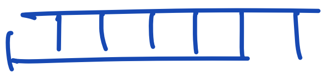
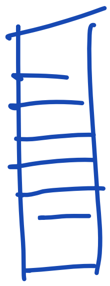
$\boxed{F06}$

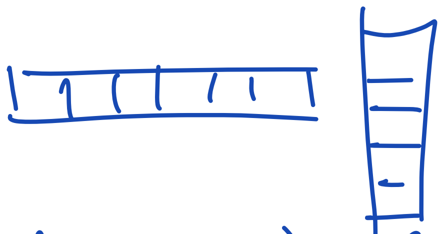
- Gradientmetoden
- Lagranges metod

- **Repetition**

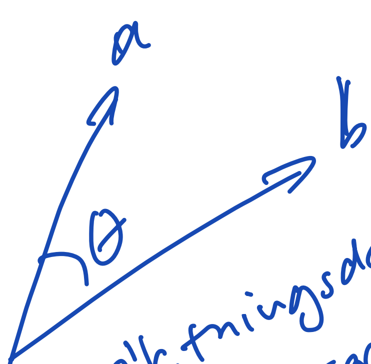
$f : \mathbb{R}^n \to \mathbb{R} \quad (m = 1)$

$f' =$  $1 \times n$ (radvektor)

$\nabla f = (f')^T =$  $= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}, \ldots \right)$

$\Delta y \approx f' \Delta x =$ 

$\Delta y \approx \nabla f \cdot \Delta x = (\cdot, \cdot, \cdot) \cdot (\cdot, \cdot, \cdot)$



$a \cdot b = \cos(\theta) \cdot \|a\| \cdot \|b\|$

Störst då $\theta = 0$

riktningsderivata

gradient

$D_{\hat{u}} f = \nabla f \cdot \hat{u}$

enhetsvektor

$\nabla f \text{ "=" } f'$

- Gradientmetoden

Numerisk metod för lösning av fria optimeringsproblem, dvs optimering utan bivillkor

Sök $\min -f(x)$

$$\min_{x \in \mathbb{R}^n} f(x) \qquad \left(\max_{x \in \mathbb{R}^n} f(x) \right)$$

Vi söker

$$\begin{cases} \min_{x \in \mathbb{R}^n} f(x) & \text{extremvärde} \\ \bar{x} = \arg\min_{x \in \mathbb{R}^n} f(x) & \text{extrempunkt} \end{cases}$$

Vi vet att $D_{\hat{u}} f$ minimal i riktningen

$$\hat{u} = -\nabla f / \|\nabla f\|$$

$$\hat{u} \parallel -\nabla f$$

Gradientmetod söker minimum genom att stega i negativa gradientens riktning.

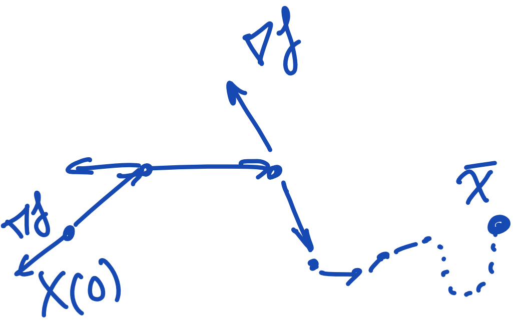
Algoritm:

h_k = steglängd

$x(0)$ = startgissning (vektor)

$$x(k+1) = x(k) - h_k \nabla f(x(k))$$

Ger lokalt minimum

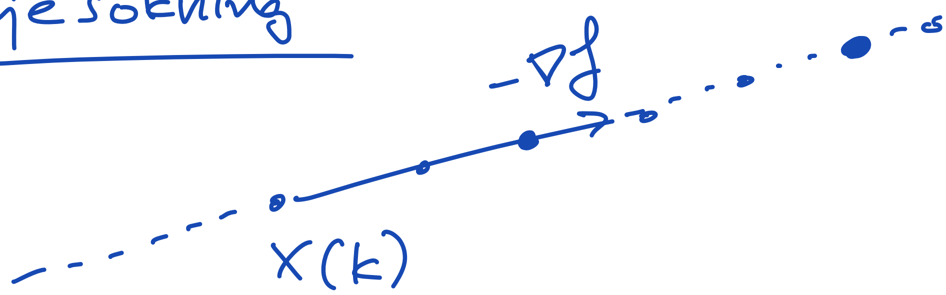


$$\bar{x} = \arg\min f(x)$$

Hur välja h_k?

