

Sammanfattning läsvecka ett flervariabelanalys

Johan Philis, Pontus Lindgren, Jakob Nilsson,
Matilda Selvaraj Tivesten, Jakob Ruderfors, Gustav Johansson.

6 februari 2021

1 Introduktion

Differentierbarhet är en egenskap hos funktioner i högre dimensioner. Det är lämpligt att läsaren introduceras till differentierbarhet genom att se på det som en förlängning av deriverbarhet eftersom dessa två är nära besläktade.

Låt $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ vara en funktion. Då definieras funktionens derivata i punkten a på följande vis.

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} \quad (1)$$

Om detta gränsvärde existerar är f deriverbar i punkten a . Vidare kan denna definition skrivas om till

$$\rho(h) = \frac{f(a+h) - f(a)}{h} - A \rightarrow 0 \text{ då } h \rightarrow 0 \quad (2)$$

Detta uttryck kan skrivas om med en felterm $\rho(h)$ till:

$$f(a+h) - f(a) = hA + h\rho(h) \text{ då } \lim_{h \rightarrow 0} \rho(h) = 0. \quad (3)$$

Skälet att man skriver om uttrycket på detta sätt är att det blir lättare att gå vidare till definitionen av differentierbarhet.

2 Differentierbarhet och partiella derivator

2.1 Funktioner av 2 variabler

Om man vill utöka till två variabler $f(x, y)$ blir motsvarande uttryck väldigt likt ekvation (3). Om $\exists A_1, A_2 \in \mathbb{R}$ sådana att

$$f(a+h, b+k) - f(a, b) = hA_1 + kA_2 + \sqrt{h^2 + k^2}\rho(h, k) \text{ då } \lim_{(h,k) \rightarrow (0,0)} \rho(h, k) = (0, 0). \quad (4)$$

då säger vi att f är differentierbar i punkten (a, b) .

En funktion sägs vara differentierbar i mängden M om funktionen är differentierbar i alla punkter i M .

Ett annat viktigt begrepp är tangentplan. För grafen $z = f(x, y)$ till en funktion av två variabler som är differentierbar i punkten (a, b) finns ett och endast ett plan som är tangent med funktionsytan i punkten (a, b) . Om planet är tangent i x -led och y -led följer att planet är tangent i alla riktningar. Om funktionsytan $z = f(x, y)$ är differentierbar i punkten (a, b) ser tangentplanet för f i (a, b) ut som följer:

$$z = f(a, b) + A_1(x - a) + A_2(y - b). \quad (5)$$

I flera dimensioner kan lutningen i en punkt skilja sig i olika riktningar, och således finns ingen entydig definition för derivata såsom i envariabelanalys. Därför introduceras partiella derivator, vilka tolkas som lutningen i en funktionspunkt när man närmar sig den parallellt med någon koordinataxel. Låt $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, vara en funktion av (x, y) , där (a, b) är inre punkt i $D(f)$. Då definieras de partiella derivatorna i (a, b) med avseende på x och y enligt

$$\frac{\partial f}{\partial x}(a, b) = \lim_{h \rightarrow 0} \frac{f(a + h, b) - f(a, b)}{h} \quad (6)$$

$$\frac{\partial f}{\partial y}(a, b) = \lim_{h \rightarrow 0} \frac{f(a, b + h) - f(a, b)}{h}, \quad (7)$$

givet att gränsvärdena existerar.

De partiella derivatorna används vid linjärisering på så sätt att om f är differentierbar i (a, b) enligt ekvation (4), är konstanterna A_1 resp. A_2 de partiella derivatorna i x - och y -led.

Sats 1. Om $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ är differentierbar i (a, b) enligt ekvation (4) gäller att $A_1 = \frac{\partial f}{\partial x}(a, b)$ och $A_2 = \frac{\partial f}{\partial y}(a, b)$.

