Konvekse mængder

Erik Christensen

Indholdsfortegnelse

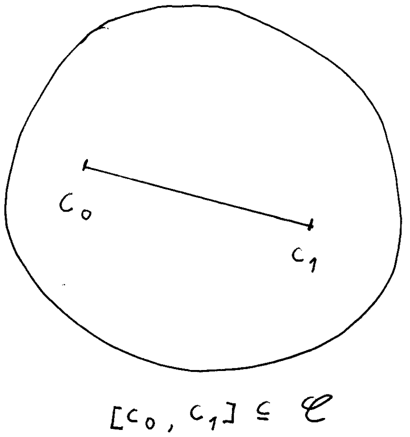
Version 2004: Denne version er ændret i forhold til 2003-udgaven efter mange gode forslag fra Christian Berg, som hermed takkes for sin grundige indsats med at læse noterne. Specielt kapitel 12 er blevet forbedret således at resultaterne nu fremstår i deres stærkeste versioner.

Erik Christensen
echris@math.ku.dk

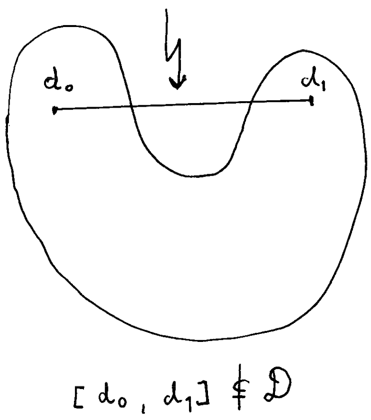
0 Elementære definitioner og det fundamentale resultat

Begrebet konveksitet fremgår forhåbentligt af nedenstående tegninger:

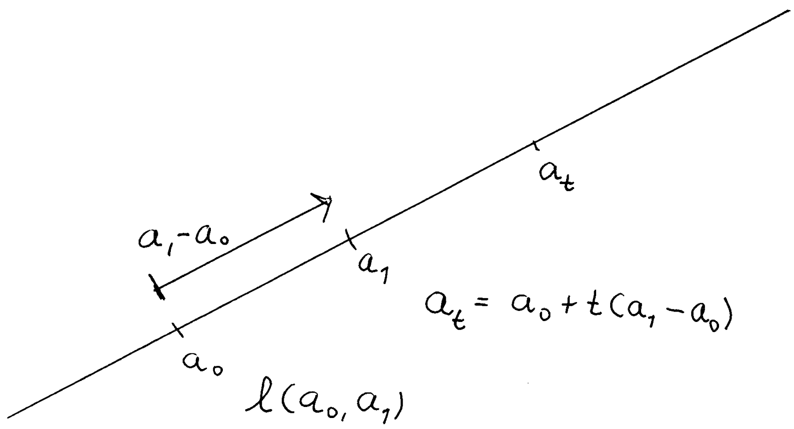
C konveks



D Ej konveks



Det ses at en delmængde af planen ($\mathbb{R}^2$) er konveks hvis forbindelseslinien mellem 2 vilkårlige punkter fra C er indeholdt i C. For at kunne generalisere dette til delmængder af $\mathbb{R}^n$ skal vi først gøre os klart hvad vi skal forstå ved et liniestykke mellem to punkter a_0, a_1 fra $\mathbb{R}^n$.



Tegningen ovenover er forhåbentlig kendt i forbindelse med den såkaldte parameterfremstiling af en linie i planen gennem punktet a_0 og med retningsvektoren $a_1 - a_0$.

Tegningen kunne naturligvis lige så gerne illustrere 2 punkter i $\mathbb{R}^n$.

Definition 0.1. Lad a_0, a_1 være punkter i $\mathbb{R}^n$, da defineres linien $\ell(a_0, a_1)$ gennem disse punkter ved

$$\ell(a_0, a_1) = \{a_0 + t(a_1 - a_0) | t \in \mathbb{R}\}.$$

Liniestykket mellem a_0 og a_1 betegnes $[a_0, a_1]$ og det er givet ved

$$[a_0, a_1] = \{a_0 + t(a_1 - a_0) | t \in [0, 1]\}.$$

Vi kan nu formulere definitionen konveksitet af delmængder af $\mathbb{R}^n$ eksakt.

Definition 0.2. Lad C være en delmængde af $\mathbb{R}^n$, da siges C at være konveks såfremt der gælder

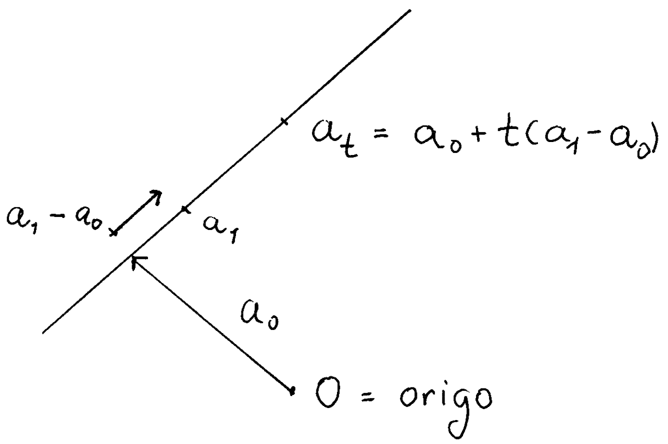
$$\forall a_0, a_1 \in C : [a_0, a_1] \subseteq C.$$

Vigtig bemærkning

$$\boxed{\text{Udtrykket } a_0 + t(a_1 - a_0) \text{ kan omskrives til } (1-t)a_0 + ta_1.}$$

Denne tilsyneladende uskyldige manipulation er faktisk en fundamental bestanddel af de følgende sider. Det første udtryk kalder på en geometrisk forståelse, medens det andet forstås bedre algebraisk.

Ved det første udtryk er det meningen at man skal forestille sig tegningen herunder, og at man får fremstillet linien $\ell(a_0, a_1)$ ved først at gå fra origo til a_0 for derefter at addere alle multipla af retningsvektoren $a_1 - a_0$.



Det andet udtryk kan algebraisk skrives som

$$\ell(a_0, a_1) = \{\alpha a_0 + \beta a_1 | \alpha, \beta \in \mathbb{R} \;\; \alpha + \beta = 1\}.$$

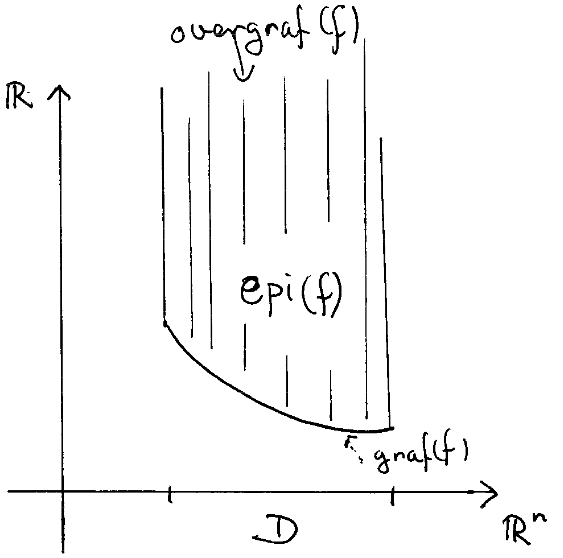
Når vi beskriver linien på denne måde er geometrien ikke synlig og vi vil senere lære at sige, at $\ell(a_0, a_1)$ er et affint underrum af $\mathbb{R}^n$ udspændt af punkterne a_0, a_1.

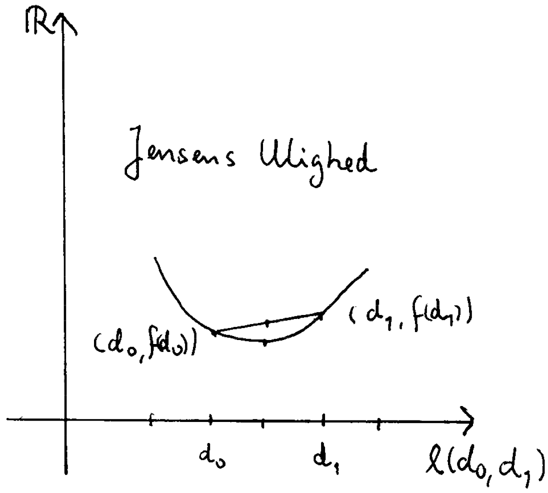
Vi skal nu lære begrebet "reel konveks funktion". Det baserer sig umiddelbart på begrebet konveks mængde, idet der til en reel funktion defineret på en delmængde af $\mathbb{R}^n$ naturligt knytter sig en delmængde af $\mathbb{R}^{n+1}$ kaldet overgrafen for f.

Definition 0.3. Lad D være en konveks delmængde af $\mathbb{R}^n$ og $f : D \to \mathbb{R}$ en funktion. Ved overgrafen eller epigrafen, $\mathrm{epi}(f)$, forstås punktmængden

$$\mathrm{epi}(f) = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} | x \in D\,,\, t \in \mathbb{R}\,,\, f(x) \leq t\}\,.$$

Funktionen f siges at være konveks hvis $\mathrm{epi}(f)$ *er konveks.*





Vigtig opgave. Simpleste version af Jensens Ulighed

Lad D være en konveks delmængde af $\mathbb{R}^n$ og $f : D \to \mathbb{R}$ en funktion. Lad d_0 og d_1 være punkter i D.

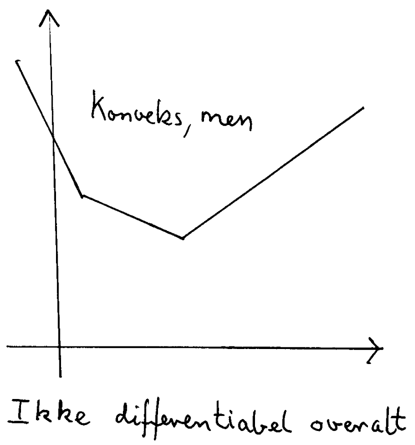
1. Beskriv både geometrisk og algebraisk liniestykket i $\mathbb{R}^{n+1}$ givet som $[(d_0, f(d_0)), (d_1, f(d_1))]$.

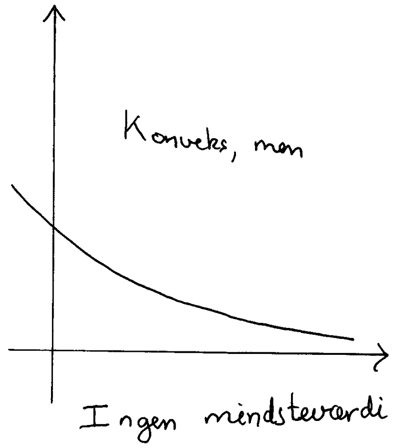
2. Vis, at f er konveks hvis og kun hvis der gælder

$$\forall d_0, d_1 \in D \forall t \in [0,1] : f((1-t)d_0 + td_1) \leq (1-t)f(d_0) + tf(d_1)\,.$$

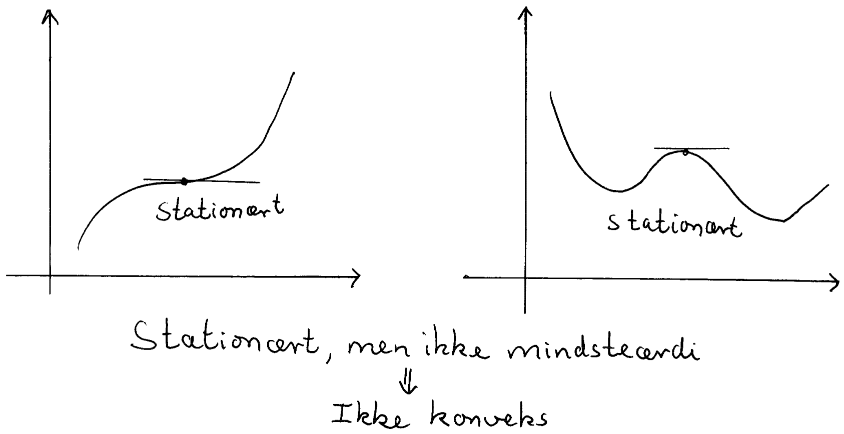
Denne sidste ulighed kaldes ofte for Jensens Ulighed, men er egentlig et specialtilfælde af denne, som vi møder i Sætning 5.11.

Når vi tegner grafen for en konveks funktion ser den typisk ud som et af billederne herunder: Konvekse funktioner er kontinuerte men ikke nødvendigvis differentiable på det indre af deres definitionsområder. En konveks funktion kan dog være diskontinuert på randen af sit definitionsområde.





De næste tegninger tjener til at illustrere, at hvis en funktion har et stationært punkt som ikke er mindsteværdipunkt, så er funktionen ikke konveks.



Mere positivt kan det formuleres, at hvis en konveks funktion har et stationært punkt har den også en mindsteværdi. Det er en meget vigtig pointe i dette kursus at dette resultat kan bevises eksakt og uden differentiabilitetsforudsætninger. Her studser du måske lidt, thi hvordan kan man tale om at gradienten er 0 hvis funktionen ikke er differentiabel ?. Dette problem klares senere, men allerede nu kan vi imidlertid vise det fundamentale resultat for differentiable funktioner definerede på åbne delmængder af $\mathbb{R}^n$.

Sætning 0.4. *Lad $D \subseteq \mathbb{R}^n$ være en konveks og åben mængde og lad $f : D \to \mathbb{R}$ være en differentiabel og konveks funktion. Hvis det for $d_0 \in D$ gælder at* $\operatorname{grad} f(d_0) = 0$ *så har f en mindsteværdi og denne er $f(d_0)$.*

Bevis. Beviset føres indirekte og er baseret på den fra Matematik 1 kendte regel for differentiation af en sammensæt funktion kaldet **kædereglen**. □

Lad os nu antage at der findes et par af punkter $d_0, d_1 \in D$ så $\operatorname{grad} f(d_0) = 0$ og $f(d_1) < f(d_0)$. Lad os nu betragte linien $\ell(d_0, d_1)$. Da D er åben og konveks, må der findes $\alpha < 0$ og $\beta > 1$ således, at $\forall t \in]\alpha, \beta[: (1-t)d_0 + td_1 \in D$. Da f er differentiabel er den sammensatte funktion $g(t) := f((1-t)d_0 + td_1)$ differentiabel i 0, og da gradienten

$\mathrm{grad} f(d_0) = 0$ fås ved kædereglen, at $g'(0) = \mathrm{grad} f(d_0) \bullet (d_1 - d_0) = 0$. Fra Jensens Ulighed fås da:

$$\begin{array}{llll}
\forall t \in]0,1[: & & & \\
f((1-t)d_0 + td_1) & & \leq (1-t)f(d_0) + tf(d_1) & \\
\text{så} & g(t) & \leq (1-t)g(0) & + tg(1) \\
\text{da} \quad t > 0 & \dfrac{g(t) - g(0)}{t} & \leq g(1) - g(0) & = f(d_1) - f(d_0) < 0\,.
\end{array}$$

Heraf ses at $g'(0) < 0$, så antagelsen $f(d_1) < f(d_0)$ er ikke sand, og $f(d_0)$ må være en mindsteværdi for f.

1 Repetition af Mat 1-stof. Afstandsmålet i $\mathbb{R}^n$ og lineære afbildninger

Formålet med dette afsnit er at skabe en basis for forståelsen af bgrebet affin afbildning, som er et meget væsentligt emne her i kurset. Det primære resultat vi skal bruge er at lineære afbildninger fra $\mathbb{R}^n$ ind i $\mathbb{R}^m$ er kontinuerte og givne ved $m \times n$–matricer. Kontinuiteten følger fra Cauchy–Schwarz' Ulighed, og sammenhængen mellem matricer og lineære afbildninger er et meget vigtigt resultatet fra Mat 1 GA. De ting der nævnes her er så fundamentale, og bruges så tit, at disse emner ikke blot er noget man skal have hørt om, men snarere ting man kan anvende umiddelbart, når de dukker op.