$$h_k = \arg\min_h \{x(k) - h\nabla f(x_k)\}$$

Kalla linjesökning

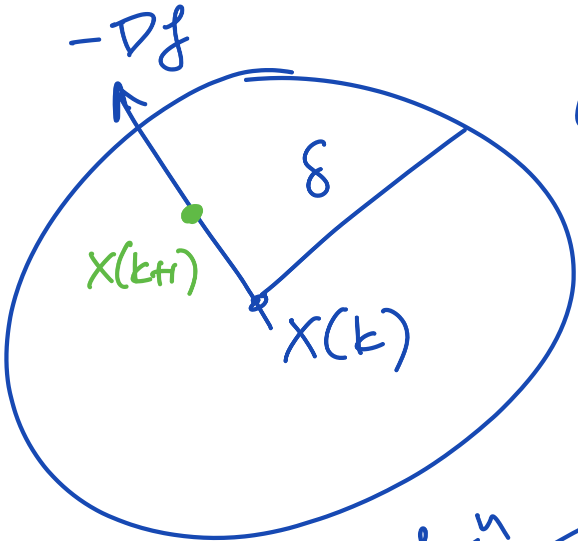


Enklare:

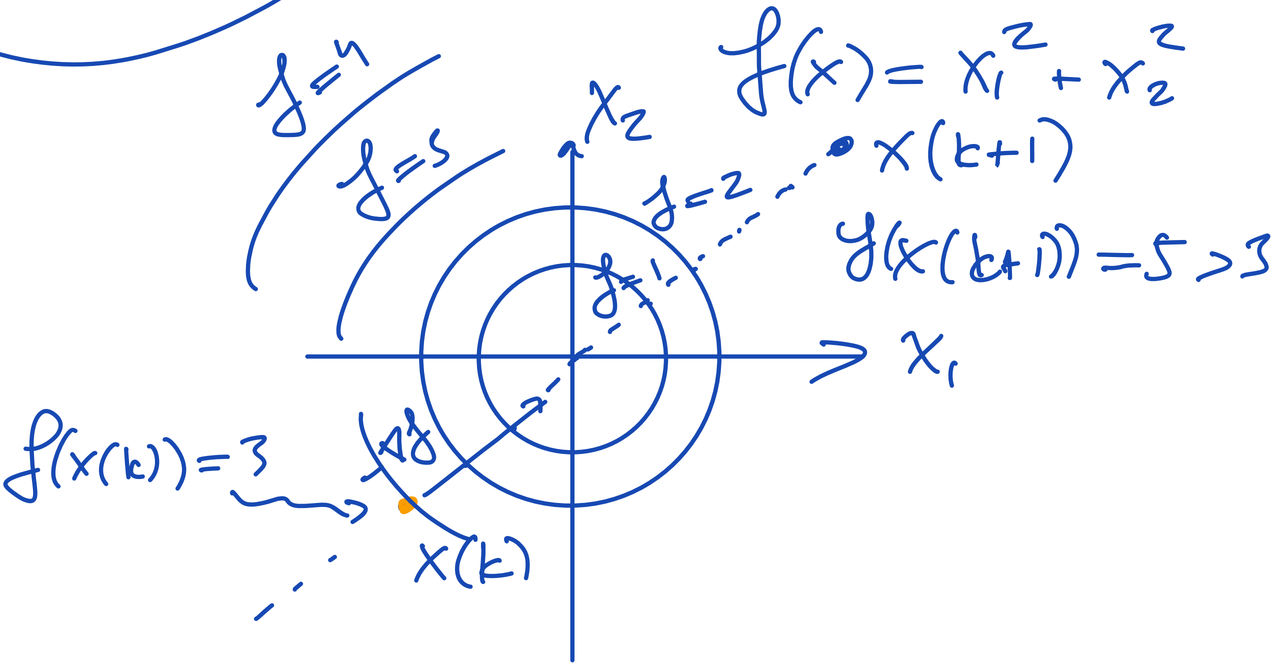
- Testa med $h=1$ (eller 0.1)
- Om $f(x(k) - h\cdot\nabla f(x(k))) < f(x(k))$

  tag $x(k+1) =$
- Annars: Tag $h \leftarrow h/2$ och upprepa

Omgivning



$$f(x(k+1)) < f(x(k))$$



· **Relation till Newtons metod**

Vill minimera $f(x)$

Nödvändigt villkor att

$$\nabla f = 0$$

$$f : \mathbb{R}^n \to \mathbb{R} \quad (m = 1)$$

$$\nabla f : \mathbb{R}^n \to \mathbb{R}^n \quad \text{tq} \quad \nabla f = \nabla f(x) \in \mathbb{R}^n$$