Bevis. Vi dividerar båda leden med h i ekvation (4) och får att

$$\frac{f(a + h, b + k) - f(a, b)}{h} = A_1 + \frac{kA_2}{h} + \frac{\sqrt{h^2 + k^2}\rho(h, k)}{h}. \quad (8)$$

Sätt $k = 0$, då gäller att

$$A_1 = \frac{f(a + h, b) - f(a, b)}{h} \pm \rho(h, 0). \quad (9)$$

Vi vet att $\rho(h, k)$ är en funktion som går mot 0 då $(h, k) \rightarrow (0, 0)$, så den termen försvinner nu när vi låter $h \rightarrow 0$. Kvar får vi att

$$A_1 = \lim_{h \rightarrow 0} \frac{f(a + h, b) - f(a, b)}{h}, \quad (10)$$

vilket per definition är lika med $\frac{\partial f}{\partial x}(a, b)$. Att $A_2 = \frac{\partial f}{\partial y}(a, b)$ kan visas analogt. $\square$

Det är däremot viktigt att en reellvärd funktion kan mycket väl vara partiellt deriverbar i en, flera eller alla riktningar utan att vara differentierbar.

2.2 Funktioner av n variabler

Vidare vill vi ha ett generellt uttryck för den partiella derivatan i n dimensioner. Vi låter $f : \mathbb{R}^n \rightarrow \mathbb{R}$, vara en funktion av $\mathbf{x} = (x_1, \dots, x_n)$, där $\mathbf{a} = (a_1, \dots, a_n)$ är inre punkt i $D(f)$, och väljer ett godtyckligt $i \in \{1, \dots, n\}$. Då gäller att den partiella derivatan m.a.p x_i i $\mathbf{a}$ är

$$\frac{\partial f}{\partial x_i}(\mathbf{a}) = \lim_{h \rightarrow 0} \frac{f(\mathbf{a} + h\mathbf{e}_i) - f(\mathbf{a})}{h}, \quad (11)$$

där $\mathbf{e}_i$ är enhetsvektorn i riktningen x_i .

Detta gränsvärde påminner till synes om derivatans definition i envariabelanalys, och f är en reellvärd funktion av n variabler. Eftersom bara x_i varierar i gränsvärdet kan övriga variabler ses som konstanter vid gränsovergång och därmed kan den partiella derivatan i godtyckligt antal dimensioner beräknas på samma sätt som i envariabelanalys.

För att finna mening i hur de partiella derivatorna i n dimensioner tillämpas vid linjärisering kring en funktionspunkt behövs en generell definition för differentierbarhet i en punkt.

Definition 2.1 (Differentierbarhet i n dim). Låt $f : \mathbb{R}^n \rightarrow \mathbb{R}$, vara en funktion av $\mathbf{x} = (x_1, \dots, x_n)$, där $\mathbf{a} = (a_1, \dots, a_n)$ är en inre punkt i $D(f)$, och $\mathbf{h} = (h_1, \dots, h_n)$ är godtycklig. Vi säger att f är differentierbar i $\mathbf{a}$ om det finns konstanter $A_1, \dots, A_n$ och en funktion $\rho : \mathbb{R}^n \rightarrow \mathbb{R}$, så att

$$f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = h_1 A_1 + \dots + h_n A_n + \|\mathbf{h}\| \rho(\mathbf{h}) \text{ och att } \rho(\mathbf{h}) \rightarrow 0 \text{ då } \mathbf{h} \rightarrow \mathbf{0}. \quad (12)$$

2.3 Gradienten

Om funktionen är differentierbar i $\mathbf{a}$ så gäller, likt i 2 variabler (med liknande bevis), att $A_i = \frac{\partial f}{\partial x_i}(\mathbf{a})$ för alla $i \in \{1, \dots, n\}$. Härifrån kan gradienten till f (∇f), motsvarigheten till derivatan i flervariabelanalys härledas, vilken samlar alla partiella derivator i en vektor.

