Obegränsad optimering

Problem: $\min_{x \in \mathbb{R}^n} f(x)$ (där $f \in C^2$ på $\mathbb{R}^n$)

↑ obs!

Optimalitetsvillkor?

Sats: (analys)

x^* är ett lokalt minimum $\Rightarrow \nabla f(x^*) = 0$

Sats: (analys)

$$\left.\begin{array}{l} \nabla f(x^*) = 0 \\ \nabla^2 f(x^*) \text{ positivt definit} \end{array}\right\} \Rightarrow x^* \text{ är ett lokalt minimum}$$

Sats:

$$\left.\begin{array}{l} f \text{ konvex på } \mathbb{R}^n \\ \nabla f(x^*) = 0 \end{array}\right\} \Rightarrow x^* \text{ är ett globalt minimum}$$

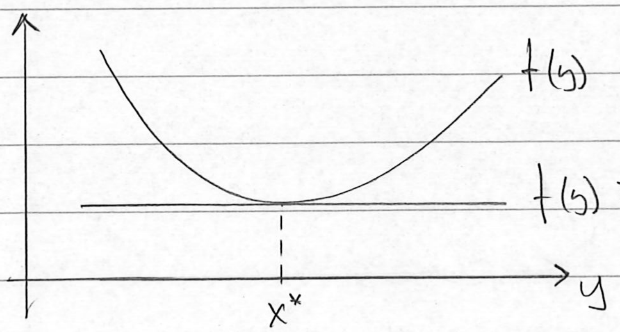
Bevis: f konvex på $\mathbb{R}^n \Rightarrow$

$\Rightarrow f(y) \geq f(x) + \nabla f(x)^T (y-x)$ för alla $x, y \in \mathbb{R}^n$

$\underset{x=x^*}{\Rightarrow} f(y) \geq f(x^*) + \nabla f(x^*)^T (y-x)$ för alla $y \in \mathbb{R}^n$ x -> x*

$\underset{\nabla f(x^*)=0}{\Rightarrow} f(y) \geq f(x^*)$ för alla $y \in \mathbb{R}^n$

$\Rightarrow x^*$ globalt minimum



$f(y) \geq f(x^*) + \nabla f(x^*)^T (y-x) = f(x^*)$

x -> x*

./.

Praktisk lösning: iterativ sökmetod.

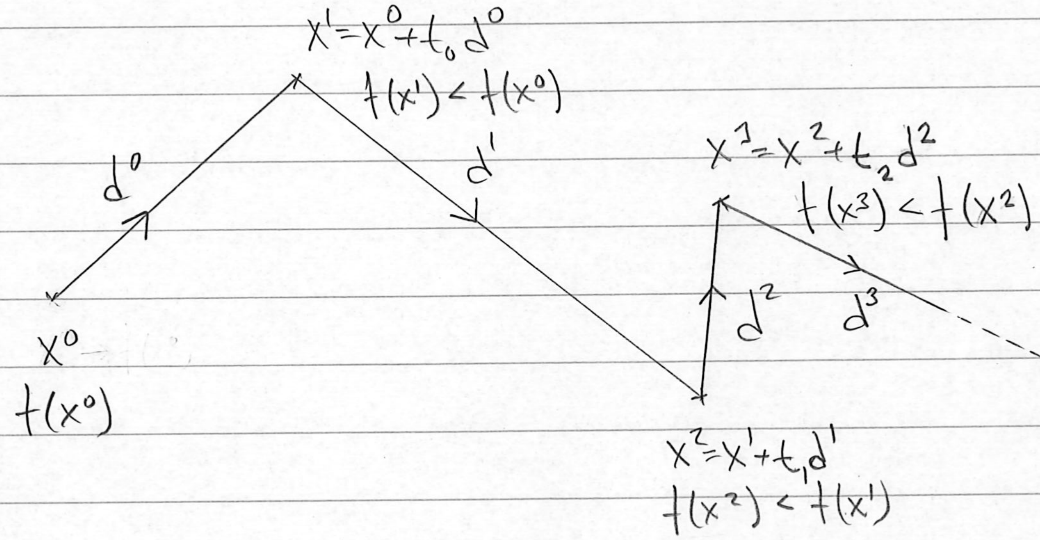
- Välj startpunkt $x^0 \in R^n$. Låt $k=0$.

