Föreläsning 12

TAOP07 – Optimeringslära grundkurs

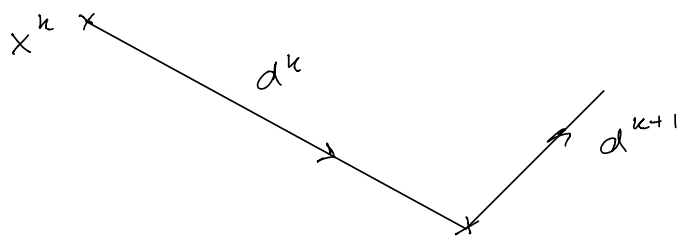
Icke-linjär optimering

Skriven av Oliver Wettergren
oliwe188@student.liu.se
https://www.instagram.com/olwettergren/

Obegränsad optimering, forts

$$\min_{x \in \mathbb{R}^n} f(x), \quad f \in C^2$$

Sökmetod:

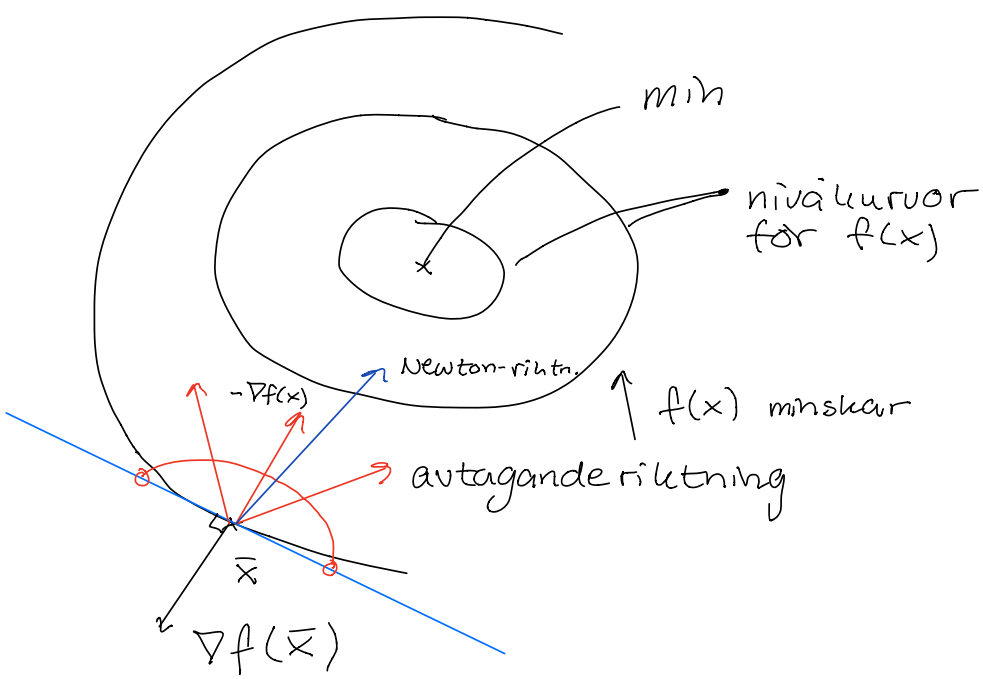


$$x^{k+1} = x^k + t_k d^k$$

Brantaste lutningsmetoden:

$$\min: \; d^k \,/\!/\, -\nabla f(x^k)$$
$$\max: \; d^k \,/\!/\, \nabla f(x^k)$$

$n = 2$



- Avtaganderiktning:

$$\nabla f(\bar{x})^T \bar{d} < 0$$

- Newtons metod:

$$d^k = -\nabla^2 f(x^k)^{-1} \nabla f(x^k)$$

för både min och max

Avtaganderiktning?

$$\nabla f(x^k)^T d^k = \left[-\nabla^2 f(x^k) d^k \right]^T d^k =$$
$$= -d^{kT} \nabla^2 f(x^k) d^k$$

- $\nabla^2 f(x^k)$ positivt definit $\Rightarrow$ avtaganderiktning, pekar mot min för $q(x)$.