$$\nabla f(\mathbf{a}) = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathbf{a}) \\ \frac{\partial f}{\partial x_2}(\mathbf{a}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\mathbf{a}) \end{bmatrix} \quad (13)$$

3 Riktningsderivator

För att definiera riktningsderivata krävs att riktning definieras. En riktning i $\mathbb{R}^n$ definieras av en enhetsvektor i $\mathbb{R}^n$, alltså är mängden av alla riktningar i $\mathbb{R}^n$ mängden $\{\hat{\mathbf{u}} \in \mathbb{R}^n : \|\hat{\mathbf{u}}\| = 1\}$. Den totala mängden av alla riktningar bildar en enhetssfär och betecknas för $\mathbb{R}^n$ som $\mathbb{S}^{n-1}$. Utifrån detta kan riktningsderivata definieras rigoröst.

Definition 3.1 (Rikttningsderivata). Låt $f : \mathbb{R}^n \rightarrow \mathbb{R}$ vara differentierbar i $\mathbf{a}$. Låt $\hat{\mathbf{u}} \in \mathbb{S}^{n-1}$. Då definieras

$$\frac{\partial f}{\partial \hat{\mathbf{u}}}(\mathbf{a}) \stackrel{\text{def}}{=} \lim_{h \rightarrow 0} \frac{f(\mathbf{a} + h\hat{\mathbf{u}}) - f(\mathbf{a})}{h}. \quad (14)$$

Notera likheten mellan den här definitionen och definitionen av partiella derivator, ekvation (11). Relaterat till rikttningsderivata finns ett par satser.

Sats 2. Låt $f : \mathbb{R}^n \rightarrow \mathbb{R}$ vara differentierbar i $\mathbf{a}$. Låt $\hat{\mathbf{u}} \in \mathbb{S}^{n-1}$. Då gäller att rikttningsderivata $\frac{\partial f}{\partial \hat{\mathbf{u}}}(\mathbf{a})$ existerar och

$$\frac{\partial f}{\partial \hat{\mathbf{u}}}(\mathbf{a}) = \hat{\mathbf{u}} \cdot \nabla f(\mathbf{a}). \quad (15)$$

Bevis. Från den allmänna definitionen för differentierbarhet, Definition 2.1, följer att

$$\frac{f(\mathbf{a} + h\hat{\mathbf{u}}) - f(\mathbf{a})}{h} = \rho(h\hat{\mathbf{u}}) \frac{\|h\hat{\mathbf{u}}\|}{h} + \hat{\mathbf{u}} \cdot \nabla f(\mathbf{a}). \quad (16)$$

Vi har att $\frac{\|h\hat{\mathbf{u}}\|}{h} = \pm 1$, ty $\hat{\mathbf{u}}$ är en enhetsvektor, samt att $\lim_{h \rightarrow 0} \rho(h\hat{\mathbf{u}}) = 0$, då f är differentierbar i $\mathbf{a}$. Därför gäller, enligt Definition 3.1, att

$$\frac{\partial f}{\partial \hat{\mathbf{u}}}(\mathbf{a}) = \lim_{h \rightarrow 0} \frac{f(\mathbf{a} + h\hat{\mathbf{u}}) - f(\mathbf{a})}{h} = \hat{\mathbf{u}} \cdot \nabla f(\mathbf{a}). \quad (17)$$

Ekvation (17) gäller $\forall \hat{\mathbf{u}} \in \mathbb{S}^{n-1}$. Detta slutför beviset. □

Av Sats 2 följer Sats 3.

Sats 3. Låt $f : \mathbb{R}^n \rightarrow \mathbb{R}$ vara differentierbar i $\mathbf{a}$. Låt $\hat{\mathbf{u}} \in \mathbb{S}^{n-1}$. Då gäller

$$\frac{\partial f}{\partial \hat{\mathbf{u}}}(\mathbf{a}) \leq \|\nabla f(\mathbf{a})\|, \quad (18)$$

där likhet gäller om $\hat{\mathbf{u}}$ är riktad i samma riktning som $\nabla f(\mathbf{a})$.