- Bestäm en sökriktning $d^k \neq 0$ sådan att $f(x^k+td^k) < f(x^k)$ gäller för små $t>0$.

- Bestäm en optimal steglängd $t_k > 0$ genom linjesökningen $\min_{t \geq 0} \varphi(t) = f(x^k+td^k)$.
  ty minimeringsproblem

- Ny iterationspunkt: $x^{k+1} = x^k + t_k d^k$
  Sätt $k=k+1$.

- Upprepa tills $\nabla f(x^{k+1}) \approx 0$.



$$\left.\begin{array}{l} \bullet\ f(x^k+td^k) < f(x^k) \text{ för små } t>0 \\ \bullet\ t_k \text{ löser } \min_{t \geq 0} \varphi(t) = f(x^k+td^k) \\ \bullet\ x^{k+1} = x^k + t_k d^k \end{array}\right\} \Rightarrow f(x^{k+1}) < f(x^k)$$

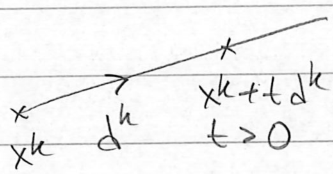
Alltså: $f(x^0) > f(x^1) > f(x^2) > f(x^3) > \ldots$

För vilka $d^k \neq 0$ gäller $f(x^k + td^k) < f(x^k)$ för små $t > 0$?

Första ordningens Taylor-utveckling av $f(x^k + td^k)$ map t:

$$f(x^k+td^k) = f(x^k) + \left[\frac{d}{dt} f(x^k+td^k)\Big|_{t=0}\right] t + O(t^2)$$

$$\frac{d}{dt} f(x^k+td^k) \underset{\substack{\uparrow \\ \text{kedjeregeln}}}{=} \nabla f(x^k+td^k)^T \frac{d}{dt}(x^k+td^k) =$$



$$= \nabla f(x^k+td^k)^T d^k$$

$$\frac{d}{dt} f(x^k+td^k)\Big|_{t=0} = \nabla f(x^k)^T d^k \quad \text{(riktningsderivata)}$$