$\therefore \quad \nabla f = 0$ en ~~ekvation~~ ekvationssystem med n obekanta

Låt $\quad F(x) = \nabla f(x) = f'(x)$

$$F : \underline{\underline{\mathbb{R}^n \to \mathbb{R}^n}}$$

Lös $\quad F(x) = \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

<u>Newtons metod:</u>

$$x(k+1) = x(k) - F'(x(k))^{-1} F(x(k))$$

$$F = f' \qquad\qquad = \text{Jacobi-matrisen}^T$$

$$F' = (f')' = f'' \qquad = \text{Hesse-matrisen}$$

$$\underbrace{x(k+1) = x(k) - \left(f''(x(k))\right)^{-1} \nabla f(x(k))}$$

Jämför med gradientmetoden:

$$\underbrace{x(k+1) = x(k) - h_k \cdot \nabla f(x(k))}$$

Kan ses som fixpunktsiteration för lösning av $\nabla f = 0$ med $\alpha = -h_k$.

- ## Lagranges metod

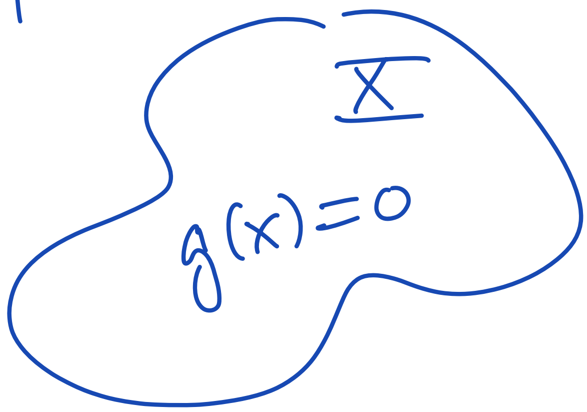
Optimering med bivillkor

Vi har då:

$$\begin{cases} f : \mathbb{R}^n \to \mathbb{R} \\ g : \mathbb{R}^n \to \mathbb{R}^p, \quad p < n \end{cases}$$

Vi söker

$$\min_{x \in X} f(x)$$



där

$$X = \{x \in \mathbb{R}^n \mid g(x) = 0\}$$

Notera: $p < n$ ty annars sök $g(x) = 0$

Exempel:

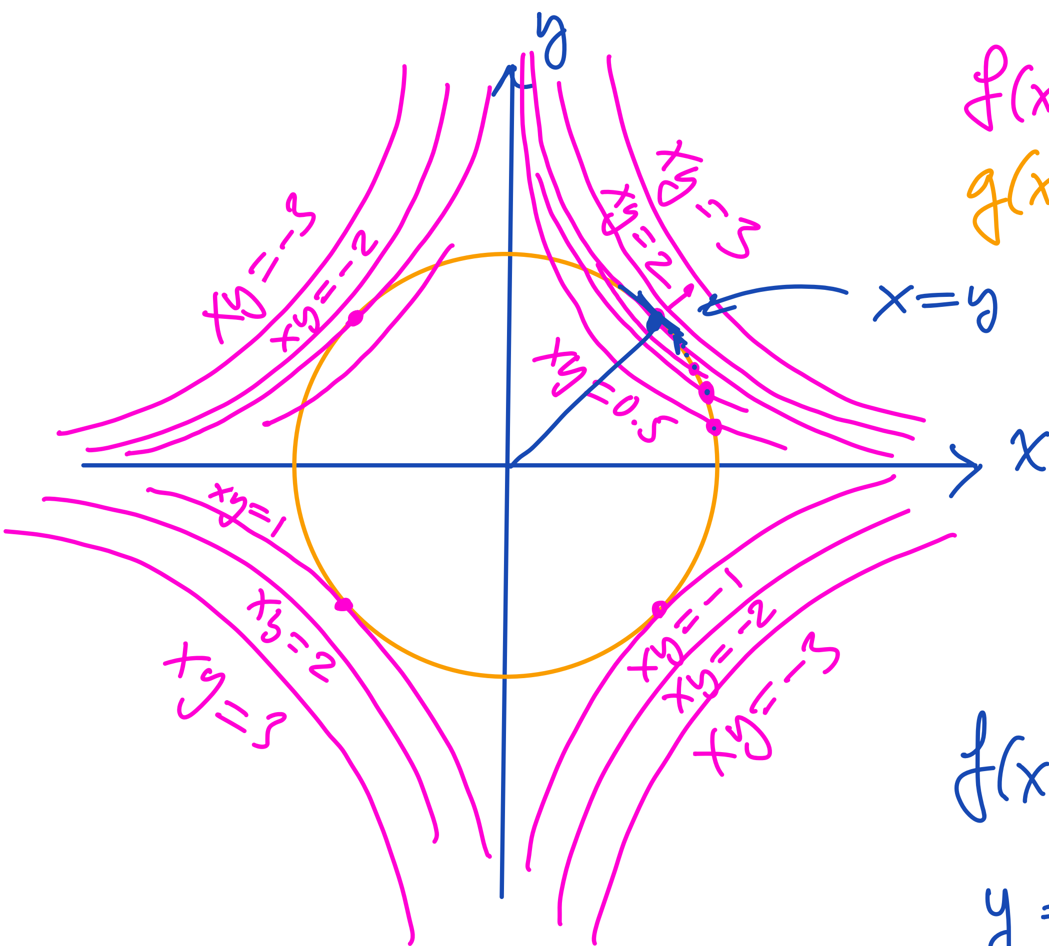
$$\begin{cases} f(x,y) = xy \\ g(x,y) = x^2 + y^2 - 1 \quad (=0) \end{cases}$$