Bevis. Enligt den tidigare bevisade satsen och definitionen av skalärprodukt gäller att

$$\frac{\partial f}{\partial \hat{\mathbf{u}}}(\mathbf{a}) = \hat{\mathbf{u}} \cdot \nabla f(\mathbf{a}) = \|\hat{\mathbf{u}}\| \|\nabla f(\mathbf{a})\| \cos \theta. \quad (19)$$

Här är θ den minsta vinkeln mellan $\hat{\mathbf{u}}$ och $\nabla f(\mathbf{a})$, således är $\theta \in [0, \pi]$. Då $\cos \theta \leq 1$ med likhet enbart då $\theta = 0$ samt att $\hat{\mathbf{u}}$ är en enhetsvektor, gäller att

$$\frac{\partial f}{\partial \hat{\mathbf{u}}}(\mathbf{a}) \leq \|\nabla f(\mathbf{a})\|, \quad (20)$$

med likhet endast om $\theta = 0$, alltså om $\hat{\mathbf{u}}$ är riktad i samma riktning som $\nabla f(\mathbf{a})$. □

4 Nivåytor

Definition 4.1 (Nivåhyperyta). Låt $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ vara en funktion av $n + 1$ variabler och $C \in \mathbb{R}$ en konstant. Ekvationen

$$f(x_1, \dots, x_{n+1}) = C \quad (21)$$

definierar i allmänhet en n -dimensionell hyperyta i $\mathbb{R}^{n+1}$. Den kallas för en nivåhyperyta till funktionen f .

Ekvationen definierar endast en hyperyta om C ingår i värdemängden till f , annars får vi den tomma mängden.

Beroende på värdet på n får vi olika saker ur ekvationen. Då $n = 1$ får vi en nivåkurva, för $n = 2$ får vi en nivåyta, och för $n > 2$ får vi en nivåhyperyta.

En viktig egenskap hos nivåytor är att i en godtycklig punkt $\mathbf{a}$ på nivåytan, är riktningsderivatan lika med noll i alla riktningar som hör till tangentplanet till nivåytan. För en riktning $\hat{\mathbf{u}} \in \mathbb{S}^n$ gäller alltså att $f_u(\mathbf{a}) = 0$ då riktningen $\hat{\mathbf{u}}$ tangerar den nivåyta som går genom punkten $\mathbf{a}$. Följdaktligen är $\nabla f(\mathbf{a})$ normal mot tangentplanet till f i $\mathbf{a}$ enligt (14).

Sats 4. Låt $f(x, y)$ vara en C^1 -funktion med $\nabla f \neq (0, 0)$. Då är $\nabla f(a, b)$ en normalvektor till den nivåkurva till f som går genom denna punkt. För en funktion $f(x, y, z)$ av tre variabler är analogt $\nabla f(a, b, c)$ normal i punkten (a, b, c) till nivåytan $f(x, y, z) = f(a, b, c)$.

5 Klasser

Definition 5.1 (Klasser). Låt $f : \mathbb{R}^n \rightarrow \mathbb{R}$ vara en funktion definierad på en domän $D \subseteq \mathbb{R}^n$ och $k \geq 0$ ett heltal.

f sägs vara av klassen $C^k(D)$ om alla dess partiella derivator av ordning mindre än eller lika med k existerar och är kontinuerliga funktioner i hela D .

Sats 5. Varje funktion av klassen C^1 är differentierbar.

5.1 Clairauts sats

Låt $f(x, y)$ vara en C^2 -funktion av 2 variabler. Eftersom f_x och f_y i sin tur är funktioner av x respektive y kan dessa funktioner deriveras med avseende på x respektive y .

