Riktningsderivata

En *riktning* $\mathbf{v}$ i $\mathbf{R}^n$ är en vektor av längd 1, $|\mathbf{v}| = 1$.
(I uppgifter kan det talas om t.ex. riktningen $u = (1,2,2)$. Det som avses är då den riktning som *bestäms* av $\mathbf{u}$, d.v.s *normeringen* $\mathbf{u}/|\mathbf{u}| = (1/3, 2/3, 2/3)$ av $\mathbf{u}$.)

Riktningsderivatan av den reellvärda funktionen $f(\mathbf{x})$ i punkten $\mathbf{a}$ i riktningen $\mathbf{v}$ definieras av

$$f'_{\mathbf{v}}(\mathbf{a}) = \lim_{t \to 0} \frac{f(\mathbf{a} + t\mathbf{v}) - f(\mathbf{a})}{t}$$

om gränsvärdet finns.

Riktningsderivatan i riktining $\mathbf{e}_i$ är, enligt definitionen, samma som $\dfrac{\partial}{\partial x_i} f(\mathbf{a})$.

Följande sats ger att man kan i allmänhet kan beräkna riktningsderivator med $\operatorname{grad} f(\mathbf{a})$.

Sats Om f är differentierbar så gäller

$$f'_{\mathbf{v}}(\mathbf{a}) = \operatorname{grad} f(\mathbf{a}) \cdot \mathbf{v}.$$

Beviset är att man använder kedjeregeln på funktionen $\phi(t) = f(\mathbf{a} + t\mathbf{v})$ och sätter $t = 0$.

Geometrisk tolkning av gradienten 1

Gradienten till en reelvärd funktion f i punkten $\mathbf{a}$ är vektorn

$$\nabla f(\mathbf{a}) = \operatorname{grad} f(\mathbf{a}) = (f'_{x_1}(\mathbf{a}), \ldots, f'_{x_n}(\mathbf{a})).$$

Att den är viktig har vi redan förstått: den bestämmer lineariseringen

$$L(\mathbf{x}) = f(\mathbf{a}) + \operatorname{grad} f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}),$$

men också alla riktningsderivator $f'_{\mathbf{v}}(\mathbf{a}) = \operatorname{grad} f(\mathbf{a}) \cdot \mathbf{v}$.
Eftersom $\mathbf{b} \cdot \mathbf{c} = |\mathbf{b}||\mathbf{c}| \cos\gamma$, där γ är vinkeln mellan $\mathbf{b}$ och $\mathbf{c}$ ger detta

$$f'_{\mathbf{v}}(\mathbf{a}) = |\operatorname{grad} f(\mathbf{a})||\mathbf{v}| \cos\gamma,$$

där γ är vinkeln mellan $\operatorname{grad} f(\mathbf{a})$ och $\mathbf{v}$. Riktningsderivatan är alltså störst när $\gamma = 0$ d.v.s när $\mathbf{v}$ är den riktning som bestäms av $\operatorname{grad} f(\mathbf{a})$.

Slutsats Gradienten $\operatorname{grad} f(\mathbf{a})$ anger den riktning i vilken funktionen f växer snabbast i punkten $\mathbf{a}$.

Tangentvektor och tangetlinje till parmetriserad kurva

Låt $\mathbf{x}(t)$, $t \in [a, b]$, vara en (paramertiserad kurva) i planet eller i rummet. Fixera tiden $t_0 \in [a, b]$.
Tangentvektorn vid tiden $t = t_0$ ges av

$$\mathbf{x}'(t_0) = (x'(t_0), y'(t_0), z'(t_0)).$$

Då parametriserar

$$\mathbf{l}(t) = \mathbf{x}(\mathbf{t_0}) + (t - t_0)\mathbf{x}'(t_0)$$

en linjen som passerar $\mathbf{x}(t_0)$ när $t = t_0$. Eftersom

$$\lim_{t \to t_0} \left| \frac{\mathbf{x}(t) - \mathbf{l}(t)}{t - t_0} \right| = \lim_{t \to t_0} \left| \frac{\mathbf{x}(t) - \mathbf{x}(t_0)}{t - t_0} - \mathbf{x}'(t_0) \right| = 0$$

