Kap. 2 ; del III

Konveksitet. Viktig begrep i optimering

Konveks mengde C : $\underbrace{(1-\lambda)\vec{x} + \lambda\vec{y}}_{\text{konvekskombinasjon}} \in C$ når $\vec{x}, \vec{y} \in C$ og $\lambda \in [0, 1]$

$B(a;\delta)$ konveks → alle punkter x s.a. $\|x - a\| < \delta$

Bevis: $\|(1-\lambda)\vec{x} + \lambda\vec{y} - \vec{a}\| = \|(1-\lambda)\vec{x} + \lambda\vec{y} - (1-\lambda)\vec{a} - \lambda\vec{a}\|$

$$= \|(1-\lambda)(\vec{x}-\vec{a}) + \lambda(\vec{y}-\vec{a})\|$$
$$\leq |1-\lambda|\,\|\vec{x}-\vec{a}\| + |\lambda|\,\|\vec{y}-\vec{a}\|$$
$$= (1-\lambda)\|\vec{x}-\vec{a}\| + \lambda\|\vec{y}-\vec{a}\|$$
$$< (1-\lambda)\delta + \lambda\delta = \delta$$

$\Rightarrow (1-\lambda)\vec{x} + \lambda\vec{y} \in B(\vec{a}, \delta) \Rightarrow B(\vec{a}, \delta)$ konveks.

Et lineært system: Et endelig system av lineære ligninger / ulikheter.

(feks $x_1 \geq 0$, $x_1 + x_2 = 1$)

Løsningene til et lineært system kalles et polyeder

prop. 12.1: Polyedre er konvekse

Viser bare med ulikheter.

Bevis: Kan beskrives ved $Ax \overset{\text{komponentvis}}{\leq} b$

Anta $A\vec{x} \leq \vec{b}$, $A\vec{y} \leq \vec{b}$

$$A((1-\lambda)\vec{x} + \lambda\vec{y}) = (1-\lambda)\underbrace{A\vec{x}}_{\leq \vec{b}} + \lambda \underbrace{A\vec{y}}_{\leq \vec{b}} \leq (1-\lambda)\vec{b} + \lambda\vec{b} = \vec{b}$$

$\Rightarrow$ $(1-\lambda)\vec{x} + \lambda\vec{y}$ ligger i polyedret også, som da er konvekst.

prop 12.2 Anta C er konveks og T er en lineær transformasjon.
Da er også $T(C)$ konveks.

Bevis: $(1-\lambda)T(\vec{x}) + \lambda T(\vec{y}) = T((1-\lambda)\vec{x} + \lambda\vec{y}) \in T(C)$

siden $(1-\lambda)\vec{x} + \lambda\vec{y} \in C$ (Konveks)

12.2 En funksjon f definert på en konveks mengde C kalles konveks hvis
$$f((1-\lambda)\vec{x} + \lambda\vec{y}) \leq (1-\lambda)f(\vec{x}) + \lambda f(\vec{y}) \quad \text{alle } \lambda \in [0,1]$$

f lineær $\Rightarrow$ f konveks (likhet over)

en norm er konveks:
$$\|(1-\lambda)\vec{x} + \lambda\vec{y}\| \leq (1-\lambda)\|\vec{x}\| + \lambda\|\vec{y}\|$$

f affin $\Rightarrow$ f konveks (likhet)

radvektor, konstant

på formen $f(x) = \vec{A}\vec{x} + \alpha$

$$f((1-\lambda)\vec{x} + \lambda\vec{y}) = A((1-\lambda)\vec{x} + \lambda\vec{y}) + \alpha = A((1-\lambda)\vec{x} + \lambda\vec{y}) + (1-\lambda)\alpha + \lambda\alpha$$
$$= (1-\lambda)(A\vec{x} + \alpha) + \lambda(A\vec{y} + \alpha)$$
$$= (1-\lambda)f(\vec{x}) + \lambda f(\vec{y}).$$

---

maks. av konvekse funksjoner f og g er også konveks:
$$f(\lambda\vec{x} + (1-\lambda)\vec{y}) \leq \lambda f(\vec{x}) + (1-\lambda)f(\vec{y}) \leq \lambda \max\{f,g\}(\vec{x}) + (1-\lambda)\max\{f,g\}(\vec{y})$$
$$g(\lambda\vec{x} + (1-\lambda)\vec{y}) \leq \lambda g(\vec{x}) + (1-\lambda)g(\vec{y}) \leq \lambda \max\{f,g\}(\vec{x}) + (1-\lambda)\max\{f,g\}(\vec{y})$$

$$\Rightarrow \max\{f,g\}(\lambda\vec{x} + (1-\lambda)\vec{y}) \leq \lambda \max\{f,g\}(\vec{x}) + (1-\lambda)\max\{f,g\}(\vec{y})$$

$\Rightarrow$ $\max\{f,g\}$ konveks.

Anta at h er konveks. Da er $e^{h(x)}$ også konveks.

$$h((1-\lambda)\vec{x}+\lambda\vec{y}) \le (1-\lambda)h(\vec{x})+\lambda h(\vec{y})$$

e^x er voksende $\Downarrow$

$$e^{h((1-\lambda)\vec{x}+\lambda\vec{y})} \le e^{(1-\lambda)h(\vec{x})+\lambda h(\vec{y})}.$$

$$e^{(1-\lambda)\underline{h(\vec{x})}+\lambda\underline{h(\vec{y})}} \le (1-\lambda)e^{h(\vec{x})}+\lambda e^{h(\vec{y})}$$

e^x er konveks:

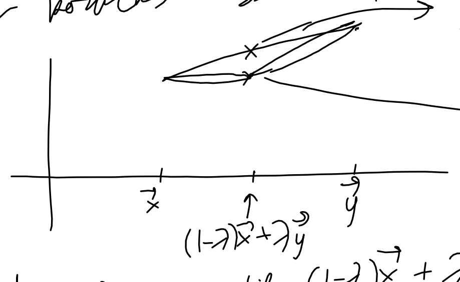
siden $e^{h((1-\lambda)\vec{x}+\lambda\vec{y})} \le (1-\lambda)e^{h(\vec{x})}+\lambda e^{h(\vec{y})}$

så er $e^{h(\vec{x})}$ konveks.

e^x er konveks siden $f''(x)\ge 0$ alle x

$\rightarrow$ linjestykket fra $f(x)$ til $f(\vec{y})$

$$(1-\lambda)f(\vec{x})+\lambda f(\vec{y})$$
$$\text{V/}$$
$$\rightarrow f((1-\lambda)\vec{x}+\lambda\vec{y})$$



sekanten fra x til $(1-\lambda)\vec{x}+\lambda\vec{y}$ har mindre stigningstall enn sekanten fra $(1-\lambda)\vec{x}+\lambda\vec{y}$ til $\vec{y}$. $\Rightarrow$ den deriverte er voksende.
$\Rightarrow$ andrederiverte ≥ 0

prop. 12.3 $f(H(\vec{x}))$ e konveks når f er konveks og H affin

Bevis: $f\Big(H\big((1-\lambda)\vec{x} + \lambda\vec{y}\big)\Big) = f\big((1-\lambda)H(\vec{x}) + \lambda H(\vec{y})\big)$

$\leq (1-\lambda)f(H(\vec{x})) + \lambda f(H(\vec{y}))$

$\Rightarrow f(H(\vec{x}))$ er konveks.

Teorem 12.4 Jensens ulikhet

Vi har at, hvis $x^{(1)}, \dots, x^{(n)}$ er i C, og $\lambda_1, \dots, \lambda_n$ s.a. $\lambda_i \geq 0$, $\sum \lambda_i = 1$

så er $f\Big(\underbrace{\sum_{i=1}^{n} \lambda_i x^{(i)}}_{\text{kalles også for en konvekskombinasjon.}}\Big) \leq \sum_{i=1}^{n} \lambda_i f(x^{(i)})$

12.3 egenskaper

Teorem 12.7 Anta: $f: \underset{\text{konveks}}{C} \to \mathbb{R}$ konveks, og C åpen

Da er f kontinuerlig på C.

Teorem 12.8 $f: C \to \mathbb{R}$ def på C, åpen og konveks.

anta f har andreordens kont. part. der

f konveks $\Leftrightarrow$ $\nabla^2 f(x)$ pos semidefinit

Kvadratiske funksjoner: $f(x) = \frac{1}{2}\vec{x}^T A\vec{x} - \vec{b}^T\vec{x} = \sum_{i,j} \frac{1}{2} x_i a_{ij} x_j - \sum b_i x_i$

Oppgave 11.9 i boka sier at $\nabla f(\vec{x}) = A\vec{x} - \vec{b}$ (når A er symmetrisk)

$$\nabla^2 f(\vec{x}) = A$$

Ser derfor: En kvadratisk funksjon er konveks

$$\Updownarrow$$

$\nabla^2 f(x)$ er positiv semidefinit

Teorem 12.9 f konveks, def. på C, åpen, konveks

Hvis de partielle deriverte til f eksisterer, så er f deriverbar i $\vec{x}$.

Teorem 12.10

Anta f deriverbar, def. på C, åpen, konveks

Følgende er ekvivalent:

1. f er konveks
2. $f(x) \geq f(x_0) + Df(x_0)^T(x-x_0)$, alle $x, x_0 \in C$
3. $(Df(x) - Df(x_0))^T(x-x_0) \geq 0$, alle $x, x_0 \in C$.

Bevis (ii) $\Rightarrow$ (iii)

Legg sammen $\begin{cases} f(x) \geq f(x_0) + Df(x_0)^T(x-x_0) \\ f(x_0) \geq f(x) + Df(x)^T(x_0-x) \end{cases}$

$\cancel{f(x)} + \cancel{f(x_0)} \geq \cancel{f(x_0)} + \cancel{f(x)} + (Df(x_0)^T - Df(x)^T)(x-x_0)$

$\Rightarrow (Df(x_0)^T - Df(x)^T)(x-x_0) \leq 0$ (gang med -1)

Bevis for $n=1$: (i) $\Rightarrow$ (ii); Antar altså at f er konveks.

$f(x_0 + t(x-x_0)) = f((1-t)x_0 + tx) \leq (1-t)f(x_0) + t f(x)$

$\Downarrow$ flytter over, deler med t på begge sider:

$$f(x) \geq f(x_0) + \frac{f(x_0 + t(x-x_0)) - f(x_0)}{t}$$

$$= f(x_0) + \frac{f(x_0 + t(x-x_0)) - f(x_0)}{\underbrace{t(x-x_0)}_{h}}(x-x_0)$$

La $t \to 0$: $f(x) \geq f(x_0) + f'(x_0)(x-x_0) \Rightarrow f$ ligger over 1. ordens Taylor.

$n=1$: (iii) $\Rightarrow$ (i)

(iii): sier at den deriverte er voksende

La $x_1 \leq y_2 \leq x_3$

Bruker middelverdisetningen: Finnes c, d, s.a. $x_1 \leq c \leq y_2 \leq d \leq y_3$

s.a. $\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c) \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2} = f'(d)$

$$\Rightarrow \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}$$

$\Rightarrow$ stigningstallet til sekanten er voksende $\Rightarrow$ konveks.