FÖ 2.3 Gradient och riktningsderivata

I förra föreläsningen:

$$f : \mathbb{R}^n \to \mathbb{R}, \quad f(x) = f(x_1, \dots, x_n)$$

Derivatan (Jacobimatrisen) är

$$f'(a) = Df(a) = [f_1'(a), \dots, f_n'(a)] \quad \text{(rad-matris)}$$

Nu ska vi använda geometriska beteckningar.

Gradienten till $f(x,y)$ är

$$\nabla f(x,y) = f_x'(x,y)\,\mathbf{i} + f_y'(x,y)\,\mathbf{j}$$

Obs att

* $\nabla f(x,y)$ är en vektorvärd funktion (vektorfält) (vector field)

* $f(x,y)$ är en <u>skalärvärd funktion</u> (<u>skalärt fält</u>) (scalar field)

* <u>nabla operatorn</u> är

$$\nabla = \frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j}$$

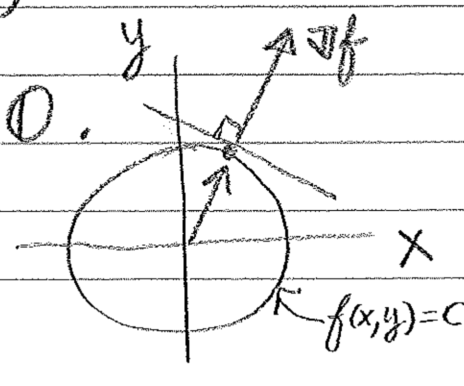
* Nabla verkar på f:

$$\nabla f = \left(\frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j}\right) f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}$$

<u>Exempel</u> $f(x,y) = x^2 + y^2$

$$\nabla f(x,y) = 2x\,\mathbf{i} + 2y\,\mathbf{j} = 2\mathbf{r}$$

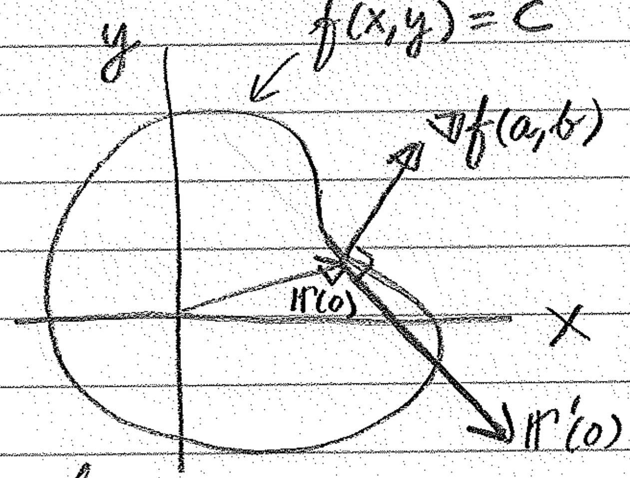
Obs: ∇f är ortogonal mot nivåkurvorna till f, utom i origo där $\nabla f(0,0) = 0$.



Sats 6 Antag att

* f är deriverbar i (a,b).
* $\nabla f(a,b) \neq \mathbb{0}$.

Då är $\nabla f(a,b)$ en normalvektor till nivåkurvan till f genom (a,b).



Bevis Parametrisera nivåkurvan,

$$\mathbb{r} = x(t)\mathbb{i} + y(t)\mathbb{j}$$

så att $x(0) = a$, $y(0) = b$.
Eftersom $f(x,y) = c$ på kurvan så gäller

$$f(x(t), y(t)) = c.$$

Kedjeregeln ger

$$0 = \frac{d}{dt} f(x(t), y(t)) =$$

$$= f'_x(x(t), y(t))\,x'(t) + f'_y(x(t), y(t))\,y'(t)$$

$$= \left(f'_x(x(t), y(t))\,\mathbb{i} + f'_y(x(t), y(t))\,\mathbb{j}\right) \bullet \left(x'(t)\mathbb{i} + y'(t)\mathbb{j}\right)$$

$$= \nabla f(x(t), y(t)) \bullet \mathbb{r}'(t).$$

Speciellt med $t=0$ får vi

$$\nabla f(a,b) \cdot \mathbf{r}'(0) = 0,$$

dvs $\nabla f(a,b)$ är ortogonal mot tangenten $\mathbf{r}'(0)$.

Obs att f måste vara deriverbar för att vi ska kunna använda kedjeregeln. Det räcker inte att f är partiellt deriverbar.