- $\nabla^2 f(x^k)$ negativt definit $\Rightarrow$ ascentriktning, pekar mot max för $q(x)$.

- $\nabla^2 f(x^k)$ indefinit $\Rightarrow$ okänt, pekar mot sadelpunkt för $q(x)$.

Om vi har ett min-problem och

$$d^k = -\nabla^2 f(x^k)^{-1} \nabla f(x^k)$$

inte är en avtaganderiktning: ersätt riktningen med

$$d^k = -\left(\nabla^2 f(x^k) + \underbrace{\nu}_{\text{"my"}} I\right)^{-1} \nabla f(x^k)$$

blir en antaganderiktning om $\nu > 0$ är tillräckligt stor. Kallas Newton-Marquardt.

Begränsad icke-linjär optimering

$$\min f(x)$$
$$\text{då } x \in X \subset \mathbb{R}^n$$

Antag $f \in C^1$

Optimalitetsvillkor?

Två sätt: primalt eller dualt.

Primalt:

Givet $x \in X$ och ett $d \in \mathbb{R}^n$ (riktning)

Om

$$\nabla f(x)^T d < 0$$

så är

$$f(x + td) < f(x)$$

för små $t > 0$.

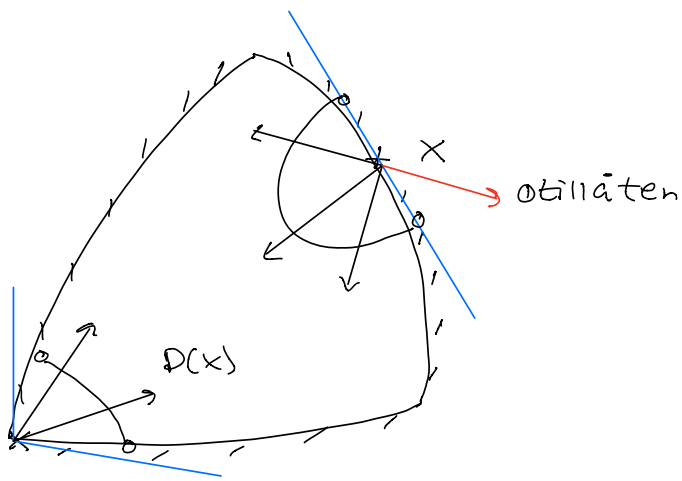
För vissa d och $t > 0$ kan $x + td \notin X$

Def:

Riktningen $d \neq 0$ är tillåten från $x \in X$ om $x + td \in X$ för små $t > 0$

Låt

$$D(x) = \text{mängden av tillåtna riktningar}$$



SATS:

$x^* \in X$ är ett lokalt minimum $\Rightarrow$

$$\nabla f(x^*)^T d \geq 0, \quad \forall d \in D(x^*)$$

Bevis:

$$p \Rightarrow q \Leftrightarrow \overset{\text{icke}}{\neg q} \Rightarrow \neg p$$