Det efterfølgende er ikke er et komplet repetitionskursus, så vil tillade os at forudsætte nedenstående begreber bekendt:

Injektiv, Surjektiv, Bijektiv, Underrum, Span, Lineær Uafhængighed, Basis, Matrix, Lineær Afbildning, Determinant, Egenværdi, Egenvektor, Kontinuitet, Norm ($:= \|x\| = (x_1^2 + \cdots + x_n^2)^{1/2}$).

1.1 Definitioner, konventioner og kommentarer

0) $:=$, Dynamisk lighedstegn, dvs. $A := B$ mener A defineres $= B$.

1) Lad $n \in \mathbb{N}$, $\mathbb{R}^n := \{(x_1, \ldots, x_n) | x_i \in \mathbb{R}\,, 1 \leq i \leq n\}$.

2) $L(\mathbb{R}^n, \mathbb{R}^m) :=$ mængden af lineære afbildninger fra

$$\mathbb{R}^n \text{ ind i } \mathbb{R}^m\,.$$

Bemærk en afbildning $F : \mathbb{R}^n \to \mathbb{R}^m$ er lineær hvis der gælder

$$\begin{aligned} &L1\ \forall x, y \in \mathbb{R}^n : F(x+y) = F(x) + F(y) \\ &L2\ \forall \alpha \in \mathbb{R}\ \forall x \in \mathbb{R}^n : F(\alpha x) = \alpha F(x)\,. \end{aligned}$$

Bemærk endvidere, at $L(\mathbb{R}^n, \mathbb{R}^m)$ er et vektorrum idet det for $F, G \in L(\mathbb{R}^n, \mathbb{R}^m)$ og $\alpha \in \mathbb{R}$ defineres

$$\begin{aligned} \forall x \in \mathbb{R}^n : (F+G)(x) &:= F(x) + G(x) \\ \forall x \in \mathbb{R}^n : (\alpha \cdot F)(x) &:= \alpha \cdot F(x)\ (= \alpha F(x))\,. \end{aligned}$$

(I det sidste udtryk $\alpha F(x)$ kan man egentlig ikke se hvad der menes, men det er altså også ligegyldigt).

3) $\mathbf{M}(m,n) :=$ mængden af $m \times n$ matricer med reelle elementer.

   Bemærk at $\mathbf{M}(m,n)$ er, som bekendt, også et vektorrum. For $A \in \mathbf{M}(m,n)$ og $B \in \mathbf{M}(n,k)$ defineres $AB \in \mathbf{M}(m,k)$ ved matrixproduktet

$$(AB)_{ij} = \sum_{t=1}^{n} a_{it} b_{tj} \,.$$

4) Standardbasen for $\mathbb{R}^n$ betegnes $e_1, \dots, e_n$, dvs.

$$e_1 = (1, 0, \dots, 0), \dots, e_j = (0 \dots 0, 1, 0 \dots 0), \dots, e_n = (0, \dots, 0, 1) \,.$$

5) For $x, y \in \mathbb{R}^n$ betegnes det indre produkt med både $x \cdot y$ og $\langle x, y \rangle$, dvs.

$$\begin{aligned} \langle x, y \rangle &= x \cdot y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \,, \\ \|x\| &:= (x_1^2 + \cdots + x_n^2)^{1/2} = (x \cdot x)^{1/2} = (\langle x, x \rangle)^{1/2} \,. \end{aligned}$$

6) $\mathbf{M}(n,1) := \mathbf{M}(1,n) := \mathbb{R}^n$.

   Pas på; naturligvis er rummet af søjlematricer med n rækker ikke det samme som rummet af rækkematricer med n søjler, men vi er alle klar over at forskellene mellem nedenstaående 3 udtryk primært består i opskrivningsmåden, og ikke i indholdet. Endvidere vil $x + y$ og αx have analoge betydninger uanset skriveformen. For $x \in \mathbb{R}^n$ bruges derfor følgende 3 skrivemåder: $x \in \mathbb{R}^n$, $x = (x_1, \dots, x_n)$

$$\begin{aligned} x &\in \mathbb{R}^n \,, & x &= (x_1, \dots, x_n) \\ x &\in \mathbf{M}(n,1) \,, & x &= \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ x &\in \mathbf{M}(1,n) \,, & x &= (x_1 \dots x_n) \,. \end{aligned}$$

   Hvis det skal understreges at x skal opfattes som henholdsvis en søjle eller en række, skrives

$$x_| := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad {}_-x := (x_1 \dots x_n) \,.$$

   For $A \in \mathbf{M}(m,n)$ skrives altså uden videre Ax selvom der naturligvis menes $Ax_|$. Analogt for $y \in \mathbb{R}^m$ skrives yA i betydningen ${}_-yA$.

   Bemærk endeligt, at med denne notation kan det indre produkt $x \cdot y$ for $x, y \in \mathbb{R}^n$ skrives

$$x \cdot y = {}_-x y_| \,.$$

7) Lad $A \in \mathbf{M}(m,n)$, $i \in \{1,\dots,m\}$, $j \in \{1,\dots,n\}$. Den i'te række i A betegnes ${}_iA$ og den j'te søjle A_j. I henhold til konventionerne fra 6) fås så

$$
\begin{aligned}
{}_iA &= \begin{pmatrix} a_{i1} & a_{i2} & & a_{in} \end{pmatrix} = \left(a_{i_1}, a_{i_2}, \dots, a_{i_n}\right) \in \mathbb{R}^n \\
A_j &= \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = \left(a_{1j}, a_{2j}, \dots, a_{mj}\right) \in \mathbb{R}^m .
\end{aligned}
$$

8) Lad $A \in \mathbf{M}(m,n)$, da defineres $\|A\|$ som den ville være hvis vi opfattede A som element af $\mathbb{R}^{mn}$, dvs.

$$
\|A\| = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2} .
$$

Med konventionerne fra 7) fås så

$$
\|A\|^2 = \sum_{i=1}^{m} \|{}_iA\|^2 = \sum_{j=1}^{n} \|A_j\|^2 .
$$

Det næste resultat, som kaldes en hovedsætning, bør man have som paratviden i ens matematiske univers!

1.2 Hovedsætning

Lad $m, n \in \mathbb{N}$.

Der defineres en bijektiv afbildning $\varphi_{m,n}$ af $\mathbf{M}(m,n)$ på $L(\mathbb{R}^n, \mathbb{R}^m)$ ved:

$$
\begin{aligned}
&\forall A \in \mathbf{M}(m,n) \quad \forall x \in \mathbb{R}^n : \\
&\varphi_{m,n}(A)(x) := Ax = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} .
\end{aligned}
$$

1.3 Kommentarer og udvidelser til hovedsætningen

Vigtig

1) Hvis $F \in L(\mathbb{R}^n, \mathbb{R}^m)$, findes $A := \varphi_{m,n}^{-1}(F)$ ved $A_j := F(e_j)$ "j'te søjle er billedet af e_j".

2) Ifølge 1.1 2) + 1.1 3) er $L(\mathbb{R}^n, \mathbb{R}^m)$ og $\mathbf{M}(m,n)$ vektorrum. Med hensyn til disse vektorrumsstrukturer er $\varphi_{m,n}$ en lineær og bijektiv afbildning af $\mathbf{M}(m,n)$ på $L(\mathbb{R}^n, \mathbb{R}^m)$.

   Opgave: Vis dette! Husk at *på*, betyder surjektiv, men φ er altså også injektiv!

3) For afbildninger $F : \mathbb{R}^n \to \mathbb{R}^m$, $G : \mathbb{R}^k \to \mathbb{R}^n$ defineres $F \circ G$ som den sammensatte afbildning, dvs. for $z \in \mathbb{R}^k$ er $F \circ G(z) := F(G(z))$. For matricer $A \in \mathbf{M}(m,n)$, $B \in \mathbf{M}(n,k)$ betyder AB blot matrixproduktet. Med disse notationer gælder der:

$$\varphi_{m,k}(AB) = \varphi_{m,n}(A) \circ \varphi_{n,k}(B)\,.$$

Opgave: Vis dette!

4) For $m = n$ defineres $I \in \mathbf{M}(m,m)$ ved $I_{ij} = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$

Der gælder $\varphi_{m,m}(I) = \mathrm{id}_{\mathbb{R}^m}$, hvor dette udtryk betegner den identitetsafbildnigen:

$$\mathrm{id}_{\mathbb{R}^m}(x) = x \text{ for alle } x \in \mathbb{R}^m\,.$$

Opgave: Vis dette!

Endelig får 3) en speciel pæn form når $m = m$. Idet vi så kan slutte, at $\varphi_{m,m}$ er en lineær bijektiv afbildning af $\mathbf{M}(m,m)$ på $L(\mathbb{R}^m, \mathbb{R}^m)$ som tillige opfylder $\varphi_{m,m}(AB) = \varphi_{m,m}(A) \circ \varphi_{m,m}(B)$.

5) Bemærk, at for $A \in \mathbf{M}(m,n)$ og $x \in \mathbb{R}^n$ er

$$\begin{aligned}
\varphi_{m,n}(A)x &= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} \\
&= x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \\
&= x_1 A_1 + \cdots x_n A_n \\
&= \sum_{j=1}^{n} x_j A_j\,.
\end{aligned}$$

Dvs. billedrummet $\varphi_{mn}(A)(\mathbb{R}^n)$ er givet ved

$$\varphi_{mn}(A)(\mathbb{R}^n) = \text{ span } \{A_1, \ldots, A_n\}\,.$$

"Billedrummet er det lineære span af søjlerne".

6) Vi kigger nu på specialtilfældet $m = 1$.

Via identifikationen $\mathbb{R}^n = \mathbf{M}(1,n)$ giver hovedsætningen en lineær bijektiv afbildning $\varphi_{1,n}$ af $\mathbb{R}^n$ på $L(\mathbb{R}^n, \mathbb{R})$. Det ses at for en vektor $y \in \mathbb{R}^n$ bliver $\varphi_{1,n}(y) \in L(\mathbb{R}^n, \mathbb{R})$ givet ved

$$\forall x \in \mathbb{R}^n : \varphi_{1,n}(y)(x) = (y_1 \ldots y_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} y_i x_i = \langle y, x \rangle\,.$$

Heraf fås altså at til enhver lineær afbildning $f : \mathbb{R}^n \to \mathbb{R}$ findes netop et $y \in \mathbb{R}^n$ så f er givet ved:

$$\forall x \in \mathbb{R}^n : f(x) = \langle y, x \rangle .$$

Ovenstående identitet kan skrives $f = \langle y, \cdot \rangle$.

Omvendt vil vi, når $y \in \mathbb{R}^n$ er givet, lade f_y betegne den lineære afbildning af $\mathbb{R}^n$ ind i $\mathbb{R}$ givet ved $f_y(x) :=< y, x >$.

Generelt for et vektorrum E kalder man vektorrummet $L(E, \mathbb{R})$ af lineære afbildninger fra E ind i $\mathbb{R}$ for E's duale rum, og dette betegnes E^*. Ovenstående viser, at $(\mathbb{R}^n)^*$ kan identificeres med $\mathbb{R}^n$ via

$$\mathbb{R}^n \ni y \leftrightarrow f_y \in (\mathbb{R}^n)^* .$$

Ovenstående kommentarer 1, 3, 6 leder måske tanken hen på Cauchy–Schwarz' Ulighed. Hvis ikke, så fortvivl ikke, uligheden kommer nu.

Sætning 1.1 (Cauchy–Schwarz' Ulighed).

1) *Lad* $x, y \in \mathbb{R}^n$
så gælder $|\langle x, y \rangle| \leq \|x\| \, \|y\|$.

2) *Hvis* $|\langle x, y \rangle| = \|x\| \, \|y\|$ *og* $x \neq 0$ *så findes* $t \in \mathbb{R}$ *så* $y = tx$.

Bevis. 1) $\forall s \in \mathbb{R} : \|y - sx\|^2 \geq 0$

$$\|y - sx\|^2 = \|x\|^2 s^2 - 2\langle x, y \rangle s + \|y\|^2 \geq 0 .$$

Diskriminantuligheden ($D \leq 0$) giver $|\langle x, y \rangle| \leq \|x\| \, \|y\|$.

2) Antag $\langle x, y \rangle = \|x\| \, \|y\|$, thi hvis $\langle x, y \rangle = -\|x\| \, \|y\|$ kan x erstattes af $-x$ i de følgende regninger.

Antag endvidere, at $x \neq 0$ så fås

$$\|y - sx\|^2 = \|x\|^2 s^2 - 2\langle x, y \rangle s + \|y\|^2 = (s\|x\| - \|y\|)^2 .$$

For $s = \frac{\|y\|}{\|x\|}$ fås da $\|y - sx\| = 0$, dvs. y er et multiplum af x.

□

Cauchy–Schwarz' Ulighed er "VIGTIG" — den bruges igen og igen i såvel teori som i anvendelser. Vores første anvendelse bliver på kontinuitet af lineære afbildninger $F \in L(\mathbb{R}^n, \mathbb{R}^m)$.

Sætning 1.2. *Lad* $F \in L(\mathbb{R}^n, \mathbb{R}^m)$ *med tilhørende matrix* $A \in \mathbf{M}(m, n)$*, da gælder*

1) $\forall x \in \mathbb{R}^n : \|F(x)\| \leq \|A\| \, \|x\|$

2) $\forall x, y \in \mathbb{R}^n \|F(x) - F(y)\| \leq \|A\| \, \|x - y\|$

3) F er Lipschitz kontinuert med konstant $\leq \|A\|$.

Bevis. Antages 1) bevist, følger 2) af $F(x)-F(y) = F(x-y)$, og 3) følger af definitionen på Lipschitz kontinuitet og 2). Nu til beviset for udsagnet 1).

Fra Hovedsætningen fås for $x \in \mathbb{R}^n$

$$\begin{aligned}
\|F(x)\|^2 = \|Ax\|^2 &= \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j \right)^2 \qquad \text{v.h.a. Cauchy–Schwarz} \\
&\leq \sum_{i=1}^m \left(\left(\sum_{j=1}^n a_{ij}^2 \right) \left(\sum_{j=1}^n x_j^2 \right) \right) \\
&= \|x\|^2 \left(\sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}^2 \right) \right) = \|x\|^2 \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 \\
&= \|A\|^2 \|x\|^2 .
\end{aligned}$$

□

Bemærkning. I videregående matematik defineres $\|A\|$ - *operatornormen* - som det mindste ikke negative reelle tal k som opfylder $\|Ax\| \leq \|k\|x\|$ for alle x i $\mathbb{R}^n$. Den norm, jeg her har defineret på A, kaldes så for Hilbert–Schmidt normen, og vi kan altså se, at Hilbert–Schmidt normen dominerer operatornormen.

Vi slutter med et par observationer mere i tilknytning til Hovedsætningen og ovenstående sætning. Som bekendt er en afbildning F af $\mathbb{R}^m$ ind i $\mathbb{R}^m$ en homeomorfi netop hvis F er bijektiv, og både F og F^{-1} er kontinuerte afbildninger.

Sætning 1.3. *Lad $m \in \mathbb{N}$ og $F \in L(\mathbb{R}^m, \mathbb{R}^m)$ være en lineær afbildning med tilhørende matrix $A \in \mathbf{M}(m,m)$. Da gælder*

1) F er bijektiv $\iff \det A \neq 0$.

2) Hvis F er bijektiv er F^{-1} også lineær, og matricen hørende til F^{-1} er A^{-1}.

3) Hvis F er bijektiv er F en homeomorfi af $\mathbb{R}^m$ på $\mathbb{R}^m$.

4) F er bijektiv $\iff$ F er injektiv $\iff$ F er surjektiv.

