1 Linjära rum

1.1 Definition och exempel

I matematiken möter man i olika sammanhang storheter som på ett naturligt sätt kan adderas och multipliceras med tal så att vissa enkla räknelagar gäller. Några exempel är geometriska vektorer, vektorer i $\mathbb{R}^n$, matriser, funktioner. Trots de yttre olikheterna har dessa och andra exempel någonting gemensamt, och för att fånga detta har man infört det generellare begreppet linjärt rum.

Definition 1.1. Ett *linjärt rum* är en icke-tom mängd V, vars element kan adderas parvis och multipliceras med tal. Den exakta innebörden av detta är följande:

(1) För alla $u \in V$ och $v \in V$ finns ett väldefinierat element $u + v \in V$.

(2) För alla $u \in V$ och $\alpha \in \mathbb{R}$ finns ett väldefinierat element $\alpha u \in V$.

Dessa två operationer, addition och multiplikation med tal (eller *skalär*) skall uppfylla följande räknelagar:

(3) $u + v = v + u$ för alla $u, v \in V$. (Kommutativ lag)

(4) $(u + v) + w = u + (v + w)$ för alla $u, v, w \in V$. (Associativ lag)

(5) Det finns ett element $\mathbf{0} \in V$ (*nollelementet*) så att $\mathbf{0} + u = u + \mathbf{0} = u$ för alla $u \in V$.

(6) För varje $u \in V$ finns ett element $-u \in V$ sådant att $u + (-u) = (-u) + u = \mathbf{0}$.

(7) $\alpha(\beta u) = (\alpha\beta)u$ för alla $\alpha, \beta \in \mathbb{R}$, $u \in V$. (Associativ lag)

(8) $\alpha(u + v) = \alpha u + \alpha v$ för alla $\alpha \in \mathbb{R}$, $u, v \in V$. (Distributiv lag)

(9) $(\alpha + \beta)u = \alpha u + \beta u$ för alla $\alpha, \beta \in \mathbb{R}$, $u \in V$. (Distributiv lag)

(10) $1u = u$ för alla $u \in V$.

Linjära rum kallas också *vektorrum*. Elementen i ett linjärt rum kallas *vektorer*.

Skalärerna kan också vara komplexa tal. Om egenskaperna (axiomen) (1) – (10) är uppfyllda med $\mathbb{C}$ i stället för $\mathbb{R}$ (där $\mathbb{C}$ betecknar mängden av alla komplexa tal), har man ett *komplext linjärt rum* (vektorrum). I de tre första kapitlen kommer vi att förutsätta reella skalärer, men bl.a. alla resultat i kapitel 1 gäller även för komplexa linjära rum. All matris- och determinantkalkyl gäller också för matriser med komplexa element.

Exempel 1.1. Några exempel på linjära rum.

(a) De geometriska vektorerna i det åskådliga rummet.

(b) $\mathbb{R}^n$, dvs. mängden av alla n-tipler

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)$$

av reella tal, där man definierar

$$\begin{aligned} \mathbf{x} + \mathbf{y} &= (x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = \\ &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n), \\ \alpha\mathbf{x} &= \alpha(x_1, x_2, \ldots, x_n) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n). \end{aligned}$$

Elementen i $\mathbb{R}^n$ kan representeras som *kolonnvektorer*, dvs. $n \times 1$-matriser,

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

eller som *radvektorer*, dvs. $1 \times n$-matriser,

$$(x_1 \; x_2 \; \ldots \; x_n).$$

I regel identifierar vi $\mathbb{R}^n$ med mängden av alla kolonnvektorer. För tydlighetens skull skriver vi ofta radvektorer som

$$(x_1, x_2, \ldots, x_n);$$

det torde framgå av sammanhanget om en n-tipel eller en radvektor avses.

(c) Mängden $\mathbb{R}^{m \times n}$ av alla $m \times n$-matriser med reella element.

(d) Mängen $F(I)$ av alla funktioner på ett intervall I. Addition och multiplikation med skalär definieras på det naturliga sättet, dvs.

$$(f + g)(x) = f(x) + g(x), \quad (\alpha f)(x) = \alpha f(x) \quad \text{för alla } x \in I.$$

Nollelementet är nollfunktionen, $0(x) = 0$ för alla $x \in I$. Axiomen (3) – (10) blir uppfyllda för $F(I)$ på grund av att motsvarande räknelagar gäller för reella (eller komplexa) tal.

(e) I (b) kan alla tal få vara komplexa, så att $\mathbb{C}^n$ (dvs. mängden av alla n-tipler av komplexa tal) är ett komplext linjärt rum. Om vi emellertid begränsar oss till reella skalärer, så ser vi att $\mathbb{C}^n$ också är ett reellt linjärt rum (dvs. ett linjärt rum med reella skalärer). ■

De två första exemplen kan sägas ha stått modell för det allmänna begreppet linjärt rum. Därför används ett geometriskt språkbruk, och man kan hela tiden ha en geometrisk bild i tankarna.

I de konkreta exemplen i Exempel 1.1 är egenskaperna (1) – (10) mer eller mindre självklara. Dessutom har man i dessa fall ytterligare egenskaper som kan förefalla lika självklara. En del av dessa kan visas vara konsekvenser av egenskaperna (1) – (10) och gäller alltså för godtyckliga linjära rum, medan andra egenskaper kan vara specifika för det speciella exemplet och sakna motsvarighet i det allmänna fallet. Som en illustration skall vi härleda några konsekvenser av (1) – (10).