Gradienten i 3 variabler:

$$\nabla f(x,y,z) = f'_x(x,y,z)\mathbf{i} + f'_y(x,y,z)\mathbf{j} + f'_z(x,y,z)\mathbf{k}$$

Riktningsderivata:

$$D_{\mathbf{u}} f(a,b,c) = \mathbf{u} \cdot \nabla f(a,b,c)$$

med $\mathbf{u} = u\mathbf{i} + v\mathbf{j} + w\mathbf{k}$, $|\mathbf{u}| = 1$.

Riktningsderivata

Låt $\mathbf{u} = u\mathbf{i} + v\mathbf{j}$ vara en enhetsvektor

dvs $|\mathbf{u}| = \sqrt{u^2 + v^2} = 1$.

Riktningsderivatan av f i (a,b) i riktningen $\mathbf{u}$ är

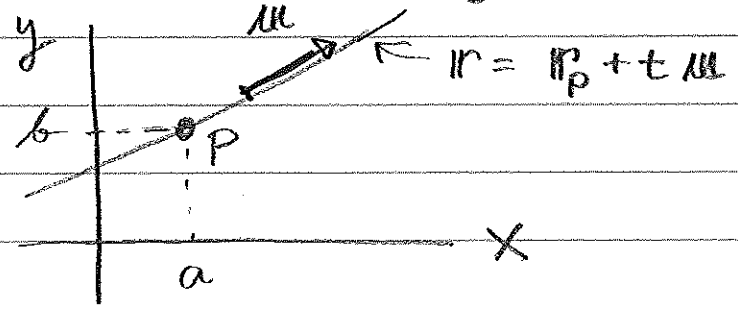
$$D_{\mathbf{u}} f(a,b) = \frac{d}{dt} f(a + tu, b + tv)\Big|_{t=0} .$$

Obs att vi deriverar envariabelfunktionen

$$g(t) = f(a + tu, b + tv)$$

i punkten $t = 0$, dvs $D_{\mathbf{u}} f(a,b) = g'(0)$.

Funktionen $g(t)$ evaluerar f längs den räta linjen genom $P = (a,b)$ med riktningsvektorn $\mathbf{u}$.



<u>Sats 7</u> Om f är deriverbar i (a,b) så gäller

$$D_{\mathbf{u}} f(a,b) = \mathbf{u} \bullet \nabla f(a,b).$$

<u>Bevis</u> Kedjeregeln ger

$$\frac{d}{dt} f(a+tu, b+tv) =$$

$$= f_x'(a+tu, b+tv)u + f_y'(a+tu, b+tv)v$$

$$= \nabla f(a+tu, b+tv) \bullet \mathbf{u}$$

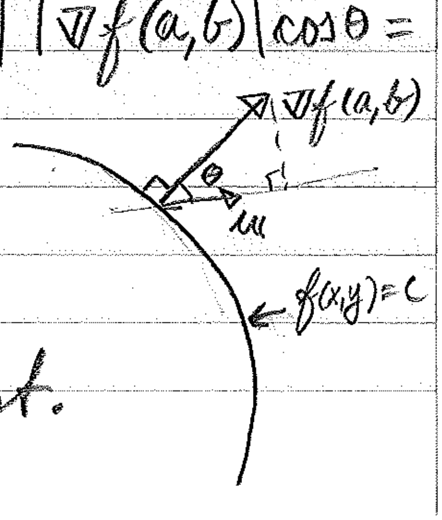
Med $t=0$ får vi $D_{\mathbf{u}} f(a,b) = \mathbf{u} \bullet \nabla f(a,b)$.

Obs: $\mathbf{u} \bullet \nabla f(a,b)$ är den skalära projektionen av gradientvektorn på riktningen $\mathbf{u}$. (Adams 10.2)

$$D_{\mathbf{u}} f(a,b) = \mathbf{u} \bullet \nabla f(a,b) = |\mathbf{u}|\,|\nabla f(a,b)| \cos\theta =$$

$$= |\nabla f(a,b)| \cos\theta$$

är maximal då $\cos\theta = 1$.



Alltså: $\nabla f(a,b)$ pekar i den riktningen där f ökar mest.

Skippa 12.8.

12.9 Taylors formel.

Endast Taylor-polynom av grad 2. (sid 737-739)

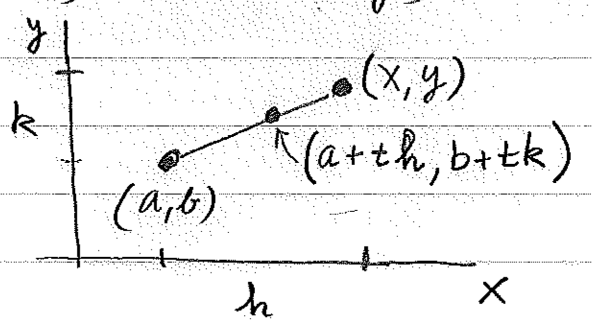
Kom ihåg Taylors formel i en variabel: $(x = a+h)$

$$F(x) = F(a+h) = F(a) + F'(a)h + \frac{1}{2}F''(a)h^2 + R_2(a,h)$$

med $R_2(a,h) = \frac{1}{3!}F'''(s)h^3$ där s är mellan a och x.