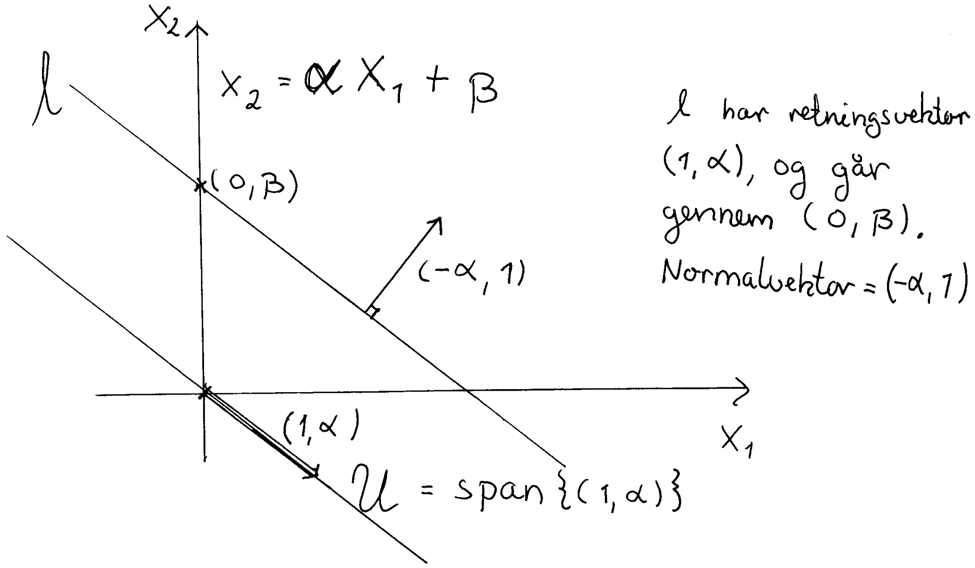
2 Affine rum og affine afbildninger

I almindelig daglig tale betegnes en funktion $f : \mathbb{R} \to \mathbb{R}$ givet ved $f(x) = \alpha x + \beta$ som en lineær funktion. Hvis $\beta \neq 0$, så er $L1$ og $L2$ på side 9 ikke opfyldte, så f er altså ikke lineær med derimod – som vi skal se – affin. Af hensyn til anvendeligheden af begrebet affin afbildning fra $\mathbb{R}^n$ ind i $\mathbb{R}^m$ vil vi godt have en abstrakt definition, men det kan godt allerede nu afsløres at man får noget meget konkret ud af anstrengelserne, idet det her, i Sætning 2.10 vises at de affine afbildninger fra $\mathbb{R}^n$ ind i $\mathbb{R}^m$ er dem der kan skrives på formen

$$F(x) = Ax + b\,;\ A \in \mathbf{M}(m,n)\,,\ b \in \mathbb{R}^m\,,\ x \in \mathbb{R}^n\,,$$

hvor vi altså ser analogien fra $m = n = 1$ klart.

For at få inspiration til forståelsen af de affine underrum af $\mathbb{R}^n$ kigger vi nu på linien ℓ i $\mathbb{R}^2$ givet ved



Denne linie ℓ er ikke et underrum *d. v. s. et lineært underrum, i betydnigen kendt fra Mat 1GA*, af $\mathbb{R}^2$ med mindre $\beta = 0$. Det ses imidlertid at når vi lader $\mathcal{U} :=$ span $\{(1,\alpha)\}$ være det lineære underrum udspændt af retningsvektoren $(1,\alpha)$ for ℓ, så er

$$\ell = (0,\beta) + \mathcal{U}\,,$$

Dette generaliseres nedenstående til affine underrum L af $\mathbb{R}^n$, idet vi i denne generelle ramme først vil give en abstrakt definition af begrebet affint underrum for så dernæst at vise at disse netop er de delmængder som har formen

$$L = x_o + \mathcal{U}\,,$$

hvor $x_0 \in \mathbb{R}^n$ og $\mathcal{U}$ er et lineært underrum i $\mathbb{R}^n$.

Af hensyn til anvendeligheden og af hensyn til at vi godt vil tillade $\emptyset$, den tomme mængde, som et affint underrum, er det hensigtsmæssigt at definere de affine underrum abstrakt for så derefter at vise at de ikke–tomme affine underrum af $\mathbb{R}^n$ får ovennævnte form.

Udgangspunktet for den abstrakte definition tages i den beskrivelse af linien $\ell(a_0, a_1)$ som blev præsenteret i afsnit 0,

$$\ell(a_0, a_1) = \{a = \alpha a_0 + \beta a_1 | \alpha + \beta = 1\}\,.$$

Definition 2.1. Lad $a_0, \ldots, a_k$ være punkter i $\mathbb{R}^n$ og $\lambda_0, \ldots, \lambda_k \in \mathbb{R}$. Hvis $\lambda_0 + \lambda_1 + \cdots + \lambda_k = 1$ kaldes linearkombinationen $\lambda_0 a_0 + \lambda_1 a_1 + \cdots + \lambda_k a_k$ en affin kombination af $a_0, \ldots, a_k$.

Definition 2.2. En delmængde $\mathcal{A}$ af $\mathbb{R}^n$ siges at være et affint underrum, hvis det for enhver affin kombination $\lambda_0 a_0 + \cdots + \lambda_k a_k$ af elementer fra $\mathcal{A}$ gælder at $\lambda_0 a_0 + \cdots + \lambda_k a_k \in \mathcal{A}$.

Den næste sætning er "typisk", og dens bevis indeholder en standardteknik, som det er værd at lære. Sætningen siger, at det er nok at $\mathcal{A}$ indeholder alle affine kombinationer af 2 elementer fra $\mathcal{A}$, for at $\mathcal{A}$ skal være et affint underrum. I henhold til beskrivelsen af linien $\ell(a_0, a_1)$ er det altså tilstrækkeligt at $\mathcal{A}$ indeholder alle linier mellem 2 af sine punkter for at $\mathcal{A}$ er et affint underrum. Når jeg fremhæver dette, er det fordi dette kan løfte lidt af sløret for hvorfor det er nødvendigt at studere de affine underrum abstrakt. Det kan nemt ske, at andre områder af matematikken producerer en mængde $\mathcal{A} \subseteq \mathbb{R}^n$ som har egenskaben

$$\forall a_0, a_1 \in \mathcal{A} : \ell(a_0, a_1) \subseteq \mathcal{A}\,.$$

Vore anstrengelser her viser så, at $\mathcal{A} = x_0 + \mathcal{U}$ med mindre $\mathcal{A} = \emptyset$. (Overvej at $\emptyset$ er et affint underrum).

Sætning 2.3. *Lad $\mathcal{A} \subseteq \mathbb{R}^n$. Følgende betingelser er ensbetydende:*

1) $\mathcal{A}$ er et affint underrum.

2) $\forall a_0, a_1 \in \mathcal{A}\ \forall t \in \mathbb{R} : (1-t)a_0 + ta_1 \in \mathcal{A}$.

3) $\forall a_0, a_1,\ a_0 \neq a_1 : \ell(a_0, a_1) \subseteq \mathcal{A}$.

Bevis. I henhold til afsnit 0, **Definition 0.1** er 2) og 3) ensbetydende.

For $t \in \mathbb{R}$ er $(1-t)a_0 + ta_1$ specielt en affin kombination af a_0 og a_1, så 1) $\Rightarrow$ 2). Det eneste vanskelige er så at vise 2) $\Rightarrow$ 1). Antag nu 2) er opfyldt, og lad $\lambda_0 a_0 + \cdots + \lambda_k a_k$ være en affin kombination af $a_0, \ldots, a_k \in \mathcal{A}$. Vi skal så vise, at $\lambda_0 a_0 + \cdots + \lambda_k a_k \in \mathcal{A}$. Dette gøres ved induktion efter k — antallet af elementer i den affine kombination minus 1. Vi betragter udsagnene $\mathcal{U}_k$, $k \in \mathbb{N}$

$$\begin{aligned} \mathcal{U}_k :=& \text{Enhver affin kombination } \lambda_0 a_0 + \cdots + \lambda_k a_k \\ & \text{af } k+1 \text{ elementer fra } \mathcal{A} \text{ er i } \mathcal{A}\,. \end{aligned}$$

For $k = 1$ er $\mathcal{U}_1$ netop udsagnet 2), som er forudsætningen i implikationen 2) $\Rightarrow$ 1).

Lad nu $k \geq 1$ og antag $\mathcal{U}_k$ er sand, vi vil så indse at $\mathcal{U}_{k+1}$ er sand, og deraf ved induktion slutte at 1) er opfyldt.

Lad $x = \lambda_0 a_0 + \cdots + \lambda_{k+1} a_{k+1}$ være en affin kombination af $k+2$ elementer fra $\mathcal{A}$. Da $k \geq 1$ er $k+2 \geq 3$, så for $0 \leq i \leq k+1$ kan ikke alle λ_i, kan være lig 1. Uden tab af generalitet kan vi derfor antage, at $\lambda_{k+1} \neq 1$. Da $\lambda_0 + \cdots + \lambda_k + \lambda_{k+1} = 1$ fås $\lambda_0 + \cdots + \lambda_k = 1 - \lambda_{k+1} \neq 0$, så vi får en affin kombination af $k+1$ punkter fra $\mathcal{A}$ ved

$$b = \frac{\lambda_0}{1-\lambda_{k+1}} a_0 + \cdots + \frac{\lambda_k}{1-\lambda_{k+1}} a_k\,.$$

Antagelsen $\mathcal{U}_k$ sikrer at $b \in \mathcal{A}$, og antagelsen 2) sikrer $(1-\lambda_{k+1})b + \lambda_{k+1}a_{k+1} \in \mathcal{A}$, men $(1-\lambda_{k+1})b + \lambda_{k+1}a_{k+1} = \lambda_0 a_0 + \cdots + \lambda_k a_k + \lambda_{k+1} a_{k+1}$. □

Som oplæg til den geometriske karakterisering af affine underrum som delmængder af $\mathbb{R}^n$ på formen $x_0 + \mathcal{U}$, starter vi med en definition af *linearkombinationer af mængder.*

Definition 2.4. Lad A, B være delmængder af $\mathbb{R}^n$ og $\alpha, \beta \in \mathbb{R}$. Ved udtrykket $\alpha A + \beta B$ forstås mængden

$$\alpha A + \beta B := \{\alpha a + \beta b | a \in A,\, b \in B\}\,.$$

□

Bemærk, at hvis bare en af mængderne er tom er $\alpha A + \beta B = \emptyset$. Begrundelsen er at hvis f.eks. $A = \emptyset$, så giver $a \in A$ aldrig et punkt i $\mathbb{R}^n$, dvs. der er ingen elementer af formen $\alpha a + \beta b$.

Lemma 2.5. *Lad $\mathcal{U}$ og $\mathcal{V}$ være lineære underrum af $\mathbb{R}^n$, $u \in \mathcal{U}$ og $\alpha \in \mathbb{R}$. Der gælder*

a) $u + \mathcal{U} = \mathcal{U}$.

(b) $\alpha\mathcal{U} + \mathcal{U} = \mathcal{U}$.

c) $\mathcal{U} + \mathcal{V}$ *er et lineært underrum.*

Før vi starter på beviset for Lemma 2.5 er det måske hensigtsmæssigt formelt at repetere, at en delmængde $\mathcal{U} \subseteq \mathbb{R}^n$ er et *lineært underrum*, hvis nedenstående 3 betingelser er opfyldte.

U0: $0 \in \mathcal{U}$.

U1: $\forall u_1, u_2 \in \mathcal{U}$: $u_1 + u_2 \in \mathcal{U}$.

U2: $\forall \alpha \in \mathbb{R}\ \forall u \in \mathcal{U}$: $\alpha u \in \mathcal{U}$.

Bemærk specielt, at $\emptyset$ *ikke* er et lineært underrum, (men derimod et affint underrum).
Undertiden bruges udtrykket *underrum* i opgaver og i tekst uden specifikation af arten af underrummet. Når dette sker betyder det at der er tale om et *lineært underrum*, i overensstemmelse med den fra Mat 1GA kendte sprogbrug.

Bevis. Lemma 2.5.

a) Fra U1 fås $u+\mathcal{U} \subseteq \mathcal{U}$. Nu er $\mathcal{U} = u+(-u+\mathcal{U})$, så igen fra U1 fås $\mathcal{U} = u+(-u+\mathcal{U}) \subseteq u+\mathcal{U}$.

b) Fra U2 fås $\alpha\mathcal{U} \subseteq \mathcal{U}$, og dermed fra U1 : $\alpha\mathcal{U}+\mathcal{U} \subseteq \mathcal{U}$. På den anden side er $0 \in \alpha\mathcal{U}$, så $\alpha\mathcal{U} + \mathcal{U} \supseteq 0 + \mathcal{U} = \mathcal{U}$.

c) Oplagt: Vis det selv!

$\square$

Opgave: Vis generelt, at for $x \in \mathbb{R}^n$ og $\mathcal{B} \subseteq \mathbb{R}^n$ gælder

$$x + (-x + \mathcal{B}) = \mathcal{B}\,.$$

Vis, at hvis $\mathcal{A}$ og $\mathcal{B}$ er delmængder af $\mathbb{R}^n$, således at $\mathcal{A}$ har mindst 2 elementer og $\mathcal{B} = \{0\}$, så gælder der ikke lighedstegn, men:

$$\mathcal{A} + (-\mathcal{A} + \mathcal{B}) \underset{\neq}{\supset} \mathcal{B}\,.$$

Sætning 2.6. *Lad $\mathcal{A}$ være en ikke tom delmængde af $\mathbb{R}^n$. Følgende betingelser er ensbetydende:*

a) Der findes $a_0 \in \mathbb{R}^n$ og et lineært underrum $\mathcal{U} \subseteq \mathbb{R}^n$ så $\mathcal{A} = a_0 + \mathcal{U}$.

b) For alle $a \in \mathcal{A}$ er $\mathcal{A} - a$ et lineært underrum af $\mathbb{R}^n$.

c) $\mathcal{A}$ er et affint underrum.

Hvis $\mathcal{A}$ er et affint underrum er det lineære underrum $\mathcal{A} - a$ uafhængigt af valget af a.

Bevis. Der vises $a) \Rightarrow b) \Rightarrow c) \Rightarrow a)$, og til sidst at $\forall a_1, a_2 \in \mathcal{A}$ gælder $a) \Rightarrow (\mathcal{A} - a_1 = \mathcal{A} - a_2)$.

Antag a), og lad $a \in \mathcal{A} = a_0 + \mathcal{U}$, dvs. der findes $u \in \mathcal{U}$ så $a = a_0 + u$. Dermed er $\mathcal{A} - a = (a_0 + \mathcal{U}) - (a_0 + u) = \mathcal{U} - u = \mathcal{U}$, ifølge Lemma 2.5 a); og b) følger.

Antag b), og lad $a_0, a_1 \in \mathcal{A}$ samt $\alpha \in \mathbb{R}$, da fås, idet $\mathcal{U} := \mathcal{A} - a$ et lineært underrum, at der findes $u_0, u_1 \in \mathcal{U}$, så $a_0 = a + u_0$, $a_1 = a + u_1$. Heraf ses $(1-\alpha)a_0 + \alpha a_1 = (1-\alpha)a + \alpha a + ((1-\alpha)u_0 + \alpha u_1) \in a + \mathcal{U} = \mathcal{A}$. Ifølge Sætning 2.3 er $\mathcal{A}$ et affint underrum af $\mathbb{R}^n$, og c) følger.

Antag c), og vælg $a_0 \in \mathcal{A}$ samt definer $\mathcal{U} = \mathcal{A} - a_0$. Så er $\mathcal{A} = a_0 + \mathcal{U}$, og vi skal blot vise at $\mathcal{U}$ er et underrum.

U1: $a_0 \in \mathcal{A}$ så $0 \in \mathcal{A} - a_0 = \mathcal{U}$.

U2: Lad $u_1 = a_1 - a_0 \in \mathcal{A} - a_0$, $u_2 = a_2 - a_0 \in \mathcal{U} = \mathcal{A} - a_0$, så fås

$$u_1 + u_2 = a_1 - a_0 + a_2 - a_0 = (-a_0 + a_1 + a_2) - a_0 .$$

Da $\mathcal{A}$ er et affint underrum er $(-a_0 + a_1 + a_2) \in \mathcal{A}$, og dermed $u_1 + u_2 \in \mathcal{U}$.