Vi börjar med att konstatera att nollelementet (nollvektorn) $\mathbf{0}$ i (5) är entydigt bestämd. (Observera att en formulering som ”det finns ett element ... så att ... ” betyder ”det finns minst ett element ... så att ... ”.) Om nämligen $\mathbf{0}^*$ uppfyller $\mathbf{0}^* + v = v + \mathbf{0}^* = v$ för alla v, får vi med $v = \mathbf{0}$ och med $u = \mathbf{0}^*$ i (5) att $\mathbf{0}^* + \mathbf{0} = \mathbf{0}$, resp. $\mathbf{0}^* + \mathbf{0} = \mathbf{0}^*$, dvs. $\mathbf{0}^* = \mathbf{0}$. Låt oss titta på en ekvation $w + u = v$ (u och v givna). Den har den entydiga lösningen $w = v + (-u)$. Att $v + (-u)$ är en lösning följer av att

$$(v + (-u)) + u \overset{(4)}{=} v + ((-u) + u) \overset{(6)}{=} v + \mathbf{0} \overset{(5)}{=} v.$$

Å andra sidan är $v + (-u)$ den enda tänkbara lösningen, ty

$$\begin{aligned}
&w + u = v \;\Rightarrow\; (w + u) + (-u) = v + (-u) \overset{(4)}{\Rightarrow} \\
&w + (u + (-u)) = v + (-u) \overset{(6)}{\Rightarrow} w + \mathbf{0} = v + (-u) \overset{(5)}{\Rightarrow} \\
&w = v + (-u).
\end{aligned}$$

För enkelhets skull skriver vi $v - u$ i stället för $v + (-u)$. Vi kan då formulera ovanstående som

$$w + u = v \;\Leftrightarrow\; w = v - u. \tag{1.1}$$

Om speciellt $v = \mathbf{0}$ har vi att $w + u = \mathbf{0} \;\Rightarrow\; w = \mathbf{0} + (-u) = -u$, vilket visar att $-u$ i (6) är entydigt bestämt. Vidare är

$$0u + u \overset{(10)}{=} 0u + 1u \overset{(9)}{=} (0 + 1)u = 1u = u,$$

varför enligt (1.1), eftersom $u - u = \mathbf{0}$,

$$0u = \mathbf{0}. \tag{1.2}$$

Observera att nollan i vänsterledet i (1.2) är det reella (eller komplexa) talet 0, medan nollan i högerledet är nollvektorn i V. Vi noterar också att

$$-u = (-1)u, \tag{1.3}$$

ty $u + (-1)u = 1u + (-1)u = (1 + (-1))u = 0u = \mathbf{0}$, vilket ger (1.3).

1.2 Underrum

Definition 1.2. Ett *underrum* M av ett linjärt rum V är en delmängd M av V sådan att M självt är ett linjärt rum m.a.p. samma operationer.

Eftersom räknelagarna (3) – (10) i Definition 1.1 gäller i hela V gäller de också i en delmängd M under förutsättning att $\mathbf{0} \in M$. Därför behöver man bara verifiera (1) och (2) för M för att inse att M är ett underrum av V, dvs. man skall verifiera att M är slutet under addition och multiplikation med tal. Vi formulerar detta som en sats.

Sats 1.1. *En icke-tom delmängd M av ett linjärt rum V är ett underrum av V om och endast om M har egenskaperna*

$$\begin{aligned} u, v \in M \quad &\Rightarrow \quad u + v \in M, \\ u \in M,\ \alpha \in \mathbb{R} \quad &\Rightarrow \quad \alpha u \in M, \end{aligned} \tag{1.4}$$

eller, ekvivalent, egenskapen

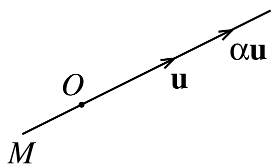
$$u, v \in M,\ \alpha, \beta \in \mathbb{R} \ \Rightarrow\ \alpha u + \beta v \in M.$$

Anmärkning. Tag ett $u_1 \in M$ (M är icke-tom) och $\alpha = 0$ i (1.4) så fås att $0u_1 = \mathbf{0} \in M$. Därmed gäller (5) för M, om (1.4) gäller.

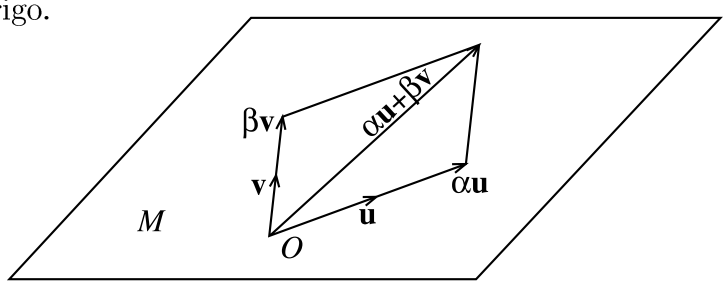
Exempel 1.2. I det åskådliga rummet, där vi som vanligt identifierar punkter och ortsvektorer, är följande delmängder underrum:

(a) Origo, eller noggrannare, den mängd vars enda element är 0.

(b) Räta linjer genom origo.



(c) Plan genom origo.



Även i allmännare situationer kan man ha denna figur i tankarna och föreställa sig ett underrum som någonting som är analogt med ett plan genom origo.

(d) Hela rummet.

När vi i nästa avsnitt har infört dimensionsbegreppet, kommer det att framgå att dessa underrum har dimensionerna 0, 1, 2 och 3, respektive. ■

Exempel 1.3. Låt V vara det linjära rummet av alla reella $n \times n$-matriser (jfr Exempel 1.1 (c)), och låt M vara den delmängd som består av alla symmetriska $n \times n$-matriser. Då är M ett underrum av V, ty om A och B är symmetriska, så är också $A + B$ symmetrisk ($(A + B)^t = A^t + B^t = A + B$), och αA är symmetrisk om α är ett reellt tal ($(\alpha A)^t = \alpha A^t = \alpha A$). ■

Exempel 1.4. I tillämpningarna möter man ofta s.k. *funktionsrum*, dvs. olika underrum av $F(I)$ (se Exempel 1.1 (d)). Några exempel:

$C[a, b]$ = mängden av alla funktioner som är kontinuerliga på $[a, b]$.