Antag att det finns

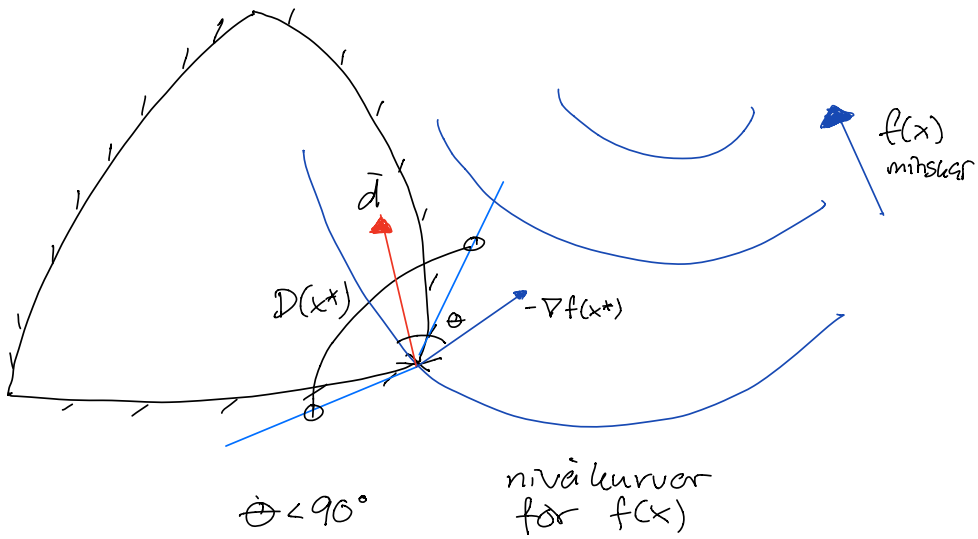
$$\bar{d} \in D(x^*) \text{ med } \nabla f(x^*)^T \bar{d} < 0.$$

1) $\bar{d} \in D(x^*) \Rightarrow x^* + t\bar{d} \in X$ för små $t > 0$

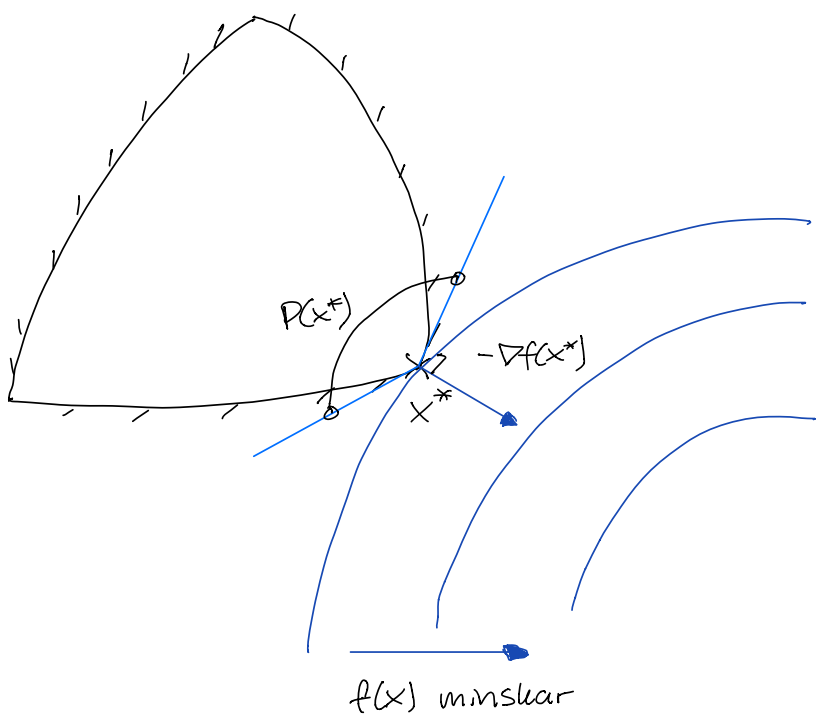
2) $\nabla f(x^*)^T \bar{d} < 0 \Rightarrow f(x^* + td) < f(x^*)$ för små $t > 0$

Alltså är x^* inte ett lokalt minimum ∎

$\bar{d}$ är en tillåten avtagande riktning



Finns tillåtna avtagande riktning $\Rightarrow$ x^* inte lokalt minimum.



Finns ingen tillåten avtagande riktning och x^* är ett lokalt minimum.