U3: Lad $u_1 = a_1 - a_0 \in \mathcal{U}$ og $\alpha \in \mathbb{R}$, så er

$$\begin{aligned} &\alpha u_1 = \alpha(a_1 - a_0) = ((1-\alpha)a_0 + \alpha a_1) - a_0 \in \mathcal{A} - a_0 = \mathcal{U} \\ &\alpha u_1 \in \mathcal{U}o . \end{aligned}$$

Dvs. $\mathcal{A} - a_0$ er et lineært underrum og a) følger.

$\square$

Beviset for at det lineære underrum $\mathcal{A} - a$ er uafhængigt af valget af a, når $\mathcal{A}$ er et affint underrum kan aflæses af beviset a) $\Rightarrow$ b), idet det heraf følger, at når $\mathcal{A} = a_0 + \mathcal{U}$, så er $\mathcal{U} = \mathcal{A} - a$ for alle $a \in \mathcal{A}$.

Eksempel 2.7. Lad $A \in \mathbf{M}(m, n)$, $b \in \mathbb{R}^m$, $x \in \mathbb{R}^n$ og definer

$$\mathcal{L} = \{x \in \mathbb{R}^n | Ax = b\} .$$

Fra teorien for lineære ligningssystemer ved vi, at enten er $\mathcal{L} = \emptyset$ (altså ingen løsninger) eller også, hvis der er en løsning x_0, så er

$$\mathcal{L} = x_0 + \ker A = x_0 + \{x \in \mathbb{R}^n | Ax = 0\} .$$

Heraf ses at $\mathcal{L}$ er et affint underrum, da $\ker A$ er et (velkendt) lineært underrum i $\mathbb{R}^n$. Selv om vi nu ikke kendte til teorien for lineære ligningssystem kunne vi godt v.h.a. Sætning 2.3 bevise, at $\mathcal{L}$ er et affint underrum af $\mathbb{R}^n$. Lad nemlig $x_0, x_1 \in \mathcal{L}$ og $\alpha \in \mathbb{R}$, så fås $A((1-\alpha)x_0 + \alpha x_1) = (1-\alpha)Ax_0 + \alpha Ax_1 = (1-\alpha)b + \alpha b = b$, så $(1-\alpha)x_0 + \alpha x_1 \in \mathcal{L}$. Fortsætter vi nu med at antage at vort kendskab til lineær algebra er borte, giver Sætning 2.6, at hvis $\mathcal{L} \neq \emptyset$ og $x_0 \in \mathcal{L}$, så er $\mathcal{L} - x_0$ et lineært underrum, og det er ret let at se, at dette lineære underrum netop må være $\ker A$.

Det bør endvidere bemærkes, at ethvert affint underrum $\mathcal{L}$ kan beskrives som

$$\mathcal{L} = \{x \in \mathbb{R}^n | Ax = b\}\,.$$

For at indse dette skrives $\mathcal{L} = x_0 + \mathcal{U}$, hvor $\mathcal{U}$ er et lineært underrum. Lad $\mathcal{U}^\perp$ betegne det ortogonale komplement til $\mathcal{U}$ i $\mathbb{R}^n$ og lad m betegne dimensionen af dette underrum. Vælg så en ortonormalbasis $a_1, \ldots, a_m$ for $\mathcal{U}^\perp$. Eftersom det ortogonale komplement til $\mathcal{U}^\perp$ netop er

$$\mathcal{U} = \{x \in \mathbb{R}^n | a_1 \cdot x = 0, \ldots, a_m \cdot x = 0\}\,.$$

ses det, at hvis vi lader A betegne den $m \times n$ matrix hvis rækker netop består af vektorerne $\{a_1, \ldots, a_m\}$ så er

$$\mathcal{U} = \{x \in \mathbb{R}^n | Ax = 0\}\,.$$

og vi får ved at definere $b := Ax_0$ at

$$\mathcal{L} = x_0 + \{x \in \mathbb{R}^n | Ax = 0\} = \{x \in \mathbb{R}^n | Ax = b\}$$

Vi vender os nu til de affine afbildninger mellem affine underrum af talrum.

Definition 2.8. Lad $\mathcal{A} \subseteq \mathbb{R}^n$ og $\mathcal{B} \subseteq \mathbb{R}^m$ være affine underrum og $F : \mathcal{A} \to \mathcal{B}$ en afbildning. Hvis det for en vilkårlig affin kombination $\alpha_0 a_0 + \cdots + \alpha_k a_k$ af elementer fra $\mathcal{A}$ gælder at

$$F(\alpha_0 a_0 + \cdots + \alpha_k a_k) = \alpha_0 F(a_0) + \cdots + \alpha_k F(a_k)\,,$$

siges F at være en affin afbildning.

Vort næste resultat er analogt til Sætning 2.3, og dets bevis også. Det udsiger at kombinationer med $k = 1$ - dvs affine kombinationer af 2 elementerer - er nok til at teste affinitet af en afbildning.

Sætning 2.9. *Lad $\mathcal{A} \subseteq \mathbb{R}^n$, $\mathcal{B} \subseteq \mathbb{R}^m$ være affine underrum og $F : \mathcal{A} \to \mathcal{B}$ en afbildning, da gælder det at F er affin hvis og kun hvis det for alle par $a_0, a_1 \in \mathcal{A}$ og $\alpha \in \mathbb{R}$ gælder, at $F((1-\alpha)a_0 + \alpha a_1) = (1-\alpha)F(a_0) + \alpha F(a_1)$.*

Bevis. Det er klart at hvis F er affin så vil $F((1-\alpha)a_0 + \alpha a_1) = (1-\alpha)F(a_0) + \alpha F(a_1)$ være opfyldt. Antag nu at ovenstående ligning gælder for vilkårlige $a_0, a_1 \in \mathcal{A}$, $\alpha \in \mathbb{R}$, vi skal så indse at F er affin. Dette gøres v.h.a. induktion hvor induktionsudsagnet $\mathcal{U}_k$ er givet ved ($k \in \mathbb{N}$)

$$\begin{aligned} \mathcal{U}_k :& \forall a_0, \ldots, a_k \in \mathcal{A}\ \forall \alpha_0, \ldots, \alpha_k \in \mathbb{R}\,,\ \alpha_0 + \cdots + \alpha_k = 1 : \\ & F(\alpha_0 a_0 + \cdots + \alpha_k a_k) = \alpha_0 F(a_0) + \cdots + \alpha_k F(a_k)\,. \end{aligned}$$

$\mathcal{U}_1$ er antaget gyldig fra starten af beviset. Lad nu $k \geq 1$ og antag $\mathcal{U}_k$ gyldig. Vi skal da vise at $\mathcal{U}_{k+1}$ er gyldig. Lad derfor $\alpha_0 a_0 + \cdots + \alpha_{k+1} a_{k+1}$ være en affin kombination af $k+2$ elementer fra $\mathcal{A}$. Vi vender tilbage til beviset fra Sætning 2.3, 2) $\Rightarrow$ 1). Som der antages det at $\alpha_{k+1} \neq 1$ og vi får

$$F(\alpha_0 a_0 + \cdots + \alpha_{k+1} a_{k+1}) = F((1-\alpha_{k+1})\left(\frac{\alpha_0}{1-\alpha_{k+1}} a_0 + \cdots + \frac{\alpha_k}{1-\alpha_{k+1}} a_k\right) + \alpha_{k+1} a_{k+1})$$

(antagelsen $\mathcal{U}_1$)

$$= (1-\alpha_{k+1}) F\left(\frac{\alpha_0}{1-\alpha k+1} a_0 + \cdots + \frac{\alpha_k}{1-\alpha_{k+1}} a_k\right) + \alpha_{k+1} F(a_{k+1})$$

(antagelsen) $\mathcal{U}_k$

$$= \alpha_0 F(a_0) + \cdots + \alpha_k F(a_k) + \alpha_{k+1} F(a_{k+1}).$$

□

Vi kan nu karakterisere de affine afbildninger fra $\mathbb{R}^n$ ind i $\mathbb{R}^m$.

Sætning 2.10. *(a) Lad $A \in \mathbf{M}(m,n)$ og $b \in \mathbb{R}^m$.*
Der defineres en affin afbildning $F : \mathbb{R}^n \to \mathbb{R}^m$ ved

$$x \in \mathbb{R}^n\,,\ F(x) := Ax + b\,.$$

(b) Enhver affin afbildning $F : \mathbb{R}^n \to \mathbb{R}^m$ har denne form, og der gælder

$$b = F(0) \quad \forall j \in \{1, \ldots, n\} : A_j = F(e_j) - F(0)\,.$$

Bevis. (a): Lad $x_0, x_1 \in \mathbb{R}^n$ og $\alpha \in \mathbb{R}$, da fås

$$\begin{aligned} F((1-\alpha)x_0 + \alpha x_1) &= A((1-\alpha)x_0 + \alpha x_1) + b \\ &= (1-\alpha)Ax_0 + \alpha A x_1 + b \\ &= (1-\alpha)(Ax_0 + b) + \alpha(Ax_1 + b) \\ &= (1-\alpha)F(x_0) + \alpha F(x_1)\,. \end{aligned}$$

(b): Lad F være affin: $\mathbb{R}^n \to \mathbb{R}^m$ og definer $b := F(0)$ samt en matrix $A \in \mathbf{M}(m,n)$ ved for $j \in \{1, \ldots, n\}$ at definere den j'te søjle $A_j = F(e_j) - b$.

Tricket er, at $x \in \mathbb{R}^n$ kan skrives som

$$\begin{aligned} &x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n \text{ så} \\ &x = (1 - x_1 - x_2 - \cdots - x_n) \cdot 0 + x_1 e_1 + x_2 e_2 + \cdots + x_n e_n\,, \end{aligned}$$

altså er x en affin konbination af $0, e_1, \ldots, e_n$, så da F er en affin afbildning er

$$\begin{aligned} F(x) &= (1 - x_1 - \cdots - x_n)F(0) + x_1 F(e_1) + \cdots + x_n F(e_n) \\ &= F(0) + x_1(F(e_1) - F(0)) + \cdots + x_n(F(e_n) - F(0))\,. \end{aligned}$$

Defineres da $b = F(0)$, og en $m \times n$ matrix A ved $A_j := F(e_j) - F(0)$, fås da ved 1.3 5), at

$$F(x) = Ax + b\,.$$

□

I den lineære algebra spiller begreber som "span" og "lineær uafhængighed" og basis en stor rolle. Begreberne har deres modstykker i teorien for affine strukturer. Vi skal ikke dyrke disse ting i mange detaljer, og en del af resultaterne vil ikke blive bevist her, men overladt til øvelserne. Først defineres det analoge begreb til span kaldet "aff". Definitionen er baseret på følgende lemma:

Lemma 2.11. *Lad $(\mathcal{A}_i)_{i\in I}$ være en familie af affine underrum af $\mathbb{R}^n$, da er $\mathcal{A} := \bigcap_{i\in I} \mathcal{A}_i$ et affint underrum.*

Bevis. Hvis $a_0, a_1 \in \mathcal{A}$ og $\alpha \in \mathbb{R}$ er $a_0, a_1 \in \mathcal{A}_i$ for alle $i \in I$ så $(1-\alpha)a_0 + \alpha a_1 \in \mathcal{A}_i$ for alle $i \in I$, og dermed $(1-\alpha)a_0 + \alpha a_1 \in \mathcal{A}$, så ifølge Sætning 2.3, er $\mathcal{A}$ et affint underrum. □

Vi kan så definere

Definition 2.12. Lad M være en delmængde af $\mathbb{R}^n$. Ved det affine hylster – $\operatorname{aff}(M)$ – af M forstås det mindste affine underrum som indeholder M.

Bemærk $\mathbb{R}^n \supseteq M$, så der findes affine underrum der indeholder M. Lemma 2.11 viser at fællesmængden af alle affine underrum som indeholder M igen er et affint underrum som indeholder M, men da dette efter sin konstruktion er indeholdt i ethvert andet affint underrum som indeholder M, må det være det mindste sådanne.

Sætning 2.13. *Lad $M \subseteq \mathbb{R}^n$, $M \neq \emptyset$ og $m_0 \in M$.*

a) $\operatorname{aff}(M) = m_0 +$ *span* $(M - m_0)$.

b) $\operatorname{aff}(M) = \{\alpha_0 n_0 + \cdots + \alpha_k n_k | k \in \mathbb{N}\,,\ \alpha_0 + \cdots + \alpha_k = 1\,,\ n_i \in M\}$

c) $\operatorname{aff}(M) = \{$*alle affine kombinationer af elementer fra* $M\}$.

Bevis. a) $m_0 +$ span $(M - m_0)$ er et affint underrum som indeholder M så $m_0 +$ span $(M - m_0) \supseteq \operatorname{aff}(M)$. Omvendt vides det, at $\operatorname{aff}(M) = m_0 + \mathcal{U}$ for et lineært underrum $\mathcal{U}$, så $M = m_0 + (M - m_0) \subseteq m_0 + \mathcal{U}$ giver $M - m_0 \subseteq \mathcal{U}$, og dermed span $(M - m_0) \subseteq \mathcal{U}$, så $m_0 +$ span $(M - m_0) \subseteq m_0 + \mathcal{U} = \operatorname{aff}(M)$.

b) Lad $\mathcal{B} := \{\alpha_0 n_0 + \cdots + \alpha_k n_k | k \in \mathbb{N}, \alpha_0 + \cdots + \alpha_k = 1, n_i \in M\}$. For $b = \alpha_0 n_0 + \cdots + \alpha_k n_k \in \mathcal{B}$ fås, idet $\alpha_0 + \cdots + \alpha_k = 1$,

$$b = m_0 + \alpha_0(n_0 - m_0) + \cdots + \alpha_k(n_k - m_0) \in m_0 + \text{ span } (M - m_0),$$

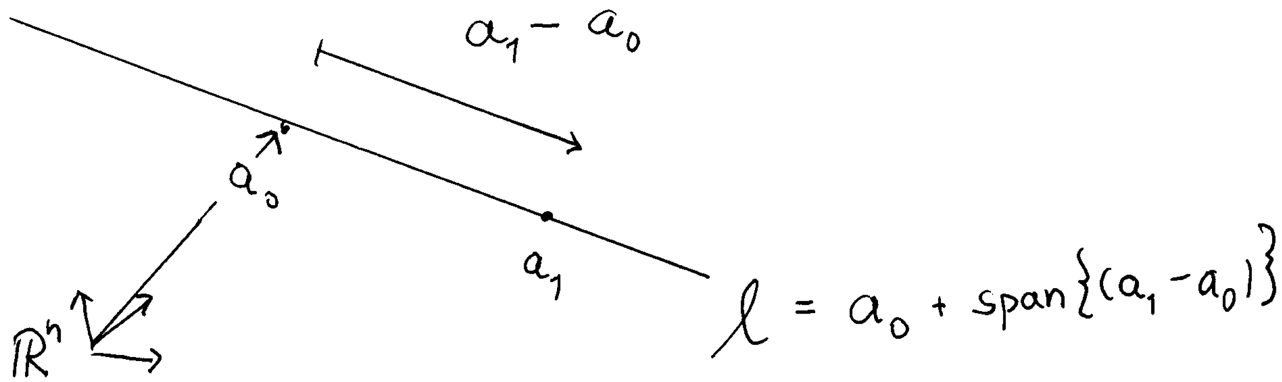
dvs. ifølge a) fås $\mathcal{B} \subseteq \text{aff}(M)$. Fra lineær algebra vides det at span $(M - m_0)$ består af samtlige mulige linearkombinationer af elementer fra $(M - m_0)$. Lad da $u \in$ span $(M - m_0)$, $u = \lambda_1(m_1 - m_0) + \cdots + \lambda(m_n - m_0)$ for et $n \in \mathbb{N}$, $\lambda_i \in \mathbb{R}$. For $a = m_0 + u \in \text{aff}(M)$ fås da

$$a = (1 - \lambda_1 - \cdots - \lambda_n)m_0 + \lambda_1 m_1 + \cdots + \lambda_n m_n \in \mathcal{B}$$

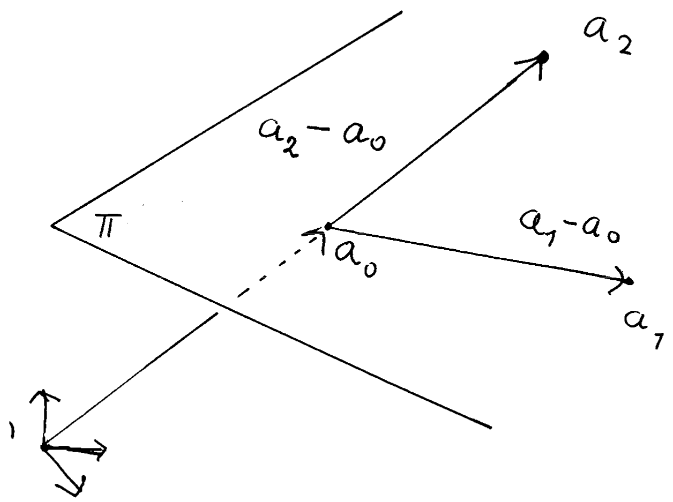
så $\text{aff}(M) \subseteq \mathcal{B}$ og b) følger.