$C^k(a, b)$ = mängden av alla funktioner som är k gånger kontinuerligt deriverbara i (a, b).

$C^\infty(a, b)$ = mängden av alla funktioner som har kontinuerliga derivator av alla ordningar i (a, b).

$\mathcal{P}$ = mängden av alla polynom.

$\mathcal{P}_n$ = mängden av alla polynom av grad högst n.

$L^2(a, b)$ = mängden av alla reellvärda funktioner f sådana att f^2 är integrerbar på (a, b). (Bokstaven L kommer från namnet Lebesgue och s.k. Lebesgueintegrerbara funktioner.)

För att visa att en mängd av funktioner är ett linjärt rum räcker det att verifiera att $f + g$ och αf tillhör mängden så snart f och g gör det. Mängden $C[a, b]$ blir ett linjärt rum, därför att $f + g$ och αf är kontinuerliga, om f och g är kontinuerliga. För C^k och C^∞ används också att $f + g$ och αf är deriverbara så snart f och g är det. För att visa att $L^2(a, b)$ är slutet under addition använder vi den enkla men nyttiga olikheten $\alpha\beta \le \frac{1}{2}(\alpha^2 + \beta^2)$ för godtyckliga reella tal α och β. Denna följer av att $\alpha^2 + \beta^2 - 2\alpha\beta = (\alpha - \beta)^2 \ge 0$. Vi får också $(\alpha + \beta)^2 = \alpha^2 + \beta^2 + 2\alpha\beta \le 2(\alpha^2 + \beta^2)$. Om f och g tillhör $L^2(a, b)$, så är $[f(x) + g(x)]^2 \le 2f^2(x) + 2g^2(x)$, vilket medför att $f + g$ tillhör $L^2(a, b)$. Att αf tillhör $L^2(a, b)$, om f gör det, är klart. ■

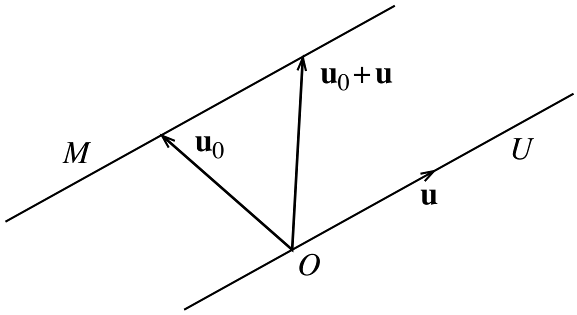
Exempel 1.5. Låt V vara det linjära rummet $C[0, 1]$ (se ovan). Sätt

$$M = \{f \in V : f(0) = f(1) = 0\}.$$

Då är M ett underrum av V, ty om f och g tillhör M och α är ett tal, så är $f + g$ kontinuerlig och satisfierar

$$(f + g)(0) = f(0) + g(0) = 0 + 0 = 0, \quad (f + g)(1) = f(1) + g(1) = 0,$$

Adjektivet ”linjär” i termen ”linjärt rum” är ju bildat från ordet ”linje”, och därför kan det kanske tyckas konstigt att (räta) linjer (och likaså plan) i allmänhet inte är linjära rum utan bara sådana som går genom origo. I det allmännare fallet använder man ofta adjektivet ”affin”. Eftersom en godtycklig rät linje är en translation (parallellförflyttning) av en linje genom origo (se fig.), så gör vi följande definition:



Definition 1.3. En delmängd M av V kallas *affin*, om det finns en vektor $u_0 \in V$ och ett underrum U av V så att

$$M = u_0 + U = \{u_0 + u : u \in U\}.$$

1.3 Några grundbegrepp

I detta avsnitt skall vi införa några centrala begrepp som linjärt beroende och oberoende, bas och dimension. Dessa begrepp har man redan stött på i någon form i samband med geometriska vektorer. Så är t.ex. tre (geometriska) vektorer linjärt oberoende om de inte ligger i ett och samma plan. En sådan uppsättning vektorer kan användas som en bas i rummet, och detta leder som bekant till begrepp som komponenter (av en vektor) och koordinater (för en punkt). Motsvarande kan göras i ett godtyckligt vektorrum (av ändlig dimension). Dimensionen är det antal vektorer som ingår i en bas. Ett centralt resultat i detta avsnitt är att varje bas innehåller lika många vektorer (detta är inte trivialt). För att komma fram till det behöver vi en del grundläggande satser. Bokstaven V betecknar genomgående ett reellt linjärt rum (men allting gäller också för komplexa linjära rum).

Definition 1.4. Låt $u_1, \ldots, u_n$ vara vektorer i V. En *linjärkombination* av $u_1, \ldots, u_n$ är en vektor av formen

$$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n$$

för vissa skalärer $\lambda_1, \ldots, \lambda_n$. Mängden av alla linjärkombinationer av $u_1, \ldots, u_n$ utgör ett underrum av V (visa detta som övning!); detta underrum betecknas $\mathrm{Span}\{u_1, \ldots, u_n\}$.

Definition 1.5. En uppsättning vektorer $u_1, \ldots, u_n$ i V sägs *generera* (eller *spänna*) V om varje element i V är en linjärkombination av $u_1, \ldots, u_n$, dvs. $V = \text{Span}\{u_1, \ldots, u_n\}$.