förtjänar linjen genom $\mathbf{x}(t_0)$ i riktning $\mathbf{x}'(t_0)$ att kallas *tangentlinjen* till $\mathbf{x}(t)$ vid tiden t_0. Det är den linje som då bäst approximerar kurvan $\mathbf{x}(t)$.

Geometrisk tolkning av gradienten 2

Antag att $F(\mathbf{x}) = C$, där C är nån konstant, är en nivåyta eller nivåkurva som går genom punkten **a** (så att $F(\mathbf{a}) = C$).
Antag att kurvan $\mathbf{x}(t)$ löper på nivåkurvan/ytan och är i **a** när $t = 0$ ($\mathbf{x}(0) = \mathbf{a}$). Då gäller att $C = F(\mathbf{x}(t))$ vid varje tidpunkt t. Derivering m.a.p t ger

$$0 = \operatorname{grad} F(\mathbf{x}(t)) \cdot \mathbf{x}'(t)$$

Sätter vi $t = 0$ ser vi att $\operatorname{grad} F(\mathbf{a})$ är vinkelrät mot tangentvektorn till varje kurva på nivåytan/kurvan. Det betyder att

> $\operatorname{grad} F(\mathbf{a})$ är vinkelrät mot tangentplanet/tangenlinjen till nivåytan/kurvan till F genom **a**.

Högre ordningens partiella derivator

En funktion $f(\mathbf{x})$ kan ha partiella derivator $\frac{\partial}{\partial x_i} f(\mathbf{x})$ som kan deriveras. Detta ger *andra ordningens* partiella derivator:

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(\mathbf{x}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}) = f''_{x_i x_j}(\mathbf{x}) = f''_{ij}$$

Om det går, kan man fortsätta derivera och få tredje, fjärde, o.s.v ordningens partiella derivator.

En funktion f är av *klass* $\mathscr{C}^n$ i en öppen delmängd D av sin definitionsmängd om samtliga partiella derivator av ordning upp till och med n finns och är *kontinuerliga*.
I allmänhet är $f_{ij} \neq f_{ji}$, d.v.s det spelar roll i vilken ordnign man deriverar m.a.p olika variabler. Följande sats underlättar:

Sats 9 Om f är av klass $\mathscr{C}^2$ gäller att

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x})$$

Upprepad användning av satsen ger att om f är av klass $\mathscr{C}^n$ spelar det ingen roll i vilken ordnign man deriverar m.a.p olika variabler i partiella derivator upp t.o.m ordning n.

Taylor utveckling av grad 2

Från envariabeln (under lämpliga förutsättningar):

$$f(t) = f(a) + f'(a)(t-a) + \frac{1}{2!}f''(a)(t-a)^2 + (t-a)^3 B(t-a),$$

där $B(t-a)$ är begränsad i närheten av $t = a$. Om man för en reellvärd flervariabel funktion $f(\mathbf{x})$ sätter

$$\phi(t) = f(\mathbf{a} + t\mathbf{h})$$

får man en envariabel funktion. Formeln ovan för $\phi(t)$ leder då till (med $t = 1$)

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \operatorname{grad} f(\mathbf{a}) \cdot \mathbf{h} + \frac{1}{2}\sum_{i,j=1}^{n} f''_{ij}(\mathbf{a}) h_i h_j + |\mathbf{h}|^3 B(\mathbf{h}),$$

där $B(\mathbf{h})$ är begränsad i en omgivning till origo. Uttrycket

$$Q(\mathbf{h}) = \frac{1}{2}\sum_{i,j=1}^{n} f''_{ij}(\mathbf{a}) h_i h_j$$

är den kvadratiska termen i f kring **a** eller den *kvadratiska formen* av f i **a**.