□

Forhåbentlig kan man nu se strukturen af et affint underrum for sig, men som illustration bringes et par tegniger som støtte for ens indre billede af dette abstrakte begreb. Linien gennem 2 forskellige punkter i $\mathbb{R}^n$: Linien er bestemt ved $\ell(a_0, a_1) = a_0 + \text{span}(a_1 - a_0)$.



Planen Π gennem 3 punkter i $\mathbb{R}^3$ eller $\mathbb{R}^n$ der ikke ligger på linie:



Planen Π er bestemt ved:

$$\begin{aligned}\Pi &= a_0 + \text{ span } \{(a_1 - a_0), (a_2 - a_0)\} = \text{aff}(\{a_0, a_1, a_2\}) \\ &= a_0 + \{s(a_1 - a_0) + t(a_2 - a_0) | s, t \in \mathbb{R}\}\end{aligned}$$

Det affine underrum Π er en 2–dimensional størrelse, men der skal 3 punkter til at udspænde planen: a_0 et punkt i planen Π og 2 retningsvektorer for Π, $a_1 - a_0$, $a_2 - a_0$.

Vort sidste begreb er den formelle definition som løser problemet med hvorledes vi skal lave en formel generalisation af udtrykket "de 3 punkter må ikke ligge på linie" når vi har mere end 3 punkter at tage hensyn til.

Definition 2.14. Lad $a_0, \dots, a_k$ være punkter i $\mathbb{R}^n$. Disse siges at være affint afhængige hvis der findes et talsæt $(\alpha_0, \dots, \alpha_k) \in \mathbb{R}^{k+1} \setminus (0, \dots, 0)$, så $\alpha_0 a_0 + \cdots + \alpha_k a_k = 0$, og $\alpha_0 + \cdots + \alpha_k = 0$. Hvis sættet ikke er affint afhængigt siges det at være affint uafhængigt.

Definitionen er lidt uhåndterlig, men hvis man tænker på Sætning 2.13 giver nedenstående sætning en forklaring på betydningen af denne definition.

Sætning 2.15. *Lad $k \in \mathbb{N}$ og $a_0, \dots, a_k$ være punkter i $\mathbb{R}^n$. Disse punkter er affint uafhængige hvis og kun hvis punkterne $(a_1 - a_0)$, $(a_2 - a_0), \dots, (a_k - a_0)$ er lineært uafhængige.*

Bevis. Det er nemmest at vise det kontraponerede udsagn, dvs. at $a_0, \dots, a_k$ er affint afhængige netop når $(a_1 - a_0), \dots, (a_k - a_0)$ er lineært afhængige. Metoden er snart set et par gange!

Antag $\alpha_0 a_0 + \cdots + \alpha_k a_k = 0$ og $\alpha_0 + \cdots + \alpha_k = 0$, og ikke alle $\alpha_i = 0$. Da fås $\alpha_0 = -\alpha_1 - \cdots - \alpha_k$ så ikke alle $\alpha_1, \dots, \alpha_k$ er nul, og

$$0 = \alpha_0 a_0 + \cdots + \alpha_k a_k = \alpha_1(a_1 - a_0) + \cdots + \alpha_k(a_k - a_0)\,.$$

Dvs.$a_0, \dots, a_k$ affin afhængig $\implies$ $(a_1 - a_0), \dots, (a_k - a_0)$ lineæt afhængig.

Antag nu $\lambda_1(a_1 - a_0) + \cdots + \lambda_k(a_k - a_0) = 0$ og ikke alle $\lambda_i = 0$, så fås for $\lambda_0 := -\lambda_1 - \cdots - \lambda_k$

$$0 = \lambda_1(a_1 - a_0) + \cdots + \lambda_k(a_k - a_0) = \lambda_0 a_0 + \lambda_1 a_1 + \cdots + \lambda_k a_k$$

$\lambda_0 + \lambda_1 + \cdots + \lambda_k = 0$, og ikke alle $\lambda_i = 0$, dvs.

$$(a_1 - a_0), \dots, (a_k - a_0) \text{ lineært afhængig } \implies a_0, a_1, \dots, a_k \text{affint afhængig}\,.$$

Sammenfattende kan vi nu ud fra Sætning 2.13 og Sætning 2.15 slutte, at hvis $a_0, a_1, \dots, a_k$ er et affint uafhængigt sæt af vektorer, så er $\text{aff}(a_0, \dots, a_k) = a_0 +$ span $((a_1 - a_0), \dots, (a_k -$

$a_0)$), og til hvert punkt $a \in \text{aff}(a_0, \dots, a_k)$ findes derfor netop et sæt af parametre for "retningsvektorerne" $(a_1 - a_0), \dots, (a_k - a_0)$, så

$$\begin{aligned} a &= a_0 + \lambda_1(a_1 - a_0) + \cdots + \lambda_k(a_k - a_0) \\ &= (1 - \lambda_1 - \cdots - \lambda_k)a_0 + \lambda_1 a_1 + \cdots + \lambda_k a_k \,. \end{aligned}$$

Altså ethvert punkt a i $\text{aff}(a_0, \dots, a_k)$ kan på netop én måde fremstilles som en affin kombination af $(a_0, \dots, a_k)$. Man siger så, at $(a_0, \dots, a_k)$ er en *affin basis* for $\text{aff}(a_0, \dots, a_k)$. De kendte sætninger for baser i underrum overføres nu let til affine underrum, blot skal man altid have et punkt mere, da affine underrum ikke går gennem 0 – nødvendigvis. Vi slutter med en oplagt definition. □

Definition 2.16. Lad $\mathcal{A}$ være et affint underrum af $\mathbb{R}^n$. Hvis $\mathcal{A} = \emptyset$ defineres $\dim(\mathcal{A}) = -1$. Hvis $\mathcal{A} = a_0 + \mathcal{U}$ hvor $\mathcal{U}$ er et lineært underrum af $\mathbb{R}^n$, defineres $\dim(\mathcal{A}) := \dim \mathcal{U}$. □

Bemærk at et punkt $x_0 \in \mathbb{R}^n$ har formen $x_0 = x_0 + \{0\}$, så et punkt er et affint underrum af dimension 0.

Vi slutter med et par bemærkninger om topologiske egenskaber for affine afbildninger og affin underrum. Det er let at se at en translation i rummet dvs. en afbildning af formen $\mathbb{R}^n \ni x \to x + a \in \mathbb{R}^n$ er en isometri og dermed kontinuert. Dens inverse består af translation med vektoren $-a$, så en translation er også en homeomorfi. Eftersom vi har set at lineære afbildninger er kontinuerte ses det at en affin afbildning må være kontinuert da den har form som en lineær afbildning efterfulgt af en translation. Med baggrund i disse observationer samt Eksempel 2.7 kan vi nu vise følgende sætning.

Sætning 2.17. *Et affint underrum $\mathcal{L}$ i $\mathbb{R}^n$ er afsluttet.*

Bevis. Fra ovenstående overvejelser samt Sætning 2.6 følger det at det er nok at vise at ethvert lineært underrum $\mathcal{U}$ af $\mathbb{R}^n$ er afsluttet. Fra Eksempel 2.7 ved vi at ethevert lineært underrum er kerne for en lineær afbildning og eftersom sådanne afbildninger er kontinuerte er deres kerner afsluttetde. □

3 Den rette linie i planen

Emnet i dette afsnit er ikke liniens ligning, men derimod 8 forskellige sammenhænge hvori en linie i planen optræder. Der er sikkert flere mulige synspunkter på en linie, men de 8 jeg nævner her er alle relevante for dette kursus, og kan tillige generaliseres til $(n-1)$–dimensionale affine underrum af $\mathbb{R}^n$; de såkaldte hyperplaner i $\mathbb{R}^n$. Dette sidste vil være emnet for kapitel 4.

De 8 opfattelser er:

1) Et matematisk objekt.

2) Et affint underrum af $\mathbb{R}^2$ af dimension $2-1=1$.

3) Det affine hylster af 2 affint uafhængige punkter i $\mathbb{R}^2$.

4) Graf for en affin funktion $\varphi : \mathbb{R} \to \mathbb{R}$.

5) Niveaumængde for en affin funktion $F : \mathbb{R}^2 \to \mathbb{R}$.

6) Randen af en halvplan.

7) Graf for den affine approksimerende funktion til en differentiabel funktion f, gennem et punkt $(x_0, f(x_0))$.

8) Tangent til grafen for en differentiabel reel funktion f af en variabel gennem et punkt $(x_0, f(x_0))$.

1): Planen identificeres med $\mathbb{R}^2$. Ifølge Euklids aksiomer for plangeometrien går der gennem 2 vilkårlige forskellige punkter i planen netop én linie. Dette er den klassiske geometris opfattelse. Dette er i virkeligheden meget abstrakt, og vi skal ikke gå ind på disse ting her, men blot sige, at som model for Euklids geometri kan man bruge analytisk geometri. Det betyder, at når vi taler om planen tænker vi på $\mathbb{R}^2$, og når vi tænker på linier i planen, tænker vi på punktmængder af formen

$$\ell = \{(x_1, x_2) \in \mathbb{R}^2 | a_1x_1 + a_2x_2 = b\} \ \ (a_1, a_2) \neq (0,0)\,.$$

Bemærk, at hvis $a_1 = a_2 = b = 0$, så er $\ell = \mathbb{R}^2$, og hvis $a_1 = a_2 = 0$, $b \neq 0$, så er $\ell = \emptyset$. Derfor kræves det $(a_1, a_2) \neq (0,0)$ for at ℓ bliver en linie.

2): Et affint underrum af dimension $2-1=1$.

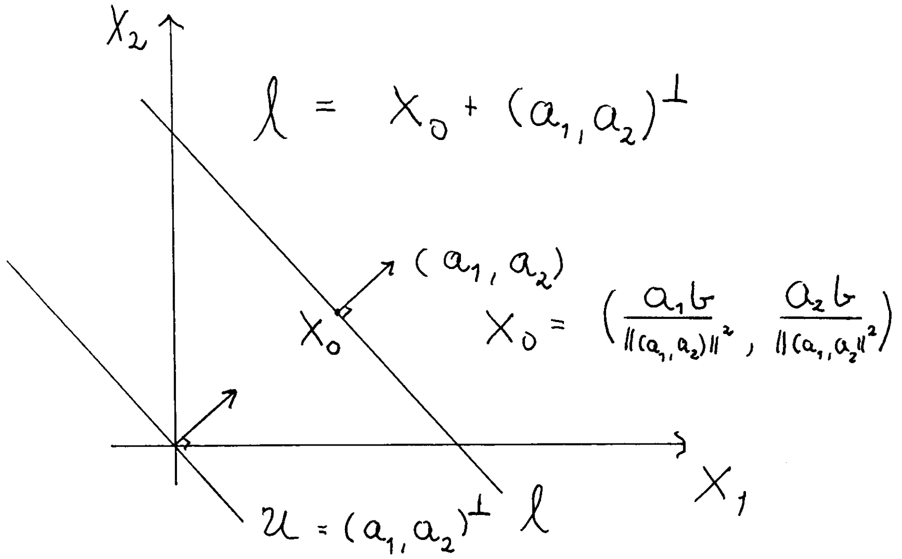
Lad $\ell = \{(x_1, x_2) \in \mathbb{R}^2 | a_1x_1 + a_2x_2 = b\}$ hvor $(a_1, a_2) \neq (0,0)$. Da er $\ell \neq \emptyset$, idet

$$x_0 := \left(\frac{ba_1}{\|(a_1,a_2)\|^2}, \frac{ba_2}{\|(a_1,a_2)\|^2} \right) \in \ell,$$

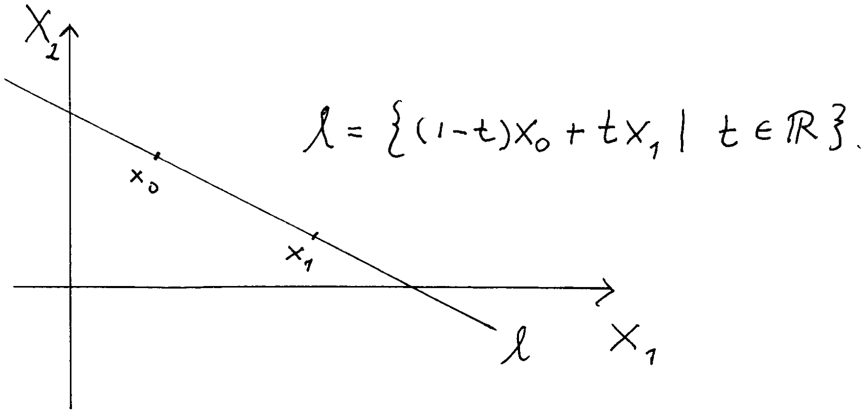
og for $\mathcal{U} := (a_1, a_2)^\perp = \{(x_1, x_2) \in \mathbb{R}^2 | a_1x_1 + a_2x_2 = 0\}$ gælder, at $\mathcal{U}$ er et lineært underrum af $\mathbb{R}^2$ som opfylder

$$\ell = x_0 + \mathcal{U}.$$

Altså ℓ er et affint underrum af $\mathbb{R}^2$ af dimension 1. Grunden til at vi skriver dimensionen som $2-1$ er, at $\mathcal{U} := (a_1, a_2)^\perp = \ker f_{(a_1,a_2)}$, hvor $f_{(a_1,a_2)}$ som bekendt betegner den lineære afbildning: $\mathbb{R}^2 \to \mathbb{R}$ givet ved $f_{(a_1,a_2)}(x_1, x_2) = a_1x_1 + a_2x_2$. Da $(a_1, a_2) \neq (0,0)$ er denne afbildning surjektiv, så ifølge dimensionssætningen er dimensionen af $\ker f_{(a_1,a_2)} = 2 - 1 = 1$.



3): Det affine hylster af 2 forskellige punkter x_0 og x_1 på ℓ.



I det konkrete tilfælde fra 2) kan man bruge $x_0 = \left(\frac{a_1b}{\|(a_1,a_2)\|^2}, \frac{a_2b}{\|(a_1,a_2)\|^2} \right)$ og $x_1 = x_0 + (a_2, -a_1)$, idet $(a_2, -a_1)$ er retningsvektor for linien.

4): Graf for en affin funktion $\varphi : \mathbb{R} \to \mathbb{R}$. OBS: Her forudsættes $a_2 \neq 0$.