Exempel 1.8. $\mathbb{R}^n$ genereras av vektorerna

$$\mathbf{e}_1 = (1, 0, \ldots, 0),\ \mathbf{e}_2 = (0, 1, \ldots, 0), \ldots,\ \mathbf{e}_n = (0, 0, \ldots, 1),$$

ty ett godtyckligt $\mathbf{x} \in \mathbb{R}^n$ kan skrivas

$$\mathbf{x} = (x_1, x_2, \ldots, x_n) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n. \qquad \blacksquare$$

Definition 1.6. En uppsättning vektorer $u_1, \ldots, u_n$ i V sägs vara *linjärt beroende* om nollvektorn kan skrivas som en icke-trivial linjärkombination av $u_1, \ldots, u_n$, dvs. om ett samband av formen

$$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n = \mathbf{0} \tag{1.5}$$

är möjligt med något $\lambda_i \neq 0$. Om å andra sidan relationen (1.5) endast är möjlig då $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$, så sägs $u_1, \ldots, u_n$ vara *linjärt oberoende*.

Exempel 1.9. Vektorerna $\mathbf{e}_1, \ldots, \mathbf{e}_n$ i $\mathbb{R}^n$ (se Exempel 1.8) är linjärt oberoende, ty om

$$\mathbf{0} = \lambda_1\mathbf{e}_1 + \cdots + \lambda_n\mathbf{e}_n = (\lambda_1, \ldots, \lambda_n),$$

så är alla $\lambda_i = 0$. $\blacksquare$

Lemma 1.1. *Vektorerna $u_1, \ldots, u_n$ är linjärt beroende om och endast om någon av dem kan skrivas som en linjärkombination av de övriga.*

Bevis. Antag att $u_1, \ldots, u_n$ är linjärt beroende så att (1.5) gäller med någon koefficient, säg λ_i, skild från noll. Då kan vi dividera med λ_i och lösa ut u_i som en linjärkombination av de övriga vektorerna:

$$u_i = -\frac{\lambda_1}{\lambda_i}u_1 - \cdots - \frac{\lambda_{i-1}}{\lambda_i}u_{i-1} - \frac{\lambda_{i+1}}{\lambda_i}u_{i+1} - \cdots - \frac{\lambda_n}{\lambda_i}u_n.$$

Omvänt, om någon av vektorerna, säg u_i, är en linjärkombination av de övriga, så är

$$u_i = \alpha_1 u_1 + \cdots + \alpha_{i-1}u_{i-1} + \alpha_{i+1}u_{i+1} + \cdots + \alpha_n u_n$$

för vissa tal $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n$. Överflyttning av u_i till högerledet ger ett samband av formen (1.5) där åtminstone någon koefficient är skild från noll (nämligen $\lambda_i = -1$). $\blacksquare$

Lemma 1.2. *Antag att $u_1, \ldots, u_m$ är linjärt oberoende vektorer i V. Om $v \in V$ inte tillhör underrummet $U = \text{Span}\{u_1, \ldots, u_m\}$, så är $u_1, \ldots, u_m, v$ linjärt oberoende.*

Bevis. Antag
$$\alpha_1 u_1 + \cdots + \alpha_m u_m + \beta v = \mathbf{0}. \tag{1.6}$$
Om $\beta \neq 0$, kan v skrivas som en linjärkombination av $u_1, \ldots, u_m$, vilket strider mot att $v \notin U$. Alltså är $\beta = 0$, och
$$\alpha_1 u_1 + \cdots + \alpha_m u_m = \mathbf{0}.$$
Då $u_1, \ldots, u_m$ är linjärt oberoende, så måste vi ha $\alpha_1 = \cdots = \alpha_m = 0$. Då är alla koefficienter i (1.6) noll, och $u_1, \ldots, u_m, v$ är linjärt oberoende. ■

Definition 1.7. En uppsättning vektorer $u_1, \ldots, u_n$ i V sägs vara en *bas* för V om vektorerna är linjärt oberoende och genererar V.

Definition 1.8. Ett linjärt rum V sägs ha *dimensionen* n, skrivet $\dim V = n$, om n är det maximala antalet linjärt oberoende vektorer i V. Om inget sådant n finns, dvs. om man kan finna delmängder av V bestående av godtyckligt många linjärt oberoende vektorer, sägs V vara *oändligdimensionellt.*

Anmärkning. Det triviala rummet $\{\mathbf{0}\}$ innehåller inga linjärt oberoende vektorer, varför dimensionen är 0.

Sats 1.2. *Antag att V har dimensionen $n > 0$. Då finns minst en uppsättning av n linjärt oberoende vektorer ur V. Varje sådan uppsättning är en bas för V.*

Bevis. Låt $u_1, \ldots, u_n$ vara en godtycklig uppsättning av n linjärt oberoende vektorer ur V. Existensen av minst en sådan uppsättning följer från Definition 1.8. Låt $v \in V$ vara godtyckligt. Om v inte tillhör $\mathrm{Span}\{u_1, \ldots, u_n\}$, så ger Lemma 1.2 att $u_1, \ldots, u_n, v$ är linjärt oberoende. Men detta strider mot maximaliteten av n i Definition 1.8. Alltså är v en linjärkombination av $u_1, \ldots, u_n$. Då $v \in V$ är godtyckligt, visar detta att $u_1, \ldots, u_n$ genererar V och alltså är en bas för V. ■

Vi har nu visat att varje linjärt rum av ändlig dimension har någon bas. Vi önskar bevisa att alla baser för ett rum V har lika många element. Då behöver vi följande resultat, som är centralt i samband med begreppen bas och dimension.