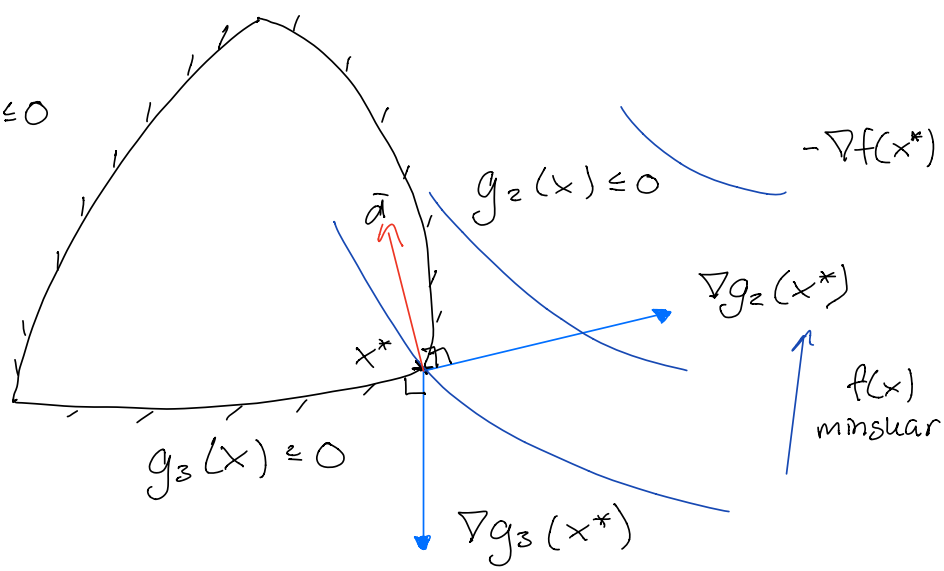
Dualt:

Låt

$$\bar{X} = \{x \mid g_i(x) \le 0, \quad i = 1, \dots, m\}$$

där alla $g_i \in C^1$.

$\bar{x}$ är inte ett lokalt minimum

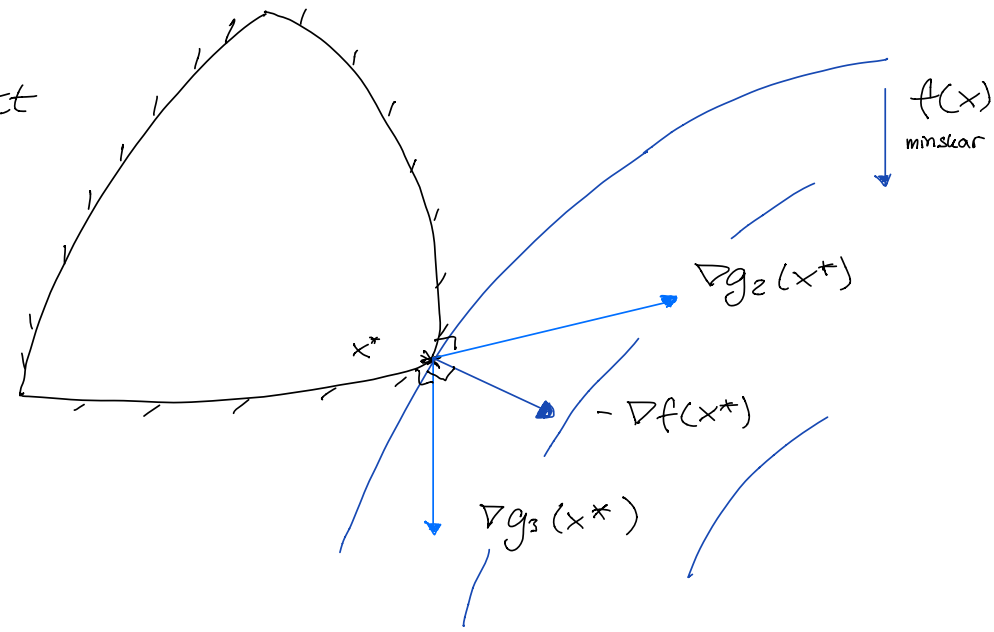


Villkoret $g_i(x) \le 0$ har utåtriktad normal $\nabla g_i(x)$

Obs: $g_2(x^*) = g_3(x^*) = 0$

$g_1(x^*) < 0$

$\bar{x}$ är ett lokalt minimum



Här gäller:

$$-\nabla f(x^*) = y_2^* \nabla g_2(x^*) + y_3^* \nabla g_3(x^*)$$

där

$$y_2^*, y_3^* > 0$$

$$\Leftrightarrow$$

$$\begin{cases} \nabla f(x^*) + y_1^* \nabla g_1(x^*) + y_2^* \nabla g_2(x^*) + \\ + y_3^* \nabla g_3(x^*) = 0 \\ y_1^*, y_2^*, y_3^* \geq 0 \\ y_i^* g_i(x^*) = 0, \quad i = 1, 2, 3 \end{cases}$$

OBS:

$$y_1^* \underbrace{g_1(x^*)}_{<0} = 0 \Rightarrow y_1^* = 0$$

$$y_2 \underbrace{g_2(x^*)}_{=0} = 0 \quad \text{ger inget}$$

$$y_3 \underbrace{g_3(x^*)}_{=0} = 0 \quad \text{ger inget}$$

SATS:

Antag att $f(x)$ och alla $g_i(x)$ är konvexa på X.
Då är x^* ett globalt minimum om det finns y_i^*, $i = 1, \dots, m$, sådana att

$$\begin{cases} \nabla f(x^*) + \sum_{i=1}^{m} y_i^* \nabla g_i(x^*) = 0 & (1) \\ y_i^* \geq 0, \ i = 1, \dots, m & (2) \\ g_i(x^*) \leq 0, \ i = 1, \dots, m & (3) \\ y_i^* g_i(x^*) = 0, \ i = 1, \dots, m & (4) \end{cases}$$

Villkoren (1) - (4) kallas
Karush-Kuhn-Tucker-villkor (KKT-villkor).

y_i, $i = 1, \dots, m$

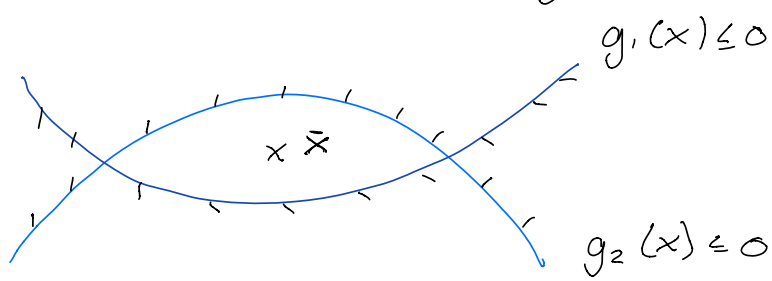
är dualvariabler.
En punkt som uppfyller (1)-(4) är en KKT-punkt.
• (1) och (2): dual tillåtenhet

- (3) primal tillåtenhet
- (4) komplementaritet

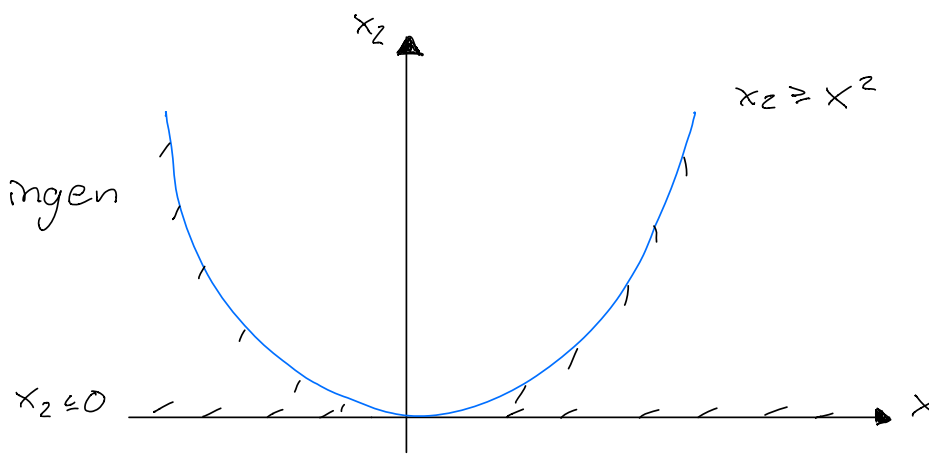
(1) – (4) är tillräckliga för global optimalitet.
Nödvändighet?
Behövs ett krav till

Def:

$\bar{x} \in X$ är en inre punkt om $g_i(\bar{x}) < 0$, $i = 1, \dots, m$.



Endast origo tillåten. Finns ingen inre punkt.



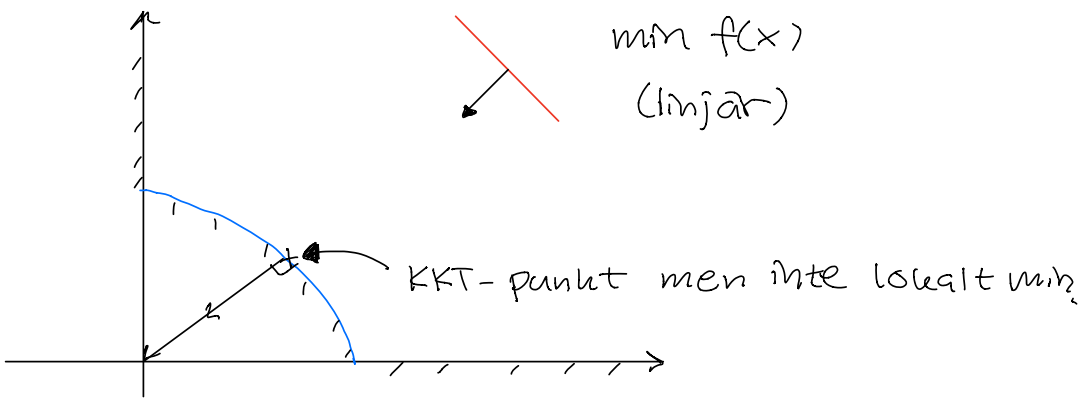
SATS:

För konvexa problem med inre punkt är KKT-villkoren nödvändiga och tillräckliga för global optimalitet.

För icke-konvexa problem:

Varning!

En KKT-punkt behöver inte ens vara ett lokalt minimum.



Straffunktionsmetoder

$$\text{(ILP)} \quad \begin{array}{l} \min f(x) \\ \text{då } g_i(x) \le 0, \quad i = 1, \dots, m \end{array}$$

Observation:

Låt

$$\sigma = \begin{cases} 0 & \text{om } g_i(x) \le 0 \quad (x \text{ tillåten}) \\ +\infty & \text{annars} \quad (x \text{ otillåten}) \end{cases}$$

sigma

Då är x^* optimal i ILP omm x^* är optimal i problemet

$$\min_{x \in \mathbb{R}^n} f(x) + \sigma$$

Alltså: ILP kan lösas som ett obegränsat problem.

Dock: $\sigma(x)$ oprakt iskt!

Approximera $\sigma(x)$ med något snällare!

Vanligast:

Yttre straffunktionen

$$\varphi(x) = \sum_{i=1}^{m} \left(\max\{0, g_i(x)\}\right)^2$$

Straffat problem:

$$\min_{x \in \mathbb{R}^n} \pi(x, \mu) = f(x) + \mu\, \varphi(x)$$

$$\mu \to \infty \Rightarrow \mu\, \varphi(x) \to \sigma(x)$$

$\min f(x)$

då $b \geq x \geq a$