$$f : \mathbb{R}^2 \to \mathbb{R} \qquad m = 1$$

$$g : \mathbb{R}^2 \to \mathbb{R} \qquad p = 1 < n = 2$$

Hitta max av $\underline{\underline{x \cdot y}}$ då $x^2+y^2-1=0$,

dvs $x^2+y^2=1$



$f(x,y)=xy$

$g(x,y)=x^2+y^2-1=0$

$x=y$ pga symmetri

$f(x,y)=xy=C$

$y=C/x$

$x=y \qquad x^2+y^2=1$

$$x^2+x^2=1$$

$$2x^2=1$$

$$x^2=1/2$$

$$x=\pm 1/\sqrt{2}$$

$$y=\pm 1/\sqrt{2}$$

$$\max f(x,y)=\frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}=\frac{1}{2}$$

Vi såg att extrempunkten inträffar då nivåkurvorna tangerar bivillkoret g.

till f

Vet att gradienten är normal till nivåytorna!

Nivåyta till f tangerar g

$$\Longleftrightarrow$$

$$\nabla f \parallel \nabla g$$

$\exists\, \lambda \in \mathbb{R}$ s.a. $\nabla f = -\lambda \nabla g$

Kontroll:

$$f(x,y) = xy \implies \nabla f = (y, x)$$

$$g(x,y) = x^2 + y^2 - 1 \implies \nabla g = (2x, 2y)$$

$$\nabla f = -\lambda \cdot \nabla g$$

$$(y, x) = -\lambda (2x, 2y) = (-2\lambda x, -2\lambda y)$$

$$\begin{cases} y = -2\lambda x & (1) \\ x = -2\lambda y & (2) \\ x^2 + y^2 - 1 = 0 & (3) \end{cases}$$

(1) $-2\lambda = \dfrac{y}{x}$

(2) $-2\lambda = \dfrac{x}{y}$

$$\implies \frac{y}{x} = \frac{x}{y} \implies x^2 = y^2$$

(3) $x^2 + x^2 = 1$

$$2x^2 = 1$$

$$x = \pm 1/\sqrt{2} \implies y = \pm 1/\sqrt{2}$$

Ger 4 st. extrempunkter

$$(x,y) = \begin{cases} \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) & \text{maximum} \\ \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) & \text{minimum} \\ \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) & \text{minimum} \\ \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) & \text{maximum} \end{cases}$$

