\hat{x} = \begin{bmatrix} \frac{1}{3} \\ 1 \\ \frac{3}{2} \end{bmatrix}$

(d) $\hat{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ (exakt lösning) (e) $\hat{x} = \begin{bmatrix} \frac{3}{2} \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

Ö4.12 (a) $\arccos(\sqrt{125/128}) \approx 0.1537 \approx 8.8062^\circ$ (b) $\arccos(0) = \pi/2 = 90^\circ$
(c) $\arccos(-\frac{\sqrt{15}}{12}) \approx 1.8994 \approx 108.82^\circ$

Ö4.13 $\frac{1}{\sqrt{k+\frac{1}{2}}}$

Ö4.14 Gram–Schmidt ger

$$\begin{aligned}
q_0(x) &= 1, \quad W_0 = \operatorname{span}\{1\} \\
q_1(x) &= x - \frac{\int_0^1 x \cdot 1 \, \mathrm{d}x}{\int_0^1 1^2 \, \mathrm{d}x} 1 = x - \tfrac{1}{2}, \quad W_1 = \operatorname{span}\{1, x - \tfrac{1}{2}\} = \operatorname{span}\{1, x\} \\
q_2(x) &= x^2 - \frac{\int_0^1 x^2 \cdot 1 \, \mathrm{d}x}{\int_0^1 1^2 \, \mathrm{d}x} 1 - \frac{\int_0^1 x^2 \cdot (x - \frac{1}{2}) \, \mathrm{d}x}{\int_0^1 (x - \frac{1}{2})^2 \, \mathrm{d}x} (x - \tfrac{1}{2}) \\
&= x^2 - \tfrac{1}{3} - 1 \cdot (x - \tfrac{1}{2}) = x^2 - x + \tfrac{1}{6} \\
W_2 &= \operatorname{span}\{1, x, x^2 - x + \tfrac{1}{6}\} = \operatorname{span}\{1, x, x^2\}
\end{aligned}$$

Normera: $\|q_0\| = 1$, $\|q_1\|^2 = \int_0^1 (x - \frac{1}{2})^2 \, \mathrm{d}x = \frac{1}{12}$, $\|q_2\|^2 = \int_0^1 (x^2 - x + \frac{1}{6})^2 \, \mathrm{d}x = \frac{1}{180}$ och $\hat{q}_0(x) = 1$, $\hat{q}_1(x) = 2\sqrt{3}(x - \frac{1}{2})$, $\hat{q}_2(x) = 6\sqrt{5}(x^2 - x + \frac{1}{6})$.

Ö4.15 Gram–Schmidt ger

$$\begin{aligned}
P_0(x) &= 1, \quad W_0 = \operatorname{span}\{1\} \\
P_1(x) &= x - \frac{\int_{-1}^1 x \cdot 1 \, \mathrm{d}x}{\int_{-1}^1 1^2 \, \mathrm{d}x} 1 = x - 0 = x, \quad W_1 = \operatorname{span}\{1, x\} \\
P_2(x) &= x^2 - \frac{\int_{-1}^1 x^2 \cdot 1 \, \mathrm{d}x}{\int_{-1}^1 1^2 \, \mathrm{d}x} 1 - \frac{\int_{-1}^1 x^2 \cdot x \, \mathrm{d}x}{\int_{-1}^1 x^2 \, \mathrm{d}x} x = x^2 - \tfrac{1}{3} - 0x = \tfrac{1}{3}(3x^2 - 1) \\
W_2 &= \operatorname{span}\{1, x, 3x^2 - 1\} = \operatorname{span}\{1, x, x^2\} \\
P_3(x) &= x^3 - \frac{\int_{-1}^1 x^3 \cdot 1 \, \mathrm{d}x}{\int_{-1}^1 1^2 \, \mathrm{d}x} 1 - \frac{\int_{-1}^1 x^3 \cdot x \, \mathrm{d}x}{\int_{-1}^1 x^2 \, \mathrm{d}x} x - \frac{\int_{-1}^1 x^3 \cdot (3x^2 - 1) \, \mathrm{d}x}{\int_{-1}^1 (3x^2 - 1)^2 \, \mathrm{d}x} (3x^2 - 1) \\
&= x^3 - 0 - \tfrac{3}{5}x - 0(3x^2 - 1) = \tfrac{1}{5}(5x^3 - 3x) \\
W_3 &= \operatorname{span}\{1, x, 3x^2 - 1, 5x^3 - 3x\} = \operatorname{span}\{1, x, x^2, x^3\}
\end{aligned}$$