Sats 1.3. *Antag att V genereras av vektorerna $u_1, \ldots, u_n$. Då är varje uppsättning av fler än n vektorer i V linjärt beroende.*

Bevis. Låt $v_1, \ldots, v_p$ vara p vektorer i V, där $p > n$. Eftersom $u_1, \ldots, u_n$ genererar V kan man skriva
$$v_i = a_{1i} u_1 + \cdots + a_{ni} u_n = \sum_{j=1}^{n} a_{ji} u_j, \quad i = 1, \ldots, p,$$
för vissa koefficienter a_{ji}. Betrakta ekvationen
$$x_1 v_1 + \cdots + x_p v_p = \mathbf{0}, \tag{1.7}$$

som kan skrivas

$$\sum_{i=1}^{p} x_i\Big(\sum_{j=1}^{n} a_{ji}u_j\Big) = \mathbf{0},$$

eller

$$\sum_{j=1}^{n}\Big(\sum_{i=1}^{p} a_{ji}x_i\Big)u_j = \mathbf{0}. \tag{1.8}$$

Ekvationen (1.8), och därmed (1.7), är satisfierad för varje val av $x_1, \ldots, x_p$ sådant att

$$\sum_{i=1}^{p} a_{ji}x_i = 0 \quad \text{för } j = 1, \ldots, n. \tag{1.9}$$

Men (1.9) är ett homogent ekvationssystem med n ekvationer och p obekanta, där $p > n$. Enligt Sats 0.1 har (1.9) en icke-trivial lösning, dvs. det finns tal $x_1, \ldots, x_p$, ej alla noll, så att (1.7) satisfieras. Detta betyder att $v_1, \ldots, v_p$ är linjärt beroende. ■

Sats 1.4. *Antag att V är ändligdimensionellt. Då har alla baser för V lika många element, och detta antal är lika med dimensionen för V.*

Bevis. Klart om $\dim V = 0$. Om $\dim V = n > 0$, så finns enligt Sats 1.2 någon bas med n element. Betrakta så en annan bas (vilken som helst) och antag att den består av m vektorer. Eftersom n är det största antal linjärt oberoende vektorer, som kan förekomma i V, och eftersom vi har m linjärt oberoende vektorer i den andra basen, så måste $m \leq n$. Dessa m vektorer genererar också V, och skulle vi ha $n > m$, så skulle Sats 1.3 ge att de n vektorerna i den första basen vore linjärt beroende, vilket är orimligt. Alltså är $m = n$, dvs. alla baser för V har n element. ■

Exempel 1.10. (a) $\mathbb{R}^n$ har dimensionen n, ty enligt Exemplen 1.8 och 1.9 har $\mathbb{R}^n$ en bas bestående av n element.

(b) $\mathbb{C}^n$ med komplexa skalärer har på samma sätt dimensionen n, ty $\mathbf{e}_1, \ldots, \mathbf{e}_n$ är en bas också för $\mathbb{C}^n$. Det blir dock annorlunda om vi betraktar reella skalärer. Det komplexa talet $x + iy$ kan identifieras med punkten (x, y) i $\mathbb{R}^2$ (det komplexa planet), varför $\dim \mathbb{C} = 2$ och $\dim \mathbb{C}^n = 2n$ med reella skalärer. I det senare fallet är vektorerna $\mathbf{e}_1, i\mathbf{e}_1, \ldots, \mathbf{e}_n, i\mathbf{e}_n$ en bas.

(c) Rummet $C[a, b]$ är oändligdimensionellt, ty vi kan finna godtyckligt många linjärt oberoende funktioner i rummet. Exempelvis är funktionerna $1, x, \ldots, x^n$ linjärt oberoende för varje n, ty ett polynom som inte är nollpolynomet har bara ändligt många nollställen. ■

Vi avslutar detta avsnitt med några nyttiga satser rörande bas och dimension. En uppsättning vektorer, som genererar V, kan tunnas ut till en bas för V, och en uppsättning linjärt oberoende vektorer kan utvidgas till en bas för V. Detta är innebörden i det två följande satserna.

Definition 2.1. Låt V vara ett reellt linjärt rum. En *skalärprodukt* i V är en reellvärd funktion $\langle u, v \rangle$ av två variabler u och v i V med följande egenskaper:

(i) $\langle u, v \rangle = \langle v, u \rangle$ för alla $u, v \in V$,
(ii) $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$ för alla $u, v \in V, \alpha \in \mathbb{R}$,
(iii) $\langle u_1 + u_2, v \rangle = \langle u_1, v \rangle + \langle u_2, v \rangle$ för alla $u_1, u_2, v \in V$,
(iv) $\langle u, u \rangle \geq 0$ för alla $u \in V$, och $\langle u, u \rangle = 0$ endast för $u = \mathbf{0}$.

Exempel 2.1. Några exempel på skalärprodukt.

(a)
$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n$$

är en skalärprodukt i $\mathbb{R}^n$ enligt ovan. Den kallas *standardskalärprodukten.*

(b)
$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + 2 x_2 y_2 + \cdots + n x_n y_n$$

är också en skalärprodukt i $\mathbb{R}^n$.

(c)
$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_1 y_2 + x_2 y_1 + 2 x_2 y_2$$

är en skalärprodukt i $\mathbb{R}^2$. Det är endast egenskap (iv) som inte är uppenbar. Vi har

$$\langle \mathbf{x}, \mathbf{x} \rangle = x_1^2 + 2 x_1 x_2 + 2 x_2^2 = (x_1 + x_2)^2 + x_2^2 \geq 0,$$

och likhet inträffar om och endast om $x_1 + x_2 = 0$ och $x_2 = 0$, dvs. $x_1 = x_2 = 0$.

Dessa exempel är olika specialfall av följande.

