Del III Ikkelineær optimering

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

f kalles objektivfunksjonen, denne skal minimeres.

x^* kalles et lokalt minimum hvis det finnes en $\varepsilon > 0$ s.a. $f(x^*) \leq f(x)$ for alle $\vec{x} \in \underbrace{B(x^*, \varepsilon)}$

alle punkter z s.a. $\|z - x^*\| < \varepsilon$

$\bar{B}(x^*, \varepsilon) : \{z \mid \|z - x^*\| \leq \varepsilon\}$

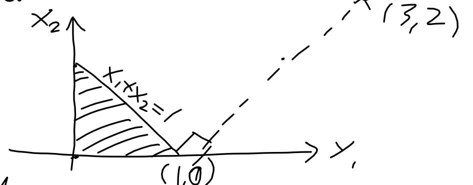
$\vec{x}^*$ er globalt minimum hvis $f(\vec{x}^*) \leq f(x)$ alle x.

(x er min for $f(x)$ $\Longleftrightarrow$ x er maks for $-f(x)$)

Skal senere ta med betingelser.

Eksempel der vi ikke trenger å regne for å finne min:

$$\left.\begin{array}{l} \min \; (x_1 - 3)^2 + (x_2 - 2)^2 \\ \quad x_1 + x_2 \leq 1 \\ \quad x_1 \geq 0, \; x_2 \geq 0 \end{array}\right\} \to \text{kvadratet av avstanden til } (3,2)$$



å minimere kvadratet av avstanden er det samme som å minimere avstanden (siden $f(x) = x^2$ er voksende på $[0, \infty)$)

$h(x) = (g(x))^2 \quad h'(x) = 2g(x)\, g'(x) = 0$ når $g'(x) = 0$

(↑ avstand)

$(1,0)$ er punktet med minst avstand, siden normalen fra $(3,2)$ ned på linjen $x_1 + x_2 = 1$ treffer linjen i $(1,0)$.

Teorem 11.1

Anta f kont. på en lukket og begrenset delmengde C. Da har f både maks. og min. på C, dvs. det finnes $x^{(1)}, x^{(2)} \in C$ s.a. $f(x^{(1)}) \leq f(x) \leq f(x^{(2)})$ for alle $x \in C$

11.2 Anvendelser

Portefølgeoptimering (målt i verdi)

x_i er her andelen av aksje i av en aksjeportefølje

$0 \le x_i \le 1 \quad \sum_{i=1}^{n} x_i = 1$

$$\boxed{\begin{array}{l} \min \ \alpha \sum_{1\le i,j\le n} c_{ij} x_i x_j - \sum_{j=1}^{n} \mu_j x_j \\ \qquad 0 \le x_i \le 1 \\ \qquad \sum x_i = 1 \end{array}}$$

fra statistikk: forventning. $\int_{-\infty}^{\infty} t R_j(t)\,dt$

R_j = avkastning av aksje j, og $\mu_j = \mathbb{E}(R_j)$ forventet avkastning av aksje j

forventet avkastning av hele porteføljen:

$$\mathbb{E}\left(\sum x_j R_j\right) = \sum_j x_j \mathbb{E}(R_j) = \sum_j \mu_j x_j$$

Vi vil finne maks av denne

$\Updownarrow$

finne min av $-\sum_j \mu_j x_j$

For enkel modell. La oss prøve å redusere risiko ved å spre investeringene

kovarians mellom R_i og R_j er definert ved

$$c_{ij} = \mathbb{E}\big((\underline{R_i - \mu_i})(R_j - \mu_j)\big)$$

kovarians til hele porteføljen: ($\sum x_i R_i$, $\mathbb{E}(\sum x_i R_i) = \sum \mu_i x_i$

$$\mathbb{E}\Big(\big(\underline{\sum x_i R_i} - \underline{\sum x_i \mu_i}\big)\big(\sum x_j R_j - \sum x_j \mu_j\big)\Big) = \dots \quad \mathbb{E}\Big(\big(\sum_i x_i (R_i - \mu_i)\big)\big(\sum x_j (R_j - \mu_j)\big)\Big)$$

$$= \sum x_i x_j \underbrace{\mathbb{E}\big((R_i - \mu_i)(R_j - \mu_j)\big)}_{c_{ij}} \qquad = \underline{\sum_{i,j} x_i x_j \mathbb{E}\big((R_i - \mu_i)(R_j - \mu_j)\big)}$$

$$= \sum c_{ij} x_i x_j$$

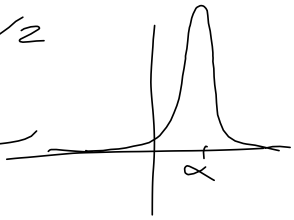
som er første del av objektivfunksjonen

maximum likelihood (ML)

X "kontinuerlig" random variabel, sannsynlighetstetthet $p(x;\alpha)$

Der α er en ukjent parameter.

X Gaussfordelt: $\underbrace{p(x;\alpha) = \frac{1}{\sqrt{2\pi}} e^{-(x-\alpha)^2/2}}$



α forventning

varians 1.

Vi vil finne α ved å ta utgangspunkt i "utfall" $x_1,\dots,x_n$

sannsynligheten for disse utfallene, (gitt at de er uavhengige), er

$$\prod_{i=1}^{n} p(x_i;\alpha) \quad \text{(kalles likelihood-funksjonen)}$$

ML finner et estimat for α ved å maksimere likelihood-funksjonen.

$$\max \prod_{i=1}^{n} p(x_i;\alpha) \iff \max \ln\left(\prod_{i=1}^{n} p(x_i;\alpha)\right) = \max \sum_{i=1}^{n} \ln(p(x_i;\alpha))$$

$$\min -\sum_{i=1}^{n} \ln(p(x_i;\alpha))$$

Eksempel: $p(x;\alpha) = \frac{1+\alpha x}{2}$, $x \in [-1,1]$.

objektivfunksjonen ~~blir~~ nå $f(\alpha) - \sum_{i=1}^{n} \ln\left(\frac{1+\alpha x_i}{2}\right)$

$$f'(\alpha) = -\sum_{i=1}^{n} \frac{\frac{x_i}{2}}{(1+\alpha x_i)/2} = -\sum_{i=1}^{n} \frac{x_i}{1+\alpha x_i}$$

$$f''(\alpha) = \sum_{i=1}^{n} \frac{x_i^2}{(1+\alpha x_i)^2} \geq 0$$

11.3 Flervariabel analyse. A kvadratisk.

A pos. semidefinit $\iff x^T A x \geq 0$ alle x

$\iff$ alle egenverdier ≥ 0

$\iff A = W^T W$ for en matrise W.

A pos. definit $\iff x^T A x > 0$ alle $x \neq 0$

$\iff$ alle egenverdier > 0

$\iff A = W^T W$ for en <u>invertebar</u> matrise W

Gradienten til en funksjon f i n variable (søylevektor)

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$$

Hessematrisen: $$\nabla^2 f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)$$ kont. 2.ordens part. der.

$\Downarrow$

symmetrisk

$$F(x) = \begin{bmatrix} F_1(x) \\ \vdots \\ F_n(x) \end{bmatrix}$$ vektorvalued funksjon med komponenter F_i

---

$$F'(x) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \\ \frac{\partial F_n}{\partial x_1} & \dots & \frac{\partial F_n}{\partial x_n} \end{pmatrix}$$

Taylorteoremer:

11.2 Førsteordens Taylor

$f: \mathbb{R}^n \to \mathbb{R}$ kont. part. der. i $B(x;r)$

For alle h i $\mathbb{R}^n$ med $\|h\| < r$, finnes en $t \in (0,1)$ s.a.

$$f(x+h) = f(x) + \underbrace{\nabla f(x+th)^T h}_{\text{restledd}}$$

11.3 Andreordens Taylor

$f: \mathbb{R}^n \to \mathbb{R}$ kont. part. 2. ordens der. i $B(x;r)$

For alle h.... finnes en $t \in (0,1)$ s.a

$$f(x+h) = f(x) + \nabla f(x)^T h + \underbrace{\frac{1}{2} h^T \nabla^2 f(x+th)\, h}_{\text{restledd}}$$

11.4 Andreordens Taylor, del 2

samme betingelser for f som i 11.3

Finnes funksjon $\varepsilon: \mathbb{R}^n \to \mathbb{R}$ s.a

$$f(x+h) = f(x) + \nabla f(x)^T h + \frac{1}{2} h^T \nabla^2 f(x)\, h + \varepsilon(h)\|h\|^2$$

der $\varepsilon(y) \to 0$ når $y \to 0$

11.5 Førsteordens Taylor for vektorvaluerte funksjoner

Anta $F: \mathbb{R}^n \to \mathbb{R}^m$ kont. deriverbar i omegn N om x. Da er

$$F(x+h) = F(x) + F'(x)h + \underbrace{O(\|h\|)}_{\text{går mot 0 som } \|h\|}$$

Kjerneregelen: $H'(x) = F'(G(x))\,G'(x)$ $\quad H(x) = F(G(x))$

Lineær konvergens av x_k til x^*

$$\|x_{k+1} - x^*\| \leq \gamma \|x_k - x^*\| \qquad (\gamma < 1)$$

Superlineær konvergens hvis i tillegg $\dfrac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0$

Kvadratisk konvergens $\|x_{k+1} - x^*\| \leq \gamma \|x_k - x^*\|^2 \qquad \gamma < 1.$