Efter normering med $P_k(1) = 1$ har vi Legendre-polynomen
$P_0(x) = 1,\ P_1(x) = x,\ P_2(x) = \frac{1}{2}(3x^2 - 1),\ P_3(x) = \frac{1}{2}(5x^3 - 3x)$

Problem

P4.2 Utveckla kvadraterna som i beviset av Pythagoras sats.

P4.3 Utveckla kvadraten och använd ortogonalitet, $u_i \cdot u_j = \delta_{ij}$:

$$\|v\|^2 = v \cdot v = \left(\sum_{j=1}^{n} c_j u_j \right) \cdot \left(\sum_{i=1}^{n} c_i u_i \right) = \sum_{j=1}^{n} \sum_{i=1}^{n} c_j c_i (u_j \cdot u_i) = \sum_{i=1}^{n} c_i^2$$

P4.4 Tag $v = [1 \; \dots \; 1]^T$ och använd Cauchy–Schwarz olikhet:

$$\sum_{i=1}^{n} u_i = \sum_{i=1}^{n} u_i v_i = \langle u, v \rangle \leqslant \|u\| \|v\| = \sqrt{\sum_{i=1}^{n} u_i^2} \sqrt{\sum_{i=1}^{n} v_i^2} = \sqrt{\sum_{i=1}^{n} u_i^2} \sqrt{n}$$

Datorövningar

D4.1 —

D4.2

MATLAB

```matlab
function P=ortoproj(A)
  U=orth(A);
  P=U*U';
end
```

D4.3

MATLAB

```matlab
function U = gram_schmidt1(A)
% ortogonalisera kolonnerna i A
% U blir ortogonal bas för kolonnrummet till A

p = size(A,2);    % antal kolonner

for i = 1:p
    U(:,i) = A(:,i);
    for j = 1:i-1
        U(:,i) = U(:,i) - U(:,i)'*U(:,j) / (U(:,j)'*U(:,j))*U(:,j);
    end
    v = U(:,i);
    a = min(abs(v(abs(v) > 0)));
    % a = 1;  % ingen skalning
    % a=norm(U(:,i));  % normera istället
    U(:,i) = U(:,i) / a;  % skalning
end
% a = abs av minsta nollskiljda komponenten
% skalning med 1/a ger minsta komponenten = 1
% snyggare resultat
```

D4.4

MATLAB

```
function U = gram_schmidt2(A)
% ortonormera kolonnerna i A
% U blir ON bas för kolonnrummet till A

n = size(A,1);    % antal rader
p = size(A,2);    % antal kolonner
I = eye(n,n);     % enhetsmatris

U(:,1) = A(:,1);
U(:,1) = U(:,1) / norm(U(:,1));

for i = 2:p
    % b  = A(:,i)
    % U(:,i) = b - U*U'*b
    U(:,i) = (I - U*U')*A(:,i);
    U(:,i) = U(:,i) / norm(U(:,i));
end
```

Kapitel 5

Övningar

Ö5.1 am=algebraisk multiplicitet, gm=geometrisk multiplicitet

(a) $\lambda = 5$, $E_\lambda = \text{span}\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$, am=1, gm=1

$\lambda = -2$, $E_\lambda = \text{span}\left\{\begin{bmatrix}-3\\4\end{bmatrix}\right\}$, am=1, gm=1