Två variabler: $x = a+h$, $y = b+k$

Bilda en en-variabelfunktion genom att evaluera $f(x,y)$ längs räta linjen från (a,b) till (x,y).



$$F(t) = f(a+th, b+tk)$$

$$0 \le t \le 1$$

$$F(0) = f(a,b)$$

$$F(1) = f(x,y) = f(a+h, b+k)$$

Skriv med Taylors formel för F:

$$F(1) = F(0) + F'(0)1 + \frac{1}{2}F''(0)1^2 + \frac{1}{6}F'''(s)1^3$$

där $0 \le s \le 1$.

Här är (kedjeregeln!)

$$F'(t) = \frac{d}{dt} f(a+th,\, b+tk) =$$

$$= f'_x(a+th,\, b+tk)h + f'_y(a+th, b+tk)k$$

$$F'(0) = f'_x(a,b)h + f'_y(a,b)k =$$

$$= [f'_x(a,b) + f'_y(a,b)]\begin{bmatrix} h \\ k \end{bmatrix}$$

$$F''(t) = f''_{xx}(a+th,\, b+tk)h^2 +$$

$$+ f''_{xy}(a+th,\, b+tk)hk \quad ''$$

$$\left(f''_{xy} = f''_{yx}\right) \qquad + f''_{yx}(a+th,\, b+tk)kh$$

$$+ f''_{yy}(a+th,\, b+tk)k^2$$

$$F''(0) = f''_{xx}(a,b)h^2 + 2f''_{xy}(a,b)hk + f''_{yy}(a,b)k^2$$

$$= [h,k]\begin{bmatrix} f''_{xx}(a,b) & f''_{xy}(a,b) \\ f''_{xy}(a,b) & f''_{yy}(a,b) \end{bmatrix}\begin{bmatrix} h \\ k \end{bmatrix}$$

Taylors polynom av grad 2:

$$P_2(x,y) = f(a,b) + f'_x(a,b)h + f'_y(a,b)k +$$
$$+ \frac{1}{2}\left(f''_{xx}(a,b)h^2 + 2f''_{xy}(a,b)hk + f''_{yy}(a,b)k^2\right)$$

$$= f(a,b) + \left[f'_x(a,b),\ f'_y(a,b)\right]\begin{bmatrix} h \\ k \end{bmatrix} +$$

$$+ \frac{1}{2}\begin{bmatrix} h & k \end{bmatrix}\begin{bmatrix} f''_{xx}(a,b) & f''_{xy}(a,b) \\ f''_{xy}(a,b) & f''_{yy}(a,b) \end{bmatrix}\begin{bmatrix} h \\ k \end{bmatrix}$$

Obs (gradientvektorn) Jacobi-matrisen

$$\nabla f(a,b) = f'(a,b) = \left[f'_x(a,b) \quad f'_y(a,b)\right]$$ (rad matris)

och Hesse-matrisen:

$$D^2 f(a,b) = \begin{bmatrix} f''_{xx}(a,b) & f''_{xy}(a,b) \\ f''_{xy}(a,b) & f''_{yy}(a,b) \end{bmatrix}$$ (symmetrisk matris)

Taylors formel (grad 2):

$$f(x,y) = P_2(x,y) + R_2(x,y)$$

där $R_2(x,y) = \frac{1}{6}F'''(s) = \frac{1}{6}\Big(f'''_{xxx}h^3 + 3f'''_{xxy}h^2k +$

$$+ 3f'''_{xyy}hk^2 + f'''_{yyy}k^3\Big)_g \quad \text{där}$$

alla derivatorna evalueras i den (okända) mellanliggande punkten $(a+sh,\, b+sk)$, $0 \le s \le 1$.

(Resttermen kan ej skrivas på matrisform. Varför?)

Mer allmänt: $f : \mathbb{R}^n \to \mathbb{R}$, $h = \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix}$

$$f(x) = f(a+h) =$$

$$= f(a) + f'(a)h + \frac{1}{2} h^T f''(a) h + R_2(x)$$

där $f'(a) = \left[\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n} \right]$

$$f''(a) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{bmatrix} \quad n \times n$$

symmetrisk: $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$

Obs formen: $f(a+h) = f(a) + [\cdot \cdot \cdot] \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} + \frac{1}{2} [\cdot \cdot \cdot] \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} + R$

$$= f'(a) + f'(a)h + \frac{1}{2} h^T f''(a) h + R_2$$