Når $a_2 \neq 0$ kan ligningen $a_1x_1 + a_2x_2 = b$ omskrives til $x_2 = -\frac{a_1}{a_2}x_1 + \frac{b}{a_2}$. Lader vi da $\varphi : \mathbb{R} \to \mathbb{R}$ være givet ved

$$\forall x_1 \in \mathbb{R} : \varphi(x_1) := -\frac{a_1}{a_2}x_1 + \frac{b}{a_2}$$

ses det at grafen for φ, kaldet $G(\varphi)$, er givet ved

$$G(\varphi) = \{(x_1, x_2) | x_2 = \varphi(x_1)\} = \ell\,.$$

5): Niveaumængden for en affin funktion $F : \mathbb{R}^2 \to \mathbb{R}$.

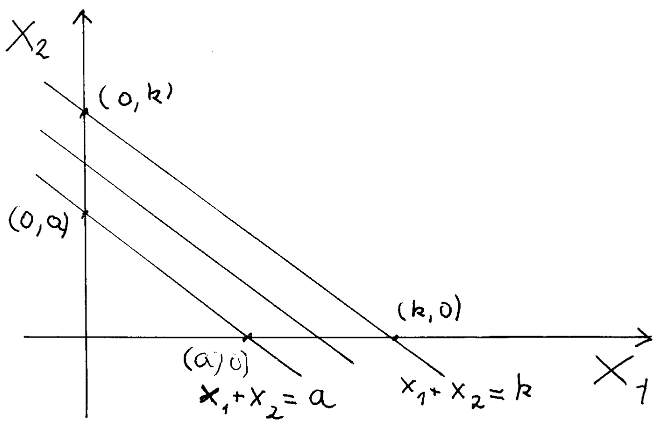
Lad $F : \mathbb{R}^2 \to \mathbb{R}$ være givet ved

$$F(x_1, x_2) := a_1x_1 + a_2x_2\,.$$

Da er $\ell = \{(x_1, x_2)|F(x_1, x_2) = b\} = F^{-1}(b)$, altså niveaumængden for F i niveau b. Bemærk, at F ikke er entydigt bestemt ud fra problemstillingen – vi kunne lige så gerne have brugt en vilkårlig af funktionerne F_c, $c \in \mathbb{R}$

$$F_c(x_1, x_2) := a_1x_1 + a_2x_2 + c\,.$$

Bemærk endelig, at alle disse niveaumængder $F^{-1}(\alpha)$, $\alpha \in \mathbb{R}$ udgør et bundt linier i planen som alle er parallelle med $\mathcal{U} = (a_1, a_2)^{\perp}$.

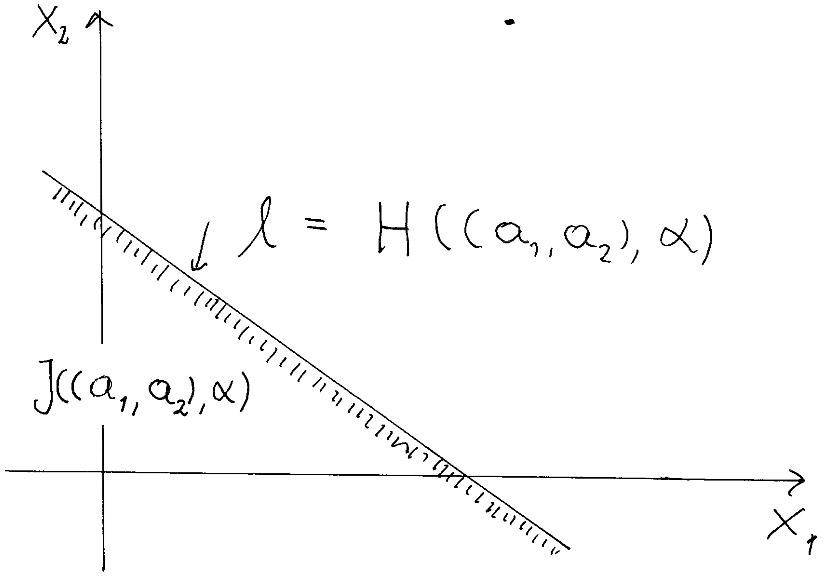


6): Randen af en halvplan.

Ved en halvplan forstås et område J af planen defineret v.h.a. en “affin” ulighed:

$$\begin{aligned} J((a_1, a_2), b) :=& \{(x_1, x_2)|a_1x_1 + a_2x_2 \leq b\} \\ \ell =& \{(x_1, x_2)|a_1x_1 + a_2x_2 = b\} \end{aligned}$$

Linien ℓ vil senere få betegnelsen $H((a_1, a_2), b)$. På tegningen herunder er $\alpha = b$.



Sådanne halvrum, eller halvplaner optræder ofte i praktiske og teoretiske modeller idet de udtrykker en begrænsning som skal overholdes. Tænker man på b som budgetbegrænsningen og ens univers af varer man kan købe består af bare to goder med stykpriser på hhv. a_1 og a_2, er forbruget (x_1, x_2) af de to goder begrænset af

$$a_1x_1 + a_2x_2 \leq b\,.$$

Bemærk, at halvrummet også kan udtrykkes som $F^{-1}(]-\infty, b])$ for den affine funktion $F(x_1, x_2) = a_1x_1 + a_2x_2$. Det ses umiddelbart af tegningen at det indre af halvrummet er $\{(x_1, x_2)|a_1x_1 + a_2x_2 < b\} = F^{-1}(]-\infty, b[)$, og at randen er ℓ.

7): Graf for den affine approksimerende funktion til en differentiabel funktion gennem $(x_0, f(x_0))$. Lad $I \subseteq \mathbb{R}$ være et åbent interval, $f : I \to \mathbb{R}$ en differentiabel funktion og $x_0 \in I$. At f er differentiabel i x_0 betyder ifølge sin definition, at for f's Taylorpolynomium af grad 1 i punktet x_0

$$p_1(x) := f(x_0) + f'(x_0)(x - x_0) \tag{$*$}$$

gælder det at

$$\frac{|f(x) - p_1(x)|}{|x - x_0|} \to 0 \text{ for } x \neq x_0, x \to x_0\,.$$

Det vil altså sige, at når f er differentiabel og $x \to x_0$, $x \neq x_0$, så vil $p_1(x)$ approksimere $f(x)$ bedre og bedre jo tættere x kommer på x_0. Denne sidste sætning er meget kvalitativ i sit udsagn, så hvis man ønsker hele emnet behandlet korrekt, bør man konsultere sin Mat. 1–lærebog. Vi kalder derfor p_1 for den approksimerende affine funktion til f i punktet x_0. Ved en omskrivning ses det at $p_1(x)$ også kan udtrykkes ved:

$$p_1(x) = f'(x_0)x + (f(x_0) - f'(x_0)x_0)\,. \tag{$**$}$$

Grafen for $p_1(x)$ bliver ifølge 4) en ret linie, og det er velkendt at denne linie er tangenten til grafen for f i punktet $(x_0, f(x_0))$. Se tegning. Erfaringsmæssigt er det nemmest at huske udtrykket $(*)$ snarere end $(**)$, som formel for den approksimerende affine funktion gennem $(x_0, f(x_0))$. Grunden til dette er at (*) umiddelbart viser at $p_1(x_0) = f(x_0)$ og $p_1'(x_0) = f'(x_0)$, og vi får:

$$\ell = \{(x_1, x_2) \in \mathbb{R}^2 | x_2 = f(x_0) + f'(x_0)(x_1 - x_0)\}\,.$$

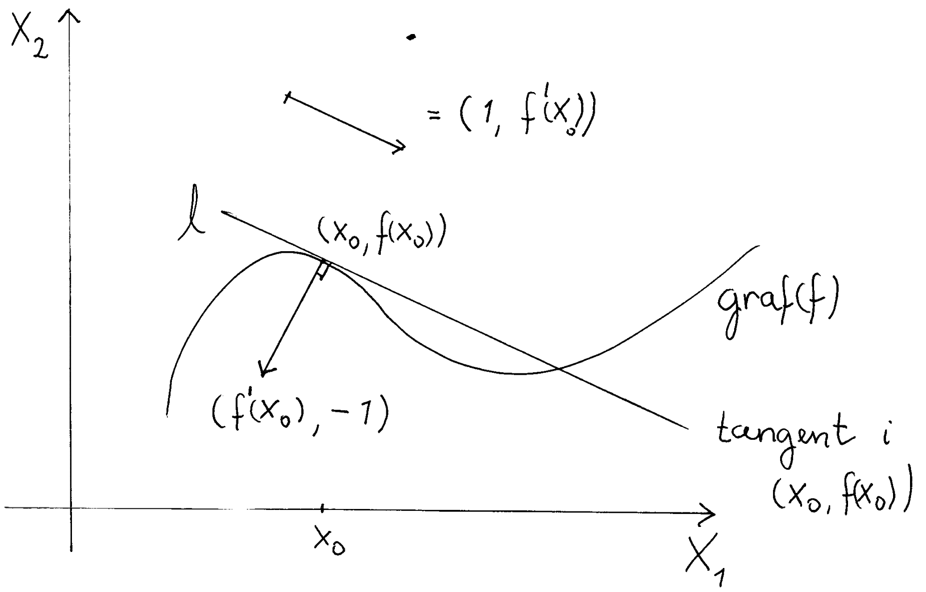
eller

$$\ell = \{(x_1, x_2) \in \mathbb{R}^2 | x_2 = p_1(x_1)\}\,.$$

8): Tangent til grafen for f i punktet $(x_0, f(x_0))$.

Emnet er næsten færdigbehandlet under 7), men der er 2 pointer som man bør erindre sig. **1** Normalvektor til tangenten for grafen for f i $(x_0, f(x_0))$ er $(-f'(x_0), 1)$ eller også $(f'(x_0), -1)$. Vi bruger den sidste tiest, da det er dette udtryk der vil blive generaliseret til funktioner af flere variable.

2 Bemærk, at f er en funktion af 1 variabel, men dens graf er en delmængde af $\mathbb{R}^2$. Tangenten forløber i $\mathbb{R}^2$, men den er graf for en affin funktion af en variabel, nemlig den approksimerende affine funktion gennem $(x_0, f(x_0))$ med hældning $f'(x_0)$.



4 Hyperplaner i $\mathbb{R}^n$, $n \geq 2$

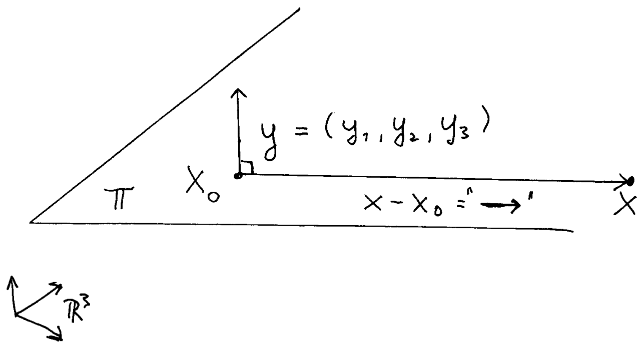
Generaliseringen af foregående afsnit til $\mathbb{R}^n$ med $n \geq 3$ benytter et sprog der er baseret på tilfældet $n = 3$. Under læsningen kan det være hensigtsmæssigt at tænke på specialtilfældet $n = 3$, idet de universelle argumenter som teksten benytter sig af så kan understøttes af almindelig rumgeometri samt skitser af rummet.

Vi starter med at rekapitulere lidt om planer i rummet $\mathbb{R}^3$. Det der kan nævnes herunder kan ord til andet generaliseres til de såkaldte hyperplaner i $\mathbb{R}^n$, som er definerede ved

Definition 4.1. Lad $y = (y_1, \ldots, y_n) \in \mathbb{R}^n \setminus \{(0, \ldots, 0)\}$ og $\alpha \in \mathbb{R}$. Hyperplanen $H(y, \alpha)$ er defineret ved

$$\begin{aligned} H(y, \alpha) &:= \{x \in \mathbb{R}^n | \langle y, x \rangle = \alpha\} \\ &= \{x \in \mathbb{R}^n | y \cdot x = \alpha\} . \end{aligned}$$

Analogien til planer i rummet er at en plan i rummet er karakteriseret ved sin normalretning (lodret), lad os sige $y = (y_1, y_2, y_3) \neq (0, 0, 0)$ og et enkelt punkt i planen $x_0 = (x_1^0, x_2^0, x_3^0)$.



Tegningen skal illustrere at når man kender en normalvektor $y = (y_1, y_2, y_3)$ til planen Π og et punkt x_0 i planen Π, så gælder

$$\begin{aligned} \forall x \in \mathbb{R}^3 : x \in \Pi &\iff (x - x_0) \perp y \\ &\iff (x - x_0) \cdot y = 0 \\ &\iff x \cdot y = x_0 \cdot y \\ &\iff x \in H(y, x_0 \cdot y) . \end{aligned}$$

Udtrykket $x \in \Pi \iff (x \cdot y = x_0 \cdot y)$ kan også omformes til at

$$\begin{aligned} \Pi &= \{x \in \mathbb{R}^3 | x \cdot y = x_0 \cdot y\} \\ &= \{x \in \mathbb{R}^3 | y_1 x_1 + y_2 x_2 + y_3 x_3 = x_0 \cdot y\} . \end{aligned}$$

Derfor kaldes udtrykket $y_1x_1 + y_2x_2 + y_3x_3 = x_0 \cdot y$ for planen Π's ligning.

Vender vi nu tilbage til $n \in \mathbb{N}$, $n \geq 2$, ser vi at for $y \in \mathbb{R}^n \setminus \{0\}$ og $\alpha \in \mathbb{R}$ gælder

$$\forall x \in \mathbb{R}^n : x \in H(y,\alpha) \iff y_1x_1 + y_2x_2 + \cdots + y_nx_n = \alpha\,,$$

så det sidste udtryk kaldes også ligningen for hyperplanen $H(y,\alpha)$. Endvidere ses det at hvis $x_0 \in H(y,\alpha)$, så er $y \cdot x_0 = \alpha$ og vi får regninger analoge til dem på forrige side, men nu for vilkårligt $n \in \mathbb{N}$, $n \geq 2$

$$\begin{aligned}\forall x \in \mathbb{R}^n : x \in H(y,\alpha) &\iff x \in H(y, x_0 \cdot y)\\ &\iff x \cdot y = x_0 \cdot y\\ &\iff y \cdot (x - x_0) = 0\\ &\iff y \perp (x - x_0)\,.\end{aligned}$$

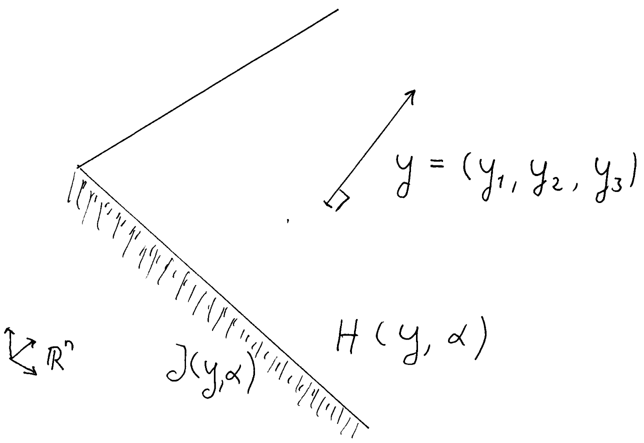
Dvs. $H(y,\alpha) = x_0 + y^{\perp}$, og hyperplanen $H(y,\alpha)$ er altså bestemt blot man kender dens normalretning og et punkt x_0 i hyperplanen.

Før vi vender os til generaliseringen er det hensigtsmæssigt at introducere begrebet, et halvrum.

Definition 4.2. Lad $y \in \mathbb{R}^n \setminus \{0\}$ og $\alpha \in \mathbb{R}$. Mængden $J(y,\alpha)$ givet ved

$$J(y,\alpha) = \{x \in \mathbb{R}^n | y \cdot x \leq \alpha\}$$

kaldes et halvrum.