(d) Låt A vara en *symmetrisk* $n \times n$-matris. Antag också att A är *positivt definit* (se avsnitt 4.5), dvs. att $\mathbf{x}^t A \mathbf{x} \geq 0$ för alla $\mathbf{x} \in \mathbb{R}^n$, och $\mathbf{x}^t A \mathbf{x} = 0$ endast för $\mathbf{x} = \mathbf{0}$. Då definierar $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^t A \mathbf{y}$ en skalärprodukt i $\mathbb{R}^n$. Låt oss verifiera egenskaperna (i) – (iv) i Definition 2.1:

(i) $\langle \mathbf{y}, \mathbf{x} \rangle = \mathbf{y}^t A \mathbf{x} = (\mathbf{y}^t A \mathbf{x})^t = \mathbf{x}^t A^t \mathbf{y} = \mathbf{x}^t A \mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle$, eftersom A är symmetrisk.
(ii) $\langle \alpha \mathbf{x}, \mathbf{y} \rangle = (\alpha \mathbf{x})^t A \mathbf{y} = \alpha \mathbf{x}^t A \mathbf{y} = \alpha \langle \mathbf{x}, \mathbf{y} \rangle$.
(iii) $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = (\mathbf{x} + \mathbf{y})^t A \mathbf{z} = (\mathbf{x}^t + \mathbf{y}^t) A \mathbf{z} = \mathbf{x}^t A \mathbf{z} + \mathbf{y}^t A \mathbf{z} = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$.
(iv) Att $\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^t A \mathbf{x} \geq 0$ för alla $\mathbf{x}$ med likhet endast för $\mathbf{x} = \mathbf{0}$ är precis definitionen av att A är positivt definit.

(e) Låt V vara det linjära rummet $C[a, b]$ av alla reella kontinuerliga funktioner på $[a, b]$. Då är

$$\langle f, g \rangle = \int_a^b f(x) g(x)\, dx$$

en skalärprodukt i V. ■

Låt oss notera några konsekvenser av Definition 2.1. Genom upprepad användning av (ii) och (iii) finner vi att

$$\begin{aligned} \langle \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k, v \rangle &= \langle \alpha_1 u_1, v \rangle + \langle \alpha_2 u_2, v \rangle + \cdots + \langle \alpha_k u_k, v \rangle = \\ &= \alpha_1 \langle u_1, v \rangle + \alpha_2 \langle u_2, v \rangle + \cdots + \alpha_k \langle u_k, v \rangle. \end{aligned} \tag{2.2}$$

Med användning av (i) får vi också

$$\langle u, \beta v\rangle = \beta\langle u, v\rangle, \tag{2.3}$$

$$\langle u, \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_m v_m\rangle = \beta_1\langle u, v_1\rangle + \beta_2\langle u, v_2\rangle + \cdots + \beta_m\langle u, v_m\rangle, \tag{2.4}$$

och genom en kombination av (2.2) och (2.4)

$$\begin{aligned}\Big\langle \sum_{i=1}^{k} \alpha_i u_i, \sum_{j=1}^{m} \beta_j v_j \Big\rangle &= \sum_{i=1}^{k} \alpha_i \Big\langle u_i, \sum_{j=1}^{m} \beta_j v_j \Big\rangle = \sum_{i=1}^{k} \alpha_i \sum_{j=1}^{m} \beta_j \langle u_i, v_j\rangle = \\ &= \sum_{i=1}^{k}\sum_{j=1}^{m} \alpha_i\beta_j\langle u_i, v_j\rangle.\end{aligned} \tag{2.5}$$

Ur (i) och (ii) med $\alpha = 0$ fås

$$\langle \mathbf{0}, v\rangle = \langle v, \mathbf{0}\rangle = 0 \ \text{ för alla } v \in V. \tag{2.6}$$

I fortsättningen av kapitel 2 betecknar V ett reellt linjärt rum försett med en skalärprodukt. I $\mathbb{R}^n$ använder vi alltid standardskalärprodukten om inget annat sägs.

I analogi med vad som gäller för geometriska vektorer kan vi införa begreppen längd och vinkel med hjälp av skalärprodukt.

Definition 2.2. Med *längden* eller *normen* av vektorn $u \in V$ menas

$$||u|| = \sqrt{\langle u, u\rangle}. \tag{2.7}$$

Observera att detta är väldefinierat p.g.a. (iv). Med *avståndet* mellan u och v menas $||u-v||$.

Ur (2.7) och (2.5) får vi

$$||u+v||^2 = \langle u+v, u+v\rangle = \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle,$$

dvs.

$$||u+v||^2 = ||u||^2 + 2\langle u, v\rangle + ||v||^2. \tag{2.8}$$

Vidare följer av (2.7), (ii) och (2.3) att

$$||\alpha v|| = |\alpha|||v||. \tag{2.9}$$

Definitionen av vinkel hänger på följande viktiga olikhet.

Sats 2.1 (Cauchy-Schwarz' olikhet).

$$|\langle u, v\rangle| \le ||u||||v|| \quad \textit{för alla } u, v \in V. \tag{2.10}$$

Likhet gäller om och endast om u och v är linjärt beroende.

2.2 Ortogonalitet. Ortogonalprojektion

Fallet $\theta = \frac{\pi}{2}$ i formel (2.12) är av särskilt intresse.

Definition 2.4.
(a) Två vektorer u och v i V är *ortogonala* om $\langle u, v\rangle = 0$. En mängd av vektorer kallas en *ortogonalmängd*, om vektorerna i mängden är parvis ortogonala.
(b) En vektor $u \in V$ kallas *normerad* om $\|u\| = 1$.
(c) En mängd av vektorer kallas en *ortonormerad* mängd, eller en *ON-mängd*, om den är en ortogonalmängd, vars samtliga vektorer är normerade.
(d) En ortogonalmängd, som är en bas för ett linjärt rum, kallas en *ortogonalbas*. Om basen är en ON-mängd, kallas den en *ON-bas*.