(b) (komplexa egenvärden)

$\lambda = \mathrm{i}$, $E_\lambda = \text{span}\left\{\begin{bmatrix}5\\3-\mathrm{i}\end{bmatrix}\right\}$, am=1, gm=1

$\lambda = -\mathrm{i}$, $E_\lambda = \text{span}\left\{\begin{bmatrix}5\\3+\mathrm{i}\end{bmatrix}\right\}$, am=1, gm=1

(c) $\lambda = 0$, $E_\lambda = \text{span}\left\{\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$, am=1, gm=1

$\lambda = 1$, $E_\lambda = \text{span}\left\{\begin{bmatrix}2\\3\\22\end{bmatrix}\right\}$, am=1, gm=1

$\lambda = -1$, $E_\lambda = \text{span}\left\{\begin{bmatrix}0\\1\\-4\end{bmatrix}\right\}$, am=1, gm=1

(d) $\det(A - \lambda I) = -\lambda^3 + 18\lambda^2 - 81\lambda = \lambda(\lambda - 9)^2$

$$\lambda = 0, E_\lambda = \text{span}\left\{\begin{bmatrix}2\\-2\\1\end{bmatrix}\right\}, \text{am=1, gm=1}$$
$$\lambda = 9, E_\lambda = \text{span}\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}-\frac{1}{2}\\0\\1\end{bmatrix}\right\}, \text{am=2, gm=2}$$

Ö5.2 (a) Ej diagonaliserbar. (b) Diagonaliserbar. $P = \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}, D = \begin{bmatrix}1 & 0\\0 & 2\end{bmatrix}$

(c) Diagonaliserbar. $P = \begin{bmatrix}\mathrm{i} & -\mathrm{i}\\1 & 1\end{bmatrix}, D = \begin{bmatrix}\mathrm{i} & 0\\0 & -\mathrm{i}\end{bmatrix}$

(d) Diagonaliserbar. $P = \begin{bmatrix}0 & 2 & 0\\0 & 3 & 1\\1 & 22 & -4\end{bmatrix}, D = \begin{bmatrix}0 & 0 & 0\\0 & 1 & 0\\0 & 0 & -1\end{bmatrix}$ (e) Ej diagonaliserbar.

Ö5.3 $\begin{bmatrix}1 & 1023\\0 & 1024\end{bmatrix}$

Ö5.4 (a) $P = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\-1 & 1\end{bmatrix}, P^TAP = \text{diag}(1,3)$

(b) $P = \frac{1}{\sqrt{5}}\begin{bmatrix}-2 & 1\\1 & 2\end{bmatrix}, P^TAP = \text{diag}(5,10)$

(c) Varje ortogonal matris P fungerar. $P^TAP = A = \text{diag}(a,a)$

(d) $P = \frac{1}{\sqrt{5}}\begin{bmatrix}1 & 2\\-2 & 1\end{bmatrix}, P^TAP = \text{diag}(0,5)$

Ö5.5 (a) $P = \frac{1}{3}\begin{bmatrix}2 & 2 & 1\\-1 & 2 & -2\\2 & -1 & -2\end{bmatrix}, P^TAP = \text{diag}(-3,3,9)$

(b) $P = \frac{1}{3}\begin{bmatrix}1 & 2 & -2\\-2 & 2 & 1\\2 & 1 & 2\end{bmatrix}, P^TAP = \text{diag}(-3,0,3)$

(c) $P = \begin{bmatrix}1 & 0 & 0 & 0\\0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\0 & 0 & 1 & 0\\0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\end{bmatrix}, P^TAP = \text{diag}(1,1,1,5)$

Ö5.6 $A^k = PD^kP^T, D^k = \text{diag}(\lambda_1^k, \ldots, \lambda_n^k)$

(a) $A^k = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\-1 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 3^k\end{bmatrix}\begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}\frac{1}{\sqrt{2}} = \frac{1}{2}\begin{bmatrix}3^k+1 & 3^k-1\\3^k-1 & 3^k+1\end{bmatrix}$