Sats 6. Låt $D \subseteq \mathbb{R}^2$ vara en domän och låt $f : D \rightarrow \mathbb{R}$. Om $f \in C^2(D)$ då gäller i en omgivning till $(a, b) \in D$ att

$$f_{xy}(a, b) = f_{yx}(a, b). \quad (22)$$

6 Kedjeregeln

Kedjeregler är mycket viktiga i till exempel variabelbyten vid lösning av PDE:er och bygger på differentierbarhet av sammansatta funktioner.

Kedjeregler. Låt f, g vara funktioner sådana att

$$\left. \begin{array}{l} f : \mathbb{R}^q \rightarrow \mathbb{R}^m \\ g : \mathbb{R}^n \rightarrow \mathbb{R}^p \end{array} \right\} \text{ där } p = q \text{ för att sammansättningen } (f \circ g) \text{ ska vara definierad.}$$

Beteckna $f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$ med $\mathbf{x} = (x_1, \dots, x_p)$ och $g(\mathbf{t}) = (g_1(\mathbf{t}), \dots, g_p(\mathbf{t}))$ med $\mathbf{t} = (t_1, \dots, t_n)$. Då har vi sammansättningen:

$$\begin{aligned} (f \circ g)(\mathbf{t}) &= f(g(\mathbf{t})) \\ &= f(g(t_1, \dots, t_n)) \\ &= f(g_1(t_1, \dots, t_n), \dots, g_p(t_1, \dots, t_n)) \\ &= (f_1(g_1(t_1, \dots, t_n), \dots, g_p(t_1, \dots, t_n)), \dots, f_m(g_1(t_1, \dots, t_n), \dots, g_p(t_1, \dots, t_n))). \end{aligned}$$

Vi formulerar kedjeregeln för skalärvärda, sammansatta funktioner i tre steg:

① $p = 1$ och $n \in \mathbb{N}$

$$f : \mathbb{R} \rightarrow \mathbb{R}$$

$$g : \mathbb{R}^n \rightarrow \mathbb{R}$$

Alltså låt $(f \circ g)(\mathbf{t}) = f(g(\mathbf{t})) = f(g(t_1, \dots, t_n))$ och vi formulerar kedjeregeln:

$$\frac{\partial}{\partial t_j} f(g(\mathbf{t})) = f'(g(\mathbf{t})) \frac{\partial}{\partial t_j} g(\mathbf{t}). \quad (23)$$

② $n = 1$ och $p \in \mathbb{N}$

$$f : \mathbb{R}^p \rightarrow \mathbb{R}$$

$$g : \mathbb{R}^n \rightarrow \mathbb{R}^p$$

Alltså låt $g(t) = (g_1(t), g_2(t), \dots, g_p(t))$ och $f(g(t)) = f(g_1(t), g_2(t), \dots, g_p(t))$.

Kedjeregeln ger då

$$\frac{d}{dt} f(g_1(t), g_2(t), \dots, g_p(t)) = \sum_{j=1}^p \frac{\partial f}{\partial x_j}(g_1(t), g_2(t), \dots, g_p(t)) \cdot \frac{dg_j(t)}{dt}. \quad (24)$$

③ $n, p \in \mathbb{N}$

$$f : \mathbb{R}^p \rightarrow \mathbb{R}$$

$$g : \mathbb{R}^n \rightarrow \mathbb{R}^p$$

Alltså låt $\mathbf{t} = (t_1, t_2, \dots, t_n)$, $\mathbf{g} = (g_1(\mathbf{t}), g_2(\mathbf{t}), \dots, g_p(\mathbf{t}))$ och $f = f(\mathbf{g}(\mathbf{t}))$. kedjeregeln ger då

$$\frac{\partial}{\partial t_j} f(\mathbf{g}(\mathbf{t})) = \sum_{k=1}^p \frac{\partial}{\partial x_k} f(\mathbf{g}(\mathbf{t})) \cdot \frac{\partial g_k(\mathbf{t})}{\partial t_j}. \quad (25)$$