De 8 aspekter som vi vil betragte hyperplaner under vil blive delt i 2 grupper. Den første gruppe 1) $- \cdots -$ 5) nedenfor, handler om hyperplaner i $\mathbb{R}^n$. Det sidste af disse aspekter, dvs. nr 5), handler om niveaumængden for en affin funktion af n–variable, og vi har så taget første skridt i undersøgelsen af sammenhænge mellem funktioner af n variable og hyperplaner i $\mathbb{R}^{n+1}$. Vi bliver nødt til at skifte til $\mathbb{R}^{n+1}$, da grafen for en reel funktion af n variable er en delmængde af $\mathbb{R}^{n+1}$. Nummerordenen er ikke helt den samme som i afsnit 3.

Aspekterne: **1 − 5** $H(y,\alpha)$, $y \neq 0$, $y \in \mathbb{R}^n$.

1) Et matematisk objekt.

2) Et affint underrum af $\mathbb{R}^n$ af dimension $n-1$.

3) Det affine hylster af n affint uafhængige punkter i $\mathbb{R}^n$.

4) Randen af et halvrum.

5) Niveaumængde for en ikke konstant affin funktion $F : \mathbb{R}^n \to \mathbb{R}$.

Aspekterne: **6 − 8** $H(y,\alpha)\,,\ y \in \mathbb{R}^{n+1}\,,\ y_{n+1} \neq 0\,,\ \alpha \in \mathbb{R}\,.$

6) Graf for en affin funktion $F : \mathbb{R}^n \to \mathbb{R}$.

7) Graf for den approksimerende affine funktion til en differentiabel funktion gennem $(x_0, f(x_0))$.

8) Tangentplan for grafen for en differentiabel funktion af n-variable i et punkt på grafen.

1) Et matematisk objekt: Er diskuteret ovenfor.

2)Et affint underrum af dimension $n-1$:

Som i tilfældet $n = 2$ ses det, at for

$$x_0 := \frac{\alpha}{\|y\|^2} y$$

gælder $x_0 \in H(y,\alpha)$, så af udregningerne fra før fås

$$H(y,\alpha) = x_0 + y^\perp\,.$$

Da $y^\perp$ er et lineært underrum af dimension $(n-1)$, ses det at $H(y,\alpha)$ er et affint underrum af dimension $(n-1)$.

Opgave:

a) Vis, at $y^\perp$ er et underrum.

b) Vis, at dimensionen af $y^\perp$ er $(n-1)$ ved hjælp af dimensionssætningen anvendt på den lineære afbildning $\mathbb{R}^n \ni x \to f_y(x) \in \mathbb{R}$.

c) Find et andet argument for at $\dim(y^\perp) = n - 1$.

Opgave:
Vis, at et affint underrum af $\mathbb{R}^n$ af dimension $(n-1)$ er en hyperplan. *Vink.* Tænk på $\mathbb{R}^3$, og hvordan y så skal findes! Tænk dernæst på at finde α.

3) Det affine hylster af n affint uafhængige punkter:

Lad $H(y,\alpha) = x_0 + y^\perp$ som vist under 2). Da er $y^\perp$ et lineært underrum af dimension $(n-1)$, og $y^\perp$ har derfor en basis $(f_1, \dots, f_{n-1})$. Punkterne definerede ved:

$$x_0\,,\ x_1 := x_0 + f_1, \dots, x_i := x_0 + f_i, \dots, x_{n-1} := x_0 + f_{n-1}$$

udgør netop n affint uafhængige punkter i $H(y,\alpha)$ ifølge sætning 2.15 At disse n punkter har $H(y,\alpha)$ som affint hylster følger af

$$\begin{aligned} H(y,\alpha) &\supseteq \operatorname{aff}(x_0,\dots,x_{n-1}) \supseteq x_0 + \text{ span } \{x_1 - x_0, \dots, x_{n-1} - x_0\} \\ &= x_0 + \text{ span } \{f_1, \dots, f_{n-1}\} = H(y,\alpha)\,. \end{aligned}$$

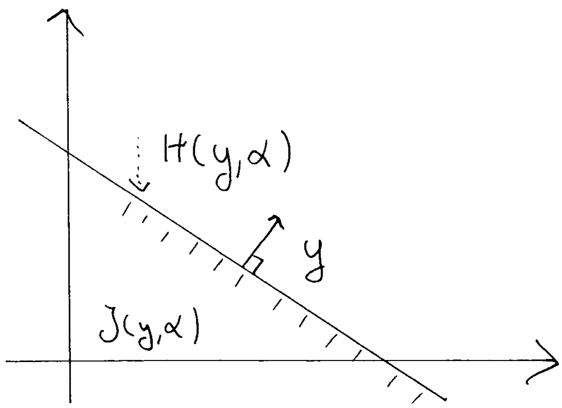
Lad omvendt $\{x_0, x_1, \dots, x_{n-1}\}$ være n affint uafhængige punkter, da fås fra sætning 2.15 at $x_1 - x_0, \dots, x_{n-1} - x_0$ er et lineært uafhængigt sæt af $(n-1)$ vektorer, og $\operatorname{aff}(x_0, \dots, x_{n-1})$
$= x_0 + \text{span } (x_1 - x_0, \dots, x_{n-1} - x_0)$. Dvs.

$$\begin{aligned} \dim(\operatorname{aff}(x_0,\dots,x_{n-1})) &= \dim(\text{span } (x_1 - x_0, \dots, x_{n-1} - x_0)) \\ &= n - 1\,, \end{aligned}$$

så ifølge 2) er $\operatorname{aff}(x_0, \dots, x_{n-1})$ en hyperplan.

4) Randen af halvrummet $J(y,\alpha) = \{x \in \mathbb{R}^n | y \cdot x \leq \alpha\}$:

Når vi skal illustrere dette begreb gøres det som regel på baggrund af situationen $n = 2$



Vi ved at afbildningen $\mathbb{R}^n \ni x \xrightarrow{f_y} f_y(x) \in \mathbb{R}$ er kontinuert (afsnit 1), og intervallet $]-\infty, \alpha[$ er åbent, så for $\mathcal{V}$, delmængden af $\mathbb{R}^n$ givet ved

$$\mathcal{V} := \{x \in \mathbb{R}^n \mid y \cdot x < \alpha\} = f_y^{-1}(]-\infty, \alpha[)$$

fås $\mathcal{V}$ åben og, analogt, $J(y,\alpha) = f_y^{-1}(]-\infty,\alpha])$ er afsluttet. Det er endvidere klart at $H(y,\alpha) = J(y,\alpha)\setminus\mathcal{V} = f^{-1}(\{\alpha\})$ er afsluttet, så for at vise at $H(y,\alpha)$ netop er randen af $J(y,\alpha)$ skal vi blot overbevise os om, at $H(y,\alpha)$ ingen indre punkter har.

Lad da $x \in H(y,\alpha)$ og $\varepsilon > 0$, da gælder for

$$z := x + \frac{\varepsilon}{2\|y\|}y\,,\ z \in K(x,\varepsilon) \text{ og } z\cdot y = x\cdot y + \frac{\varepsilon}{2}\|y\| > \alpha\,.$$

Dvs. $z \notin J(y,\alpha)$ så $x \notin J(y,\alpha)^0$, og

$$J(y,\alpha)^0 = \mathcal{V} = \{x \in \mathbb{R}^n \mid y\cdot x < \alpha\} \text{ og } \partial(J(y,\alpha)) = H(y,\alpha).$$

5) Niveaumængden for en ikke konstant affin funktion:

Når $y \in \mathbb{R}^n\setminus\{0\}$ er givet, er den lineære funktion $f_y : \mathbb{R}^n \longrightarrow \mathbb{R}$ ikke konstant da $f_y(y) = \|y\|^2 \neq 0$, og vi ser, at $H(y,\alpha) = f_y^{-1}(\alpha)$.

Lad nu $F : \mathbb{R}^n \to \mathbb{R}$ være en ikke konstant affin funktion. Ifølge Sætning 2.10 findes $A \in \mathbf{M}(1,n)$, $\beta \in \mathbb{R}$ så $F(x) = Ax + \beta$. Da F ikke er konstant, må $A \neq 0$, $A = (y_1,\ldots,y_n) \in \mathbb{R}^n\setminus(\{0\})$. Følgelig er $F^{-1}(\alpha)$ givet ved

$$F^{-1}(\alpha) = \{x|(y_1,\ldots,y_n)\cdot(x_1,\ldots,x_n)+\beta = \alpha\} = H(y,\alpha-\beta)\,.$$

OBS: Under omtalen af aspekterne 6, 7 og 8 er $y \in \mathbb{R}^{n+1}$ hvor tillige $y_{n+1} \neq 0$

6) Graf for en affin afbildbning $F : \mathbb{R}^n \to \mathbb{R}$:

I 5) så vi at en affin funktion $F : \mathbb{R}^n \to \mathbb{R}$ har formen $F(x) = a\cdot x + \beta$ for et $a \in \mathbb{R}^n$. Vi udelukker ikke konstante afbildninger mere, så $a = (0,\ldots,0) \in \mathbb{R}^n$ er tilladt.

Grafen for F,betegnes $G(F)$ og er som bekendt givet ved

$$\begin{aligned} G(F) &= \{(x,t) \in \mathbb{R}^{n+1}|x \in \mathbb{R}^n \text{ og } t = F(x)\}\\ &= \{(x,t) \in \mathbb{R}^{n+1}|x \in \mathbb{R}^n \text{ og } t = a\cdot x + \beta\}\\ &= \{(x,t) \in \mathbb{R}^{n+1}|a\cdot x - t = -\beta\}\\ &= \{(x,t) \in \mathbb{R}^{n+1}|(a_1,\ldots,a_n,-1)\cdot(x_1,\ldots,x_n,t) = -\beta\}\\ &= H(y,-\beta) \text{ hvor } y \in \mathbb{R}^{n+1},\ y = (a_1,\ldots,a_n,-1) \text{ og } \ y_{n+1} = -1 \neq 0\,. \end{aligned}$$

7) Graf for den approksimerende affine funktion i punktet $(x_0, f(x_0))$

Lad $D \subseteq \mathbb{R}^n$ være en åben mængde $x_0 \in D$ og $f : D \to \mathbb{R}$ en differentiabel funktion.

Fra Mat 1 vides det, at den approksimerende affine funktion $p_1(x)$ for $f(x)$ i punktet $(x_0, f(x_0))$ er givet ved

$$p_1(x) = f(x_0) + \bigtriangledown f(x_0)\cdot(x - x_0).$$

Differentiabiliteten af $f(x)$ i punktet x_0 betyder netop at der gælder:

$$\lim_{\substack{x\to x_0\\ x\neq x_0}} \frac{|f(x)-p_1(x)|}{\|x-x_0\|} = 0\,.$$

Dette betyder at overgangen mellem grafen for f og grafen for p_1 bliver glat i omegnen af punktet $(x_0, f(x_0))$. I den matematiske disciplin *Differentialgeometri* lærer man at grafen for p_1 netop bliver tangentplanen for $f's$ graf i punktet $(x_0, f(x_0))$, men her kan vi kun se det gennem ovenstående grænseværdi.

Ifølge 6) er grafen for p_1 en hyperplan i $\mathbb{R}^{n+1}$ og ved at gentage regningerne fra 6) i denne sammenhæng fås

$$p_1(x) = f(x_0) + \left(\frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_n}(x_0)\right) \cdot (x - x_0)$$
$$\left(\frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_n}(x_0)\right) \cdot x - p_1(x) = \bigtriangledown f(x_0) \cdot x_0 - f(x_0)$$
$$\left(\frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_n}(x_0), -1\right) \cdot (x, p_1(x)) = \bigtriangledown f(x_0) \cdot x_0 - f(x_0)$$

og vi ser

$$\text{Graf } (p_1) = H\left(\left(\frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_n}(x_0), -1\right), \bigtriangledown f(x_0) \cdot x_0 - f(x_0)\right)$$

8) Tangentplan for en graf for en differentiabel reel funktion af n reelle variable:

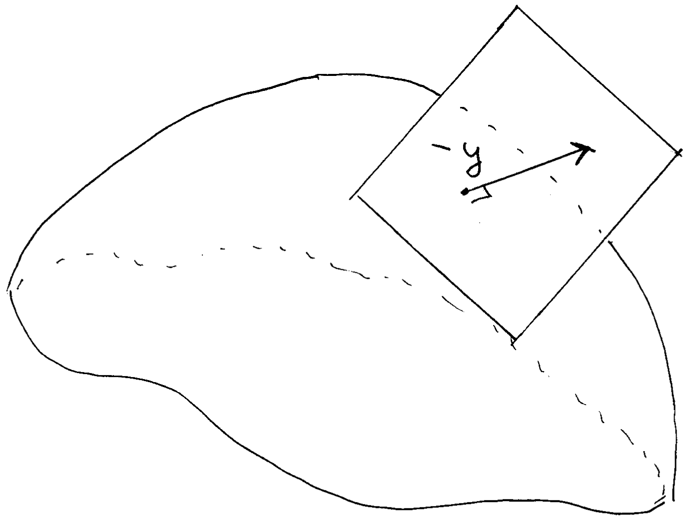
Lad situationen være som i 7), da defineres tangentplanen for $G(f)$ i punktet $(x_0, f(x_0))$ som grafen for den approksimerende affine funktion gennem $(x_0, f(x_0))$.

Af 7) aflæses, at normalvektoren til tangentplanen i $(x_0, f(x_0))$ er

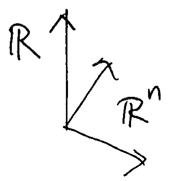
$$\left(\frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_n}(x_0), -1\right),$$

og da tangentplanen går gennem punktet $(x_0, f(x_0))$ ses det igen, at denne plan må være nedenstående mængde

$$\{(x,t) \in \mathbb{R}^{n+1} | ((x,t) - (x_0, f(x_0)) \perp (\bigtriangledown f(x_0), -1)\}$$
$$= H((\bigtriangledown f(x_0), -1), \bigtriangledown f(x_0) \cdot x_0 - f(x_0)).$$



OBS: $-y$ tegnet.



Tangentplan til grafen i $(x_0, f(x_0))$

normalvektor $y = \left(\frac{\partial f}{\partial x_1}(x_0), \dots, \frac{\partial f}{\partial x_n}(x_0), -1\right)$

Tangentplanen selv er graf for

$$P_1(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0)$$

5 Lidt om konvekse mængder og konvekse funktioner

I kapitel 0 defineredes konveksitet ved at alle forbindelseslinier mellem punkter i en mængde $\mathcal{C}$ er helt indeholdt i $\mathcal{C}$. I kapitel 2 defineredes et affint underrum af $\mathbb{R}^n$ ved at det er en delmængde af $\mathbb{R}^n$ som er stabil overfor affine kombinationer af elementer fra mængden (Definition 2.2). Siden vistes det så i Sætning 2.3, at en delmængde af $\mathbb{R}^n$ er et affint underrum, netop hvis det med 2 punkter a_0 og a_1 fra mængden, indeholder mængden hele linien $\ell(a_0, a_1)$. Altså definitionen er abstrakt, men sætningen viser at begrebet er konkret geometrisk. Vi vil følge samme rute her, og først definere konveksitet abstrakt, for så umiddelbart derefter i Sætning 5.3 at vise, at konveksitet netop er det konkrete begreb vi mødte i kapitel 0, og som er nævnt ovenfor.

Definition 5.1. Lad $k \in \mathbb{N}_0$; $x_0, \ldots, x_k \in \mathbb{R}^n$; $\lambda_0, \ldots, \lambda_k \in \mathbb{R}$. Hvis alle $\lambda_i \geq 0$ og $\lambda_0 + \cdots + \lambda_k = 1$, kaldes linearkombinationen

$$\lambda_0 x_0 + \cdots + \lambda_k x_k$$

en konveks kombination af $x_0, \cdots, x_k$.

Bemærk, at en konveks kombination også er en affin kombination, og at de konvekse kombinationer netop er de affine kombinationer, hvor alle koefficienterne er ikke–negative.