(b) $A^k = \frac{1}{\sqrt{5}}\begin{bmatrix}-2 & 1\\1 & 2\end{bmatrix}\begin{bmatrix}5^k & 0\\0 & 10^k\end{bmatrix}\begin{bmatrix}-2 & 1\\1 & 2\end{bmatrix}\frac{1}{\sqrt{5}} = 5^{k-1}\begin{bmatrix}2^k+4 & 2^{k+1}-2\\2^{k+1}-2 & 2^{k+2}+1\end{bmatrix}$

(c) $A^k = \text{diag}(a^k, a^k)$

(d) $A^k = 3^{k-2}\begin{bmatrix} 4 & -2 & -4 \\ -2 & 1 & 2 \\ -4 & 2 & 4 \end{bmatrix} + (-3)^{k-2}\begin{bmatrix} 1 & -2 & 2 \\ -2 & 4 & -4 \\ 2 & -4 & 4 \end{bmatrix}$

Ö5.7 (a) Positivt definit. Datorberäkning: $\lambda \approx 0.4158,\ 2.2943,\ 6.2899$, positiva

(b) Positivt definit. $\lambda = \frac{1}{2},\ \frac{1}{2},\ 2$

(c) Indefinit. $\lambda = -2,\ 0,\ 3$, ett positivt och ett negativt

(d) Indefinit. Datorberäkning: $\lambda \approx -0.8070,\ 1.1444,\ 8.6625$, ett negativt och två positiva

Ö5.8 $\mathcal{R}(f) = [3, 9]$

$P = \frac{1}{3}\begin{bmatrix} -1 & 2 & -2 \\ -2 & 1 & 2 \\ -2 & -2 & -1 \end{bmatrix}$, $P^TAP = \text{diag}(3, 6, 9)$

$x = Py$, $f(x) = \dfrac{\lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2}{y_1^2 + y_2^2 + y_3^2} = \dfrac{3y_1^2 + 6y_2^2 + 9y_3^2}{y_1^2 + y_2^2 + y_3^2}$; minimum då $y_2 = y_3 = 0$, min $f = \frac{3y_1^2}{y_1^2} = 3$; maximum då $y_1 = y_2 = 0$, max $f = \frac{9y_3^2}{y_3^2} = 9$; värdemängden är intervallet från min till max.

Ö5.9 Minimum är -16 och inträffar då $x = \pm\begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix}$

$x = Py$, $\|y\|^2 = \|x\|^2 = 8$, $\lambda = -2,\ 0,\ 3$, $f(x) = x^TAx = y^TDy = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2$ är minimal då $y_2 = y_3 = 0$, $y_1 = \pm\sqrt{8}$, min $f = \lambda_1 y_1^2 = -2 \cdot 8 = -16$, $y = \pm\sqrt{8}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $x = Py = [v_1\ v_2\ v_3]\begin{bmatrix} \pm\sqrt{8} \\ 0 \\ 0 \end{bmatrix} = \pm\sqrt{8}v_1 = \pm\sqrt{8}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \pm\begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix}$

Problem

P5.1 Vi resonerar som i beviset av sats 5.3:
$\det(AB - \lambda I) = \det((A - \lambda B^{-1})B) = \det((A - \lambda B^{-1})\det(B) = \det(B)\det((A - \lambda B^{-1}) = \det(B(A - \lambda B^{-1})) = \det(BA - \lambda BB^{-1}) = \det(BA - \lambda I)$

P5.2 $\lambda = 0$ är egenvärde $\iff A - 0I = A$ är singulär.

P5.3 $Av = \lambda v \implies A^2v = \lambda Av = \lambda^2 v$ och så vidare.