Alltså: $f(x^k+td^k) = f(x^k) + \nabla f(x^k)^T d^k + O(t^2)$

lägg till t

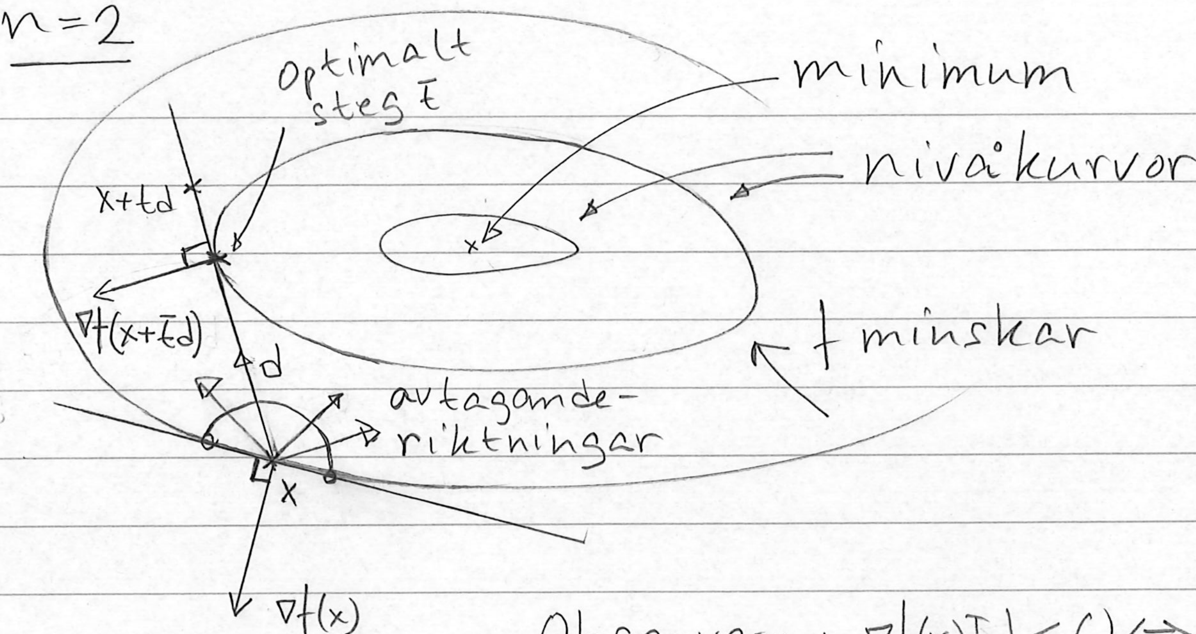
Slutsats: $f(x^k+td^k) < f(x^k)$ gäller för små $t>0$ omm $\nabla f(x^k)^T d^k < 0$!

Definition: En riktning $d \neq 0$ med $\nabla f(x)^T d < 0$ kallas en avtagenderiktning (eller descentriktning) för f i x.

Riktningen $d^k \neq 0$ ska alltså väljas som en avtagenderiktning.

Maximering: $d^k \neq 0$ ska uppfylla $\nabla f(x^k)^T d^k > 0$ och ger då $f(x^k+td^k) > f(x^k)$ för små $t>0$. Ett $d \neq 0$ med $\nabla f(x)^T d > 0$ kallas en ascentriktning för f i x.

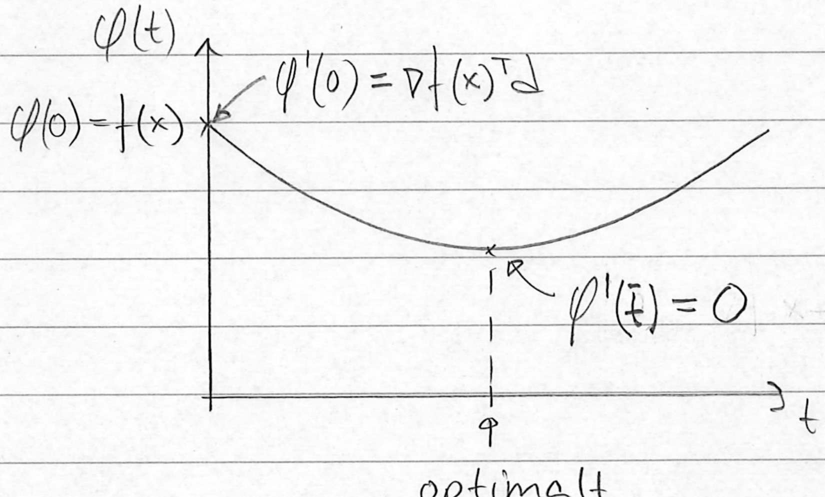
$n=2$



Observera: $\nabla f(x)^T d < 0 \Leftrightarrow$ vinkeln mellan d och $\nabla f(x)$ är strikt större än $\pi/2$.

$\varphi(t) = f(x+td), \; t \geq 0$

$\varphi'(t) = \nabla f(x+td)^T d$



Observera: $\varphi'(\bar{t}) = \nabla f(x+\bar{t}d)^T d = 0 \Leftrightarrow \nabla f(x+\bar{t}d) \perp d.$