Lemma 2.1. *En ON-mängd $e_1, \ldots, e_n$ är linjärt oberoende. Om $e_1, \ldots, e_n$ är en ON-bas för V, så gäller att varje vektor $u \in V$ kan skrivas*

$$u = \langle u, e_1\rangle e_1 + \langle u, e_2\rangle e_2 + \cdots + \langle u, e_n\rangle e_n. \tag{2.14}$$

Bevis. Antag att en vektor u är en linjärkombination av $e_1, \ldots, e_n$, så att

$$u = \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n \tag{2.15}$$

för vissa tal $\lambda_1, \ldots, \lambda_n$. Multiplicera (2.15) skalärt med $e_j, 1 \leq j \leq n$, så fås

$$\langle u, e_j\rangle = \sum_{i=1}^{n} \lambda_i \langle e_i, e_j\rangle = \lambda_j.$$

Om vi väljer $u = \mathbf{0}$ i (2.15), så får vi $\lambda_j = \langle \mathbf{0}, e_j\rangle = 0$ för alla j, dvs. $e_1, \ldots, e_n$ är linjärt oberoende. Om $e_1, \ldots, e_n$ är en ON-bas för V, så kan varje $u \in V$ skrivas på formen (2.15), och vi får $\lambda_j = \langle u, e_j\rangle$ för alla j, dvs. vi har visat (2.14). ■

Sats 2.3. *Låt V vara ändligdimensionellt. Då har V en ON-bas.*

Bevis. Beviset består av en beskrivning av en metod, *Gram-Schmidts ortogonaliseringsmetod*, som också är användbar för konkreta räkningar.

Antag $\dim V = n$ och antag att $v_1, \ldots, v_n$ är en bas för V. Sätt $e_1' = v_1$ och ansätt

$$e_2' = \alpha_{21} e_1' + v_2. \tag{2.16}$$

Bestäm α_{21} så att e_2' och e_1' blir ortogonala. Multiplicera alltså (2.16) skalärt med e_1':

$$0 = \langle e_2', e_1'\rangle = \alpha_{21}\langle e_1', e_1'\rangle + \langle v_2, e_1'\rangle = \alpha_{21}\|e_1'\|^2 + \langle v_2, e_1'\rangle.$$

Då $\|e_1'\| = \|v_1\| \neq 0$, måste

$$\alpha_{21} = -\frac{\langle v_2, e_1'\rangle}{\|e_1'\|^2}.$$

Av (2.16) framgår att $e_2' \neq 0$ (annars vore v_1 och v_2 linjärt beroende) och även att e_1', e_2' genererar samma underrum som v_1, v_2. Vi fortsätter på samma sätt. Antag att $e_1', \dots, e_k'$ $(1 \leq k \leq n-1)$ har konstruerats så att $e_1', \dots, e_k'$ är parvis ortogonala och så att $e_1', \dots, e_k'$ genererar samma underrum som $v_1, \dots, v_k$. Ansätt

$$e_{k+1}' = \alpha_{k+1,1} e_1' + \cdots + \alpha_{k+1,k} e_k' + v_{k+1}. \tag{2.17}$$

Talen $\alpha_{k+1,1}, \dots, \alpha_{k+1,k}$ skall väljas så att e_{k+1}' blir ortogonal mot $e_1', \dots, e_k'$. Multiplicera (2.17) skalärt med e_j' för $j = 1, \dots, k$, så fås, eftersom $\langle e_i', e_j' \rangle = 0$ för $i \neq j$,

$$\begin{aligned} 0 = \langle e_{k+1}', e_j' \rangle &= \alpha_{k+1,j} \|e_j'\|^2 + \langle v_{k+1}, e_j' \rangle, \\ \alpha_{k+1,j} &= -\frac{\langle v_{k+1}, e_j' \rangle}{\|e_j'\|^2}, \quad j = 1, \dots, k. \end{aligned}$$

Formel (2.17) tillsammans med förutsättningarna om $e_1', \dots, e_k'$ visar att $e_1', \dots, e_{k+1}'$ genererar samma underrum som $v_1, \dots, v_{k+1}$ (speciellt är $e_{k+1}' \neq \mathbf{0}$). Processen fortsättes tills $e_1', \dots, e_n'$ har konstruerats. Sammanfattningsvis får vi alltså

$$\begin{aligned} e_1' &= v_1, \\ e_2' &= v_2 - \frac{\langle v_2, e_1' \rangle}{\|e_1'\|^2} e_1', \\ &- - - - - \\ e_n' &= v_n - \frac{\langle v_n, e_1' \rangle}{\|e_1'\|^2} e_1' - \cdots - \frac{\langle v_n, e_{n-1}' \rangle}{\|e_{n-1}'\|^2} e_{n-1}', \end{aligned}$$

och dessa vektorer utgör en ortogonalbas för V. Sätt slutligen

$$e_i = \frac{e_i'}{\|e_i'\|}, \quad i = 1, \dots, n.$$

Då blir $e_1, \dots, e_n$ en ON-bas för V. ■

Sats 2.4 ("Pythagoras' sats"). (a) *Om u och v är ortogonala, så är*

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

(b) *Om $\{u_1, u_2, \dots, u_n\}$ är en ortogonalmängd i V, så är*

$$\|u_1 + u_2 + \cdots + u_n\|^2 = \|u_1\|^2 + \|u_2\|^2 + \cdots + \|u_n\|^2.$$

Bevis. (a) följer ur (2.8), ty $\langle u, v \rangle = 0$.