Lokala extrempunkter

Funktionen $f(\mathbf{x})$ har ett *lokalt maximum* i punkten **a** om $f(\mathbf{x}) \leq \mathbf{a}$ för alla $\mathbf{x}$ i en omgivning till **a**.
Om $f(\mathbf{x}) < f(\mathbf{a})$ när $\mathbf{x} \neq \mathbf{a}$ är det ett lokalt *strängt* maximum.
Punkten **a** är då en *lokal maximipunkt* till f.
Lokal extrempunkt är synonymt med lokal maximi- eller minimipunkt.
Funktionen har ett *lokalt extremvärde* i en sådan punkt.
Följande sats ger ett sätt att hitta de lokala extrempunkterna som ligger i det *inre* av f:s defintionsmängd.

Sats 11 Om f har lokalt extremvärde i $\mathbf{a}$, som är en inre punkt i definitionsmängden gäller att

$$\operatorname{grad} f(\mathbf{a}) = \mathbf{0}.$$

En punkt där gradienten blir nollvektorn kallas en *stationär punkt* till f.
En lokal extrempunkt i det inre av defintionsmängden är alltså en stationär punkt. Observera att stationära punkter inte behöver vara lokala extrempunkter.

Lokala extrempunkter (forts.)

Kring en stationär punkt **a** har vi

$$f(\mathbf{a}+\mathbf{h}) = Q(\mathbf{h}) + |\mathbf{h}|^3 B(\mathbf{h}|$$

Under gynnsamma omständigheter kan man med

$$Q(\mathbf{h}) = \frac{1}{2}\sum_{i,j=1}^{n} f''_{ij}(\mathbf{a})h_i h_j$$

avgöra om f har ett lokalt extremvärde i **a**.

Kvadratiska former i två variabler

1. $Q(h,k)$ är positivt definit om $Q(h,k) > 0$ när $(h,k) \neq (0,0)$. f har då ett lokalt minimum i a.

2. $Q(h,k)$ är negativt definit om $Q(h,k) > 0$ när $(h,k) \neq (0,0)$. f har då ett lokalt maximum i a.

3. Om Q antar både postiva och negativa värden har f inte lokal extremvärde i a (sadelpunkt).

4. Om $Q(h,k) \geq 0$ eller $Q(h,k) \leq 0$, men antar värdet 0 också i andra punkter än origo är den semindefinit.
   Om Q är semindefinit kan man inte avgöra om f har lokalt extremvärde eller ej i a med hjälp av Q.

Karaktären på Q kan avgöras med kvadratkomplettering.

Differentialer

Vi har sett att $f(\mathbf{x})$ i punkten **a** har lineariseringen

$$L(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \operatorname{grad} f(a) \cdot \mathbf{h}$$

Den väsentliga delen är avbildningen $\mathbf{h} \mapsto \operatorname{grad} f(\mathbf{a}) \cdot \mathbf{h}$, som är linjär i samma mening som i kursen i linjär algebra.
Denna linjära avbildning kallas *differentialen* $df(\mathbf{a})$ till f i punkten **a**.
Vi har alltså

$$df(\mathbf{a})(\mathbf{h}) = \operatorname{grad} f(\mathbf{a}) \cdot \mathbf{h}$$

Tänk på x_i som funktionen $\mathbf{x} \mapsto x_i$ då gäller att $dx_i(\mathbf{a})(\mathbf{h}) = h_i$.
av detta följer att

$$df(\mathbf{a}) = f_1'(\mathbf{a})dx_1(\mathbf{a}) + f_2'(\mathbf{a})dx_2(\mathbf{a}) + \cdots + f_n'(\mathbf{a})dx_n(\mathbf{a})$$

eller enklare

$$df = f_1'dx_1 + f_2'dx_2 + \cdots f_n'dx_n.$$