Kap. 4: del III.

minimering uten betingelser

<u>Nødvendige</u> og tilstrekkelige betingelser for at x^* er minpunkt.

Teorem 14.1: Anta $f: \mathbb{R}^n \to \mathbb{R}$ kont. part. der., x^* lokalt min.

1. Da er $\nabla f(x^*) = 0$ | førsteordens bet.
2. Hvis f også har andreordens kont. part. der. | andreordens bet.
   Da er $\nabla^2 f(x^*)$ pos. semidefinit.

<u>Bevis</u>: 1. Anta x^* lokalt min og anta for motsigelse at $\nabla f(x^*) \neq 0$

Sett $\vec{h} = -\alpha \nabla f(x^*)$, $\alpha > 0$

$$\nabla f(x^*)^T \vec{h} = -\alpha \nabla f(x^*)^T \nabla f(x^*) = -\alpha \|\nabla f(x^*)\|^2 < 0$$

$\Rightarrow \nabla f(\vec{x})^T \vec{h} < 0$ for alle $\vec{x}$ tilstrekkelig nær x^*.

Førsteordens Taylor om x^*: Finnes en $t \in (0,1)$ s.a.

$f(x^*+h) - f(x^*) = \nabla f(x^*+th)\vec{h} < 0$ for $\vec{h}$ liten nok

$\Rightarrow f(x^*+h) < f(x^*)$, motsigelse siden x^* antas lokalt min.

2. andreordens Taylor:

$$\underline{f(x^*+h)} = f(x^*) + \underbrace{\nabla f(x^*)^T}_{0}\vec{h} + \tfrac{1}{2} h^T \nabla^2 f(x^*+th) h \quad (t \in (0,1))$$

$$= f(x^*) + \tfrac{1}{2} h^T \nabla^2 f(x^*+th) h$$

$$\leq \underline{f(x^*+h)} + \tfrac{1}{2} h^T \nabla^2 f(x^*+th) h$$

$\tfrac{1}{2} h^T \nabla^2 f(x^*+th) h \geq 0 \Rightarrow \nabla^2 f(x^*+th)$ pos semidefinit.

$\Rightarrow \nabla^2 f(x^*)$ pos semidefinit. (La $h \to 0$)

$f(x) = \tfrac{1}{2} x^T A x - b^T x$ A symmetrisk A invertibel kan kun være ett min,

førsteordens bet: $\nabla f(x) = Ax - b = 0$ $Ax = b$, $x = A^{-1} b$

andreordens bet: $\nabla^2 f(x) = A$, må være pos semidefinit.

Lemma 14.2 Anta $A^{n\times n}$ symmetrisk pos. def., La $\lambda_n > 0$ være minste egenverdi. Da er

$$h^T A h \geq \lambda_n \|h\|^2 \quad \text{for alle } h \in \mathbb{R}^n.$$

Bevis: A er ort. diagonaliserbar $\Rightarrow$ $A = PDP^T$, D diagonal, P ortogonal

$$h^T A h = h^T P D P^T h = (P^T h)^T D \underbrace{P^T h}_{y} = y^T D y = \sum_{i=1}^n \lambda_i y_i^2 \geq \sum_{i=1}^n \lambda_n y_i^2$$

$$= \lambda_n \sum_{i=1}^n y_i^2 = \lambda_n \|y\|^2$$

$$= \lambda_n \|P^T h\|^2 = \underline{\lambda_n \|h\|^2}$$

Teorem 14.3: Anta f kont andreordens part. der. i et område rundt x^*. Anta og at $Df(x^*) = 0$, $D^2 f(x^*)$ pos. def. Da er x^* lokalt min.

Bevis: Lemma 14.2 + andreordens Taylor:

$$f(x^*+h) = f(x^*) + \underbrace{Df(x^*)^T}_{0} h + \frac{1}{2} \underbrace{h^T \overbrace{D^2 f(x^*)}^{A} h}_{\geq \lambda_n \|h\|^2,\ \lambda_n \text{ minste egenverdi i } D^2 f(x^*)} + \varepsilon(h) \|h\|^2$$

$$\geq f(x^*) + \frac{1}{2}\lambda_n \|h\|^2 + \varepsilon(h)\|h\|^2$$

$$\frac{f(x^*+h) - f(x^*)}{\|h\|^2} \geq \frac{1}{2}\lambda_n + \varepsilon(h) \quad \Big| \text{ velger } h \text{ så liten at } |\varepsilon(h)| < \frac{1}{4}\lambda_n$$

$$\frac{f(x^*+h) - f(x^*)}{\|h\|^2} \geq \frac{1}{2}\lambda_n - \frac{1}{4}\lambda_n = \frac{1}{4}\lambda_n > 0 \Rightarrow f(x^*+h) > f(x^*) \Rightarrow x^* \text{ lokalt min}$$