P5.4 $Av = \lambda v$ och $A^2 = A \implies \lambda v = Av = A^2v = \lambda Av = \lambda^2 v \implies (\lambda^2 - \lambda)v = 0 \implies \lambda(\lambda - 1) = 0$ ty $v \neq 0 \implies \lambda = 0$ eller $\lambda = 1$

Kapitel 6

Övningar

Ö6.1 Se övning 2.1.

(a) $L = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, U = \begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix}, y = \begin{bmatrix} 4 \\ -7 \end{bmatrix}, x = \begin{bmatrix} 6 \\ -7 \end{bmatrix}$

(b) $L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & 1 \end{bmatrix}, U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{bmatrix}, y = \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}, x = \begin{bmatrix} \frac{3}{4} \\ \frac{3}{4} \\ \frac{1}{2} \end{bmatrix}$

(c) $L = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0.5 & 0.125 & 1 \end{bmatrix}, U = \begin{bmatrix} 2 & 3 & -4 \\ 0 & 4 & -12 \\ 0 & 0 & 7.5 \end{bmatrix}, y = \begin{bmatrix} -2 \\ -2 \\ 2.25 \end{bmatrix}, x = \begin{bmatrix} -1 \\ 0.4 \\ 0.3 \end{bmatrix}$

(d) $L = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{3} & 1 & 0 \\ \frac{1}{2} & \frac{1}{4} & 1 \end{bmatrix}, U = \begin{bmatrix} 6 & 3 & 12 \\ 0 & 2 & -16 \\ 0 & 0 & -12 \end{bmatrix}, y = \begin{bmatrix} -3 \\ 2 \\ 3 \end{bmatrix}, x = \begin{bmatrix} \frac{1}{2} \\ -1 \\ -\frac{1}{4} \end{bmatrix}$

Problem

P6.1 $\det(A) = \det(LU) = \det(L)\det(U) = \det(U) = u_{11} \cdots u_{nn}$, ty L är triangulär med 1 på diagonalen och U är triangulär med u_{ii} på diagonalen.

P6.2 Sätt in $K = A - L$. Med $L = A$ blir ekvationen $x = A^{-1}b$.

P6.3 1. $\langle v, u\rangle = v^TAu = (v^TAu)^T = u^TA^Tv = u^TAv = \langle u, v\rangle$ där $v^TAu = (v^TAu)^T$ eftersom v^TAu är en skalär.

2. $\langle u + v, w\rangle = (u + v)^TAw = u^TAw + v^TAw = \langle u, w\rangle + \langle v, w\rangle$

3. $\langle \alpha u, v\rangle = (\alpha u)^TAv = \alpha(u^TAv) = \alpha\langle u, v\rangle$

4. $\langle u, u\rangle = u^TAu > 0 \,\forall\, u \neq 0$ för A är positivt definit.

Kapitel 7

Övningar

Ö7.1 (a) $x(t) = c_1e^{5t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2e^{-2t}\begin{bmatrix} -3 \\ 4 \end{bmatrix}$

(b) $x(t) = c_1e^{\mathrm{i}t}\begin{bmatrix} 5 \\ 3 - \mathrm{i} \end{bmatrix} + c_2e^{-\mathrm{i}t}\begin{bmatrix} 5 \\ 3 + \mathrm{i} \end{bmatrix}$

(c) $x(t) = c_1\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + c_2e^{t}\begin{bmatrix} 2 \\ 3 \\ 22 \end{bmatrix} + c_3e^{-t}\begin{bmatrix} 0 \\ 1 \\ -4 \end{bmatrix}$

(d) $x(t) = c_1\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} + c_2e^{9t}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_3e^{9t}\begin{bmatrix} -\frac{1}{2} \\ 0 \\ 1 \end{bmatrix}$

Problem

P7.1 Med angivna data blir $n = 10.0206$ och $\tau = 2.3285 \times 10^{25}$ s.

Sakregister

■