(b) Enligt formel (2.5) är

$$\|u_1 + u_2 + \cdots + u_n\|^2 = \Big\langle \sum_{i=1}^n u_i, \sum_{j=1}^n u_j \Big\rangle = \sum_{i=1}^n \sum_{j=1}^n \langle u_i, u_j \rangle = \sum_{i=1}^n \|u_i\|^2,$$

ty $\langle u_i, u_j \rangle = 0$ då $i \neq j$. ■

3 Linjära avbildningar

3.1 Definitioner och exempel

Definition 3.1. Låt U och V vara reella linjära rum. En avbildning (funktion) från U till V kallas *linjär*, om

$$F(u+v) = F(u) + F(v) \quad \text{för alla } u, v \in U$$

och

$$F(\alpha u) = \alpha F(u) \quad \text{för alla } \alpha \in \mathbb{R},\ u \in U,$$

eller ekvivalent,

$$F(\alpha u + \beta v) = \alpha F(u) + \beta F(v) \quad \text{för alla } \alpha, \beta \in \mathbb{R},\ u, v \in U. \tag{3.1}$$

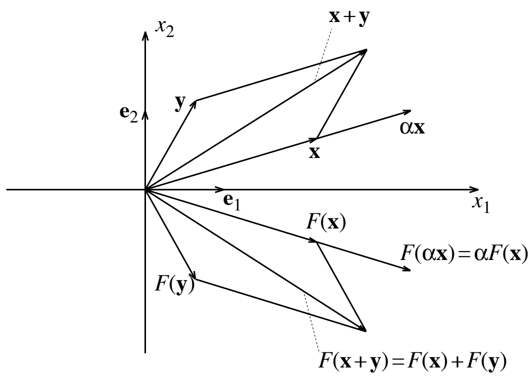
Om $U = V$ säger man att F är en linjär avbildning på V.

Genom upprepad användning av (3.1) (induktion) finner man att

$$F\Big(\sum_{i=1}^{k} \alpha_i u_i\Big) = \sum_{i=1}^{k} \alpha_i F(u_i) \tag{3.2}$$

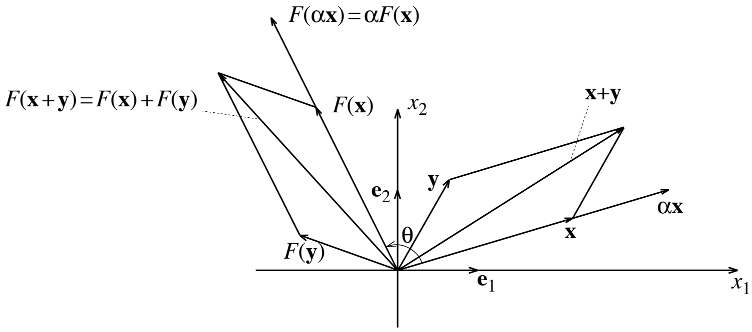
för godtyckliga $\alpha_i \in \mathbb{R}$, $u_i \in U$, $k \geq 1$. Vidare ger $\alpha = 0$ i definitionen att $F(\mathbf{0}) = \mathbf{0}$.

Exempel 3.1. $U = V =$ ett plan med en ON-bas $\mathbf{e}_1, \mathbf{e}_2$, $F(\mathbf{x}) =$ spegling i x_1-axeln.



Linjariteten illustreras i figuren. ■

Exempel 3.2. $U = V =$ ett plan med en ON-bas $\mathbf{e}_1, \mathbf{e}_2$, $F(\mathbf{x}) =$ rotation av vektorn $\mathbf{x}$ vinkeln θ i positiv led. Även här framgår linjariteten av figuren.



■

Exempel 3.3. $U = \mathbb{R}^n$, $V = \mathbb{R}^m$,

$$F(\mathbf{x}) = A\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n,$$

där A är en $m \times n$-matris. Linjariteten följer direkt av reglerna för matrismultiplikation:

$$F(\alpha\mathbf{x} + \beta\mathbf{y}) = A(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha A\mathbf{x} + \beta A\mathbf{y} = \alpha F(\mathbf{x}) + \beta F(\mathbf{y}). \qquad \blacksquare$$

Exempel 3.4. (a) $U = C^1[0,1]$ = mängden av alla kontinuerligt deriverbara funktioner på $[0,1]$, $V = C[0,1]$,

$$F(f) = Df = f'.$$

Här följer linjariteten av räknereglerna för derivata:

$$F(\alpha f + \beta g) = D(\alpha f + \beta g) = \alpha f' + \beta g' = \alpha F(f) + \beta F(g).$$

(b) Låt $P(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$ vara ett polynom. Då är

$$\begin{aligned} F(f) &= P(D)f = (D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0)f \\ &= f^{(n)} + a_{n-1}f^{(n-1)} + \cdots + a_1 f' + a_0 f \end{aligned}$$

en linjär avbildning från $U = C^n(-\infty,\infty)$ = mängden av alla n gånger kontinuerligt deriverbara funktioner på $(-\infty,\infty)$ till $V = C(-\infty,\infty)$. Detta visas som i (a). ■

Exempel 3.5. (a) $U = V = C[a,b]$. För $f \in C[a,b]$ definieras $F_1(f) = g_1$ och $F_2(f) = g_2$, där funktionerna g_1 och g_2 i $C[a,b]$ ges av

$$g_1(x) = \int_a^x f(t)dt, \quad g_2(x) = \int_a^b (x-t)^2 f(t)dt, \quad x \in [a,b].$$

Den enkla verifikationen att F_1 och F_2 är linjära lämnas som övning åt läsaren. En avbildning som (likt dessa och avbildningarna i Exempel 3.4) avbildar en funktion på en funktion brukar kallas en *operator*.

(b) $U = C[a,b]$, $V = \mathbb{R}$, $\quad F(f) = \displaystyle\int_a^b f(x)dx.$

En avbildning som (likt denna) avbildar en funktion på ett tal brukar kallas en *funktional*.