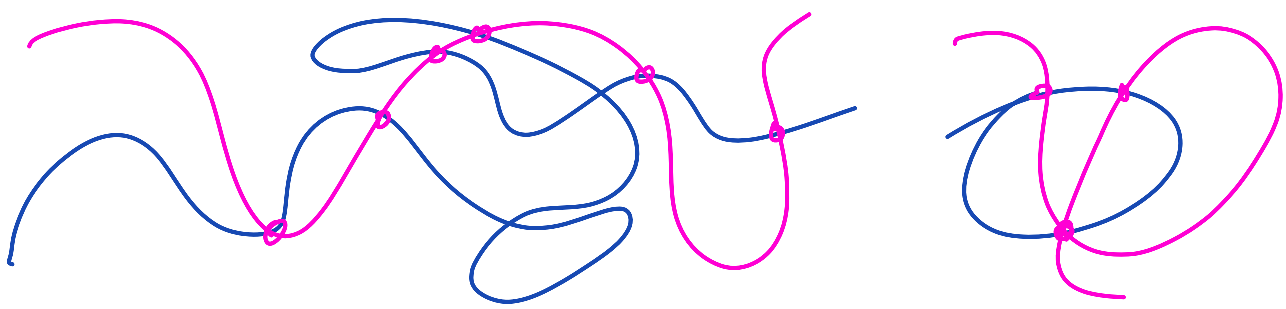
$f = x \cdot y$

Lagranges metod: (med ett bivillkor)

$f: \mathbb{R}^n \to \mathbb{R}$ deriverbar

$g: \mathbb{R}^n \to \mathbb{R}$ deriverbar

$p = 1$ (antal bivillkor)



Bilda Lagrange-funktionen:

$$\boxed{L(x, \lambda) = f(x) + \lambda g(x)}$$

Om f har extrempunkt i $\bar{x}$ med bivillkor g, så har L en stationär punkt i $\bar{x}$, dvs:

$$\boxed{\nabla_{(x,\lambda)} L(\bar{x},\bar{\lambda}) = 0}$$

Bevis: $n+1$ variabler: $x_1\ x_2\ x_3 \dots x_n\ \lambda$

$$L(\underbrace{x}, \lambda) = f(x) + \lambda g(x)$$

$$\Rightarrow \nabla L = \begin{bmatrix} \frac{\partial f}{\partial x_1} + \lambda \frac{\partial g}{\partial x_1} \\ \frac{\partial f}{\partial x_2} + \lambda \frac{\partial g}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} + \lambda \frac{\partial g}{\partial x_n} \\ g(x) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

$\nabla L = 0$

De första n ekvationerna säger att

$$\begin{cases} \frac{\partial f}{\partial x_1} + \lambda \frac{\partial g}{\partial x_1} = 0 \\ \vdots \\ \frac{\partial f}{\partial x_n} + \lambda \frac{\partial g}{\partial x_n} = 0 \end{cases}$$

$$\nabla f + \lambda \nabla g = 0$$

$$\nabla f = -\lambda \nabla g \Leftrightarrow \nabla f \parallel \nabla g$$

Sista ekvationen: $g(x)=0$ (bivillkoret)

$\blacksquare$

Exempel:

minimera (eller maximera)

$$f(x,y) = \underline{\underline{xy}}$$

då

$$g(x,y) = x^2+y^2-1=0$$

Bilda Lagrange-funktionen

$$L(x,y,\lambda) = f(x,y) + \lambda g(x,y)$$

$$= \underbrace{xy + \lambda\cdot(x^2+y^2-1)}$$

$$\nabla L = \begin{bmatrix} \frac{\partial L}{\partial x} \\ \frac{\partial L}{\partial y} \\ \frac{\partial L}{\partial \lambda} \end{bmatrix} = \begin{bmatrix} y+2\lambda x \\ x+2\lambda y \\ x^2+y^2-1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{cases} y + 2\lambda x = 0 \\ x + 2\lambda y = 0 \\ x^2 + y^2 = 1 \end{cases}$$

Lösning:

$$(x, y) = \underline{\underline{\left(\pm 1/\sqrt{2}, \pm 1/\sqrt{2}\right)}}$$

Lagranges metod med $p > 1$

$$f : \mathbb{R}^n \to \mathbb{R}$$

$$g : \mathbb{R}^n \to \mathbb{R}^p \quad (p < n)$$

$$\boxed{L(x, \lambda) = f(x) + \lambda^T g}$$

$$\lambda^T g = [\lambda_1 \ \lambda_2 \ \cdots \ \lambda_p] \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_p \end{bmatrix}$$

$$= \lambda_1 g_1 + \lambda_2 g_2 + \cdots \lambda_p g_p$$

∇L m.a.p. x och λ $= 0$