Teorem 14.4 Anta $f: \mathbb{R}^n \to \mathbb{R}$ konveks

1. Hvis x^* er lokalt min, så er x^* et globalt min.
2. Hvis f også er deriverbar så gjelder at

$$f \text{ lokalt min} \Leftrightarrow Df(x^*) = 0$$

Bevis: 1. Anta x_1 lokalt min. Hvis x_1 ikke er globalt min så finnes en x_2 s.a. $f(x_2) < f(x_1)$

For $0 < \lambda < 1$ har vi da

$$f(\underbrace{(1-\lambda)x_1 + \lambda x_2}_{\text{kan velges vilkårlig nær } x_1 \text{ (velg } \lambda \text{ liten)}}) \le (1-\lambda)f(x_1) + \lambda f(x_2) < (1-\lambda)f(x_1) + \lambda f(x_1) = f(x_1)$$

så x_1 er <u>ikke</u> lokalt min, motsigelse $\Rightarrow$ x_1 er globalt min.

2. $\Rightarrow$ er allerede vist

$\Leftarrow$: Anta $Df(x^*) = 0$

Teorem 12.10: f konveks så er $f(x) \ge f(x^*) + \underbrace{Df(x^*)}_{0}{}^T(x - x^*) = f(x^*)$

$\Rightarrow x^*$ lokalt min.

14.2 Linjesøkmetoder:

Gitt tilnærming x_k til min for f.
Finn bedre tilnærming x_{k+1} ved først å finne en søkeretning d_k, så en steglengde $\alpha_k > 0$, og sett

$$x_{k+1} = x_k + \alpha_k d_k$$

Eksakt linjesøk betyr at α_k velges som den α som gir minst verdi for $f(x_k + \alpha d_k)$

Hovedmetoder for linjesøk:

1. Steepest descent: Velg $d_k = -\nabla f(x_k)$ → avtar raskest.

$x_{k+1} = x_k - \alpha_k \nabla f(x_k)$

Andreordens Taylor: $f(x+h) = \underbrace{f(x) + \underline{\nabla f(x)^T h}}_{\text{dominerer for små } \vec{h}} + \frac{1}{2} h^2 \nabla^2 f(x+th) h$

$\nabla f(x)^T h \geq -\|\nabla f(x)\| \|h\|$

Likhet når $\underbrace{h = -\alpha \nabla f(x)}_{\text{liten steglengde, gir mest mulig reduksjon i } f}$ ⇒ denne h gir mest reduksjon

2. Newton: $f(x_k + h) \approx \underbrace{f(x_k) + \nabla f(x_k)^T h + \frac{1}{2} h^T \nabla^2 f(x_k) h}$

La oss finne min for denne:

gradient: $\nabla f(x_k) + \nabla^2 f(x_k) \vec{h}$

Hessematrise: $\nabla^2 f(x_k)$

Teorem 14.3: hvis $\nabla^2 f(x_k)$ er pos. def. og gradient $= 0$ for et valg av h, d_k, så er punktet lokalt min.

$\nabla f(x_k) + \nabla^2 f(x_k) \vec{h} = 0$

$\vec{h} = -(\nabla^2 f(x_k))^{-1} \nabla f(x_k)$

$x_{k+1} = \underbrace{x_k - (\nabla^2 f(x_k))^{-1} \nabla f(x_k)}_{\text{fullt Newton steg, dvs. } \alpha_k = 1} = x_k + \alpha_k d_k$ med $\frac{\alpha_k = 1}{d_k = -(\nabla^2 f(x_k))^{-1} \nabla f(x_k)}$

Newton med steglengde: $x_{k+1} = \underline{x_k - \alpha_k (\nabla^2 f(x_k))^{-1} \nabla f(x_k)}$

Def 14.5 "Backtracking" linjesøk

Anta vi har valgt $\_\ s \leq 1$, $0 < \beta < 1$, $0 < \sigma < 1$

Vi setter

$$m_k = \min \{ m : m \geq 0 \mid \underbrace{f(x_k) - f(x_k + \beta^m s d_k) \geq -\sigma \beta^m s \nabla f(x_k)^T d_k}_{\text{stoppbetingelse}} \}$$

$$\alpha_k = \beta^{m_k} s$$

Teorem 14.8

Anta $f(x) = x^T A x$, A pos. def $(\min f(0) = 0)$

steepest descent med eksakt linjesøk gir at $\xrightarrow{}$ største egenverdi

$\underline{f(x_{k+1}) \leq m_A f(x_k)}$, der $m_A = \left(\dfrac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^2 < 1$ $\xrightarrow{}$ minste egenverdi

lineær konvergens mot 0

$$|f(x_{k+1}) - \tilde{f}(0)| \leq m_A |f(x_k) - \tilde{f}(0)|$$

Beviser ikke