Definition 5.2. En delmængde $\mathcal{C}$ af $\mathbb{R}^n$ siges at være konveks hvis det for enhver konveks kombination $\lambda_0 c_0 + \cdots + \lambda_k c_k$ af elementer fra $\mathcal{C}$ gælder, at $\lambda_0 c_0 + \cdots + \lambda_k c_k \in \mathcal{C}$.

Som bemærket under emnet affine underrum, betyder en sådan definition at $\emptyset$, den tomme mængde, er konveks. Der er nemlig ingen konvekse kombinationer af elementer fra $\emptyset$, så udsagnet gælder for alle konvekse kombinationer af elementer fra $\emptyset$. I analogi med Sætning 2.3 gælder der følgende sætning

Sætning 5.3. *Lad $\mathcal{C}$ være en delmængde af $\mathbb{R}^n$, da er $\mathcal{C}$ konveks hvis og kun hvis*

$$\forall c_0, c_1 \in \mathcal{C}\ \forall \lambda \in [0,1] : (1-\lambda)c_0 + \lambda c_1 \in \mathcal{C}\,.$$

Bevis. Hvis $\mathcal{C}$ er konveks er det klart at $(1-\lambda)c_0+\lambda c_1 \in \mathcal{C}$, da udtrykket $(1-\lambda)c_0+\lambda c_1$ er en konveks kombination af 2 elementer fra $\mathcal{C}$. Antager vi nu at $(1-\lambda)c_0+\lambda c_1 \in \mathcal{C}$ når $c_0, c_1 \in \mathcal{C}$ og $0 \leq \lambda \leq 1$, så kan man ved induktion vise, at alle udsagnene $\mathcal{U}_k$ som defineres nedenfor gælder.

For $k \in \mathbb{N}$ defineres udsagnet $\mathcal{U}_k$ ved

$$\begin{aligned}\mathcal{U}_k : &\text{ Enhver konveks kombination } \lambda_0 c_0 + \cdots + \lambda_k c_k \\ &\text{af } \quad k+1 \text{ elementer fra } \mathcal{C} \text{ er indeholdt i } \mathcal{C}\,.\end{aligned}$$

Da $\mathcal{U}_1$ er antaget sand lige ovenfor er induktionsstarten er klar. Induktionstrinnet $\mathcal{U}_k \implies \mathcal{U}_{k+1}$ for $k \geq 1$ følger ord til andet som i beviset for Sætning 2.3. Prøv selv at gentage argumenterne, og sætningen er vist. $\square$

Den næste sætning viser hvorledes konveksitet harmonerer med forskellige operationer på delmængder af $\mathbb{R}^n$.

Sætning 5.4.

1) *Lad $\mathcal{C}$ og $\mathcal{D}$ være konvekse delmængder af $\mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$, så er $\alpha\mathcal{C}+\beta\mathcal{D}$ en konveks delmængde af $\mathbb{R}^n$.*

2) *Lad $\mathcal{C} \subseteq \mathbb{R}^n$, $\mathcal{D} \subseteq \mathbb{R}^m$ og $\mathcal{C} \times \mathcal{D} \subseteq \mathbb{R}^{n+m}$*

$$\mathcal{C} \times \mathcal{D} = \{(c_1, \ldots, c_n, d_1, \ldots, d_m) | (c_1, \ldots, c_n) \in \mathcal{C}, (d_1, \ldots, d_m) \in \mathcal{D}\}\,.$$

Der gælder, hvis $\mathcal{C} \neq \emptyset$ og $\mathcal{D} \neq \emptyset$

$$(\mathcal{C} \text{ og } \mathcal{D} \text{ konvekse }) \iff \mathcal{C} \times \mathcal{D} \text{ er konveks}\,.$$

3) *Lad $(\mathcal{C}_i)_{i \in I}$ være en familie af konvekse delmængder af $\mathbb{R}^n$, da er*

$$\mathcal{C} := \bigcap_{i \in I} \mathcal{C}_i$$

en konveks delmængde af $\mathbb{R}^n$.

4) *Lad $\mathcal{C}$ være en konveks delmængde af $\mathbb{R}^n$ og $F : \mathbb{R}^n \to \mathbb{R}^m$ en affin afbildning, da er $F(\mathcal{C})$ en konveks delmængde af $\mathbb{R}^m$.*

5) *Lad $\mathcal{D} \subseteq \mathbb{R}^m$ være konveks og $F : \mathbb{R}^n \to \mathbb{R}^m$ en affin afbildning, da er $F^{-1}(\mathcal{D})$ en konveks delmængde af $\mathbb{R}^n$.*

6) *Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være konveks, så er afslutningen $\overline{\mathcal{C}}$ også konveks.*

7) *Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være konveks, så er det indre $\overset{\circ}{\mathcal{C}}$ også konveks.*

Beviser:

Ad 1) Lad $x_0 = \alpha c_0 + \beta d_0 \in \alpha\mathcal{C} + \beta\mathcal{D}$, $x_1 = \alpha c_1 + \beta d_1 \in \alpha\mathcal{C} + \beta\mathcal{D}$ og $\lambda \in [0,1]$, så er

$$\begin{aligned}(1-\lambda)x_0 + \lambda x_1 &= (1-\lambda)(\alpha c_0 + \beta d_0) + \lambda(\alpha c_1 + \beta d_1)\\ &= \alpha((1-\lambda)c_0 + \lambda c_1) + \beta((1-\lambda)d_0 + \lambda d_1)\end{aligned}$$

Da $\mathcal{C}$ og $\mathcal{D}$ er konvekse fås $(1-\lambda)c_0+\lambda c_1 \in \mathcal{C}$ og $(1-\lambda)d_0+\lambda d_1 \in \mathcal{D}$, så $(1-\lambda)x_0+\lambda x_1 \in \alpha\mathcal{C} + \beta\mathcal{D}$, og ifølge Sætning 5.3 er $\alpha\mathcal{C} + \beta\mathcal{D}$ konveks.

Ad 2) Antag at $\mathcal{C}$ og $\mathcal{D}$ er konvekse og ikke–tomme, og lad $x_0 = (c_1^0, \dots, c_n^0, d_1^0, \dots, d_m^0)$ samt $x_1 = (c_1^1, \dots, c_n^1, d_1^1, \dots, d_m^1)$ være elementer fra $\mathcal{C}\times\mathcal{D}$. For $0 \le \lambda \le 1$ fås, idet $c^0 := (c_1^0, \dots, c_n^0) \in \mathcal{C}$, $c^1 := (c_1^1, \dots, c_n^1) \in \mathcal{C}$, $d^0 := (d_1^0, \dots, d_m^0) \in \mathcal{D}$, $d^1 := (d_1^1, \dots, d_m^1) \in \mathcal{D}$. Med denne notation fås $x_0 = (c^0, d^0)$ og $x_1 = (c^1, d^1)$ i $\mathcal{C}\times\mathcal{D}$ og ved koordinatvise regninger ses det at

$$\begin{aligned}(1-\lambda)x_0 + \lambda x_1 &= ((1-\lambda)c_1^0 + \lambda c_1^1, \dots, (1-\lambda)c_n^0 + \lambda c_n^1,\\ &\qquad (1-\lambda)d_1^0 + \lambda d_1^1, \dots, (1-\lambda)d_m^0 + \lambda d_m^1)\\ &= ((1-\lambda)c^0 + \lambda c^1, (1-\lambda)d^0 + \lambda d^1) \in \mathcal{C}\times\mathcal{D}\,,\end{aligned}$$

så $\mathcal{C}\times\mathcal{D}$ er konveks.

Antages omvendt $\mathcal{C}\times\mathcal{D}$ konveks, kan vi indse at $\mathcal{C}$ er konveks først, og dernæst ved symmetri — der er ingen forskel på $\mathcal{C}$ og $\mathcal{D}$'s roller — slutte, at også $\mathcal{D}$ er konveks. Lad da $c_0, c_1 \in \mathcal{C}$, $\lambda \in [0,1]$ og $d \in \mathcal{D}$ da $\mathcal{D} \neq \emptyset$, så er $x_0 := (c_0, d) \in \mathcal{C}\times\mathcal{D}$ og $x_1 := (c_1, d) \in \mathcal{C}\times\mathcal{D}$, så da $\mathcal{C}\times\mathcal{D}$ er konveks, fås $(1-\lambda)x_0 + \lambda x_1 \in \mathcal{C}\times\mathcal{D}$. Nu er

$$\begin{aligned}(1-\lambda)x_0 + \lambda x_1 &= ((1-\lambda)c_0 + \lambda c_1, (1-\lambda)d + \lambda d)\\ &= ((1-\lambda)c_0 + \lambda c_1, d) \in \mathcal{C}\times\mathcal{D}\,,\end{aligned}$$

så $(1-\lambda)c_0 + \lambda c_1 \in \mathcal{C}$, og $\mathcal{C}$ er konveks.

Ad 3) Antag $c_0, c_1 \in \bigcap_{i\in I} \mathcal{C}$ og $\lambda \in [0,1]$, så gælder det for alle $i \in I$ at $(1-\lambda)c_0 + \lambda c_1 \in \mathcal{C}_i$, da $c_0, c_1 \in \mathcal{C}_i$ og $\mathcal{C}_i$ er konveks. Altså $(1-\lambda)c_0 + \lambda c_1 \in \bigcap_{i\in I} \mathcal{C}_i$ og $\bigcap_{i\in I} \mathcal{C}_i$ er konveks

Ad 4) Lad $\mathcal{C} \subseteq \mathbb{R}^n$ konveks, $F : \mathbb{R}^n \to \mathbb{R}^m$ affin $c_0, c_1 \in \mathcal{C}$ og $d_0 = F(c_0)$, $d_1 = F(c_1)$, $0 \le \lambda \le 1$. Da fås

$$(1-\lambda)d_0 + \lambda d_1 = (1-\lambda)F(c_0) + \lambda F(c_1) = F((1-\lambda)c_0 + \lambda c_1)\,.$$

Da $\mathcal{C}$ er konveks er $(1-\lambda)c_0 + \lambda c_1 \in \mathcal{C}$, så $(1-\lambda)d_0 + \lambda d_1 \in F(\mathcal{C})$, og da c_0, c_1 var tilfældigt valgte, ses det at $F(\mathcal{C})$ er en konveks delmængde af $\mathbb{R}^m$.

Ad 5) Lad $c_0, c_1 \in F^{-1}(\mathcal{D})$, hvor $F : \mathbb{R}^n \to \mathbb{R}^m$ er affin og $\mathcal{D} \subseteq \mathbb{R}^m$ er konveks, så gælder for $d_0 := F(c_0)$ og $d_1 := F(c_1)$ at $d_0, d_1 \in \mathcal{D}$. For et $\lambda \in [0,1]$ fås $F((1-\lambda)c_0 + \lambda c_1) = (1-\lambda)d_0 + \lambda d_1 \in \mathcal{D}$, da F er affin og $\mathcal{D}$ er konveks. Følgelig er $(1-\lambda)c_0 + \lambda c_1 \in F^{-1}(\mathcal{D})$ og $F^{-1}(\mathcal{D})$ er konveks.

Ad 6) Lad $d_0, d_1 \in \overline{\mathcal{C}}$, $\lambda \in [0,1]$ hvor $\mathcal{C} \subseteq \mathbb{R}^n$ er en konveks mængde. Da findes følger $(c_n^0)_{n\in\mathbb{N}}$, $(c_n^1)_{n\in\mathbb{N}}$ så c_n^0 og c_n^1 ligger i $\mathcal{C}$ for alle n i $\mathbb{N}$, og

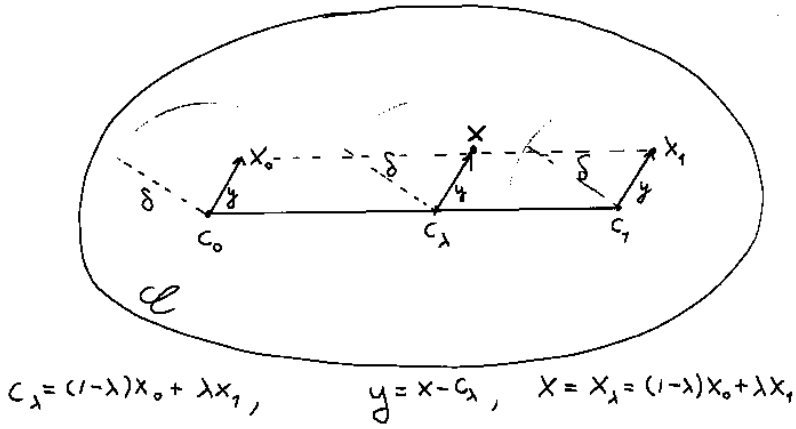
$$c_n^0 \to d_0\,, \quad c_n^1 \to d_1 \quad \text{for} \quad n \to \infty\,.$$

Da $\mathcal{C}$ er konveks er $x_n := (1-\lambda)c_n^0 + \lambda c_n^1 \in \mathcal{C}$, men

$$x_n = (1-\lambda)c_n^0 + \lambda c_n^1 \to (1-\lambda)d_0 + \lambda d_1 \text{ for } n \to \infty\,,$$

så $(1-\lambda)d_0 + \lambda d_1 \in \overline{\mathcal{C}}$ og $\overline{\mathcal{C}}$ er konveks.

Ad 7) Hvis $\overset{\circ}{\mathcal{C}} = \emptyset$, så er $\overset{\circ}{\mathcal{C}}$ konveks. Vi kan derfor antage at $\mathcal{C} \subseteq \mathbb{R}^n$ er konveks med ikke tomt indre for derefter at vise, at $\overset{\circ}{\mathcal{C}}$ er konveks. Lad $c_0, c_1 \in \overset{\circ}{\mathcal{C}}$ og $\lambda \in [0,1]$. Da c_0 og c_1 er indre punkter findes $\delta_0 > 0$, $\delta_1 > 0$, så $K(c_0, \delta_0) \subseteq \mathcal{C}$ og $K(c_1, \delta_1) \subseteq \mathcal{C}$. Lad da $\delta := \min\{\delta_0, \delta_1\}$, så er $K(c_0, \delta) \subseteq \mathcal{C}$ og $K(c_1, \delta) \subseteq \mathcal{C}$, og vi har en situation som nedenfor.



Lad nu $x \in K(c_\lambda, \delta)$, vi skal så indse at $x \in \mathcal{C}$ og vil dermed have vist, at $K(c_\lambda, \delta) \subseteq \mathcal{C}$, så c_λ er et indre punkt i $\mathcal{C}$. Til den ende defineres $y := x - c_\lambda$ så $\|y\| < \delta$ og $x_0 := c_0 + y$, $x_1 := c_1 + y$. Heraf fås $x_0 \in K(c_0, \delta) \subseteq \mathcal{C}$, $x_1 \in K(c_1, \delta) \subseteq \mathcal{C}$, så $(1-\lambda)x_0 + \lambda x_1 \in \mathcal{C}$, men $(1-\lambda)x_0 + \lambda x_1 = (1-\lambda)c_0 + \lambda c_1 + y = x$ så $x \in \mathcal{C}$ som ønsket. $\square$

Eksempel 5.5. Dette eksempel har til formål at gøre opmærksom på en teknik som kan anvendes i adskillige sammenhænge. Teknikken er baseret på resultaterne 5.4 2) og 5.4 4) og består i at repræsentere mange konvekse mængder som én ved hjælp af 5.4 2), og vise egenskaber om dem alle ved at anvende en affin afbildning på denne produktmængde. For at være konkret vises resultatet i Sætning 5.4 1) ved hjælp af en sådan teknik. Som bekendt siger 5.4 1), at $\alpha\mathcal{C} + \beta\mathcal{D}$ er konveks når $\mathcal{C}$ og $\mathcal{D}$ er konvekse. Lad os nu antage at $\mathcal{C}$ og $\mathcal{D}$ er ikke–tomme konvekse mængder indeholdt i $\mathbb{R}^n$, så er ifølge 5.4 2) $\mathcal{C} \times \mathcal{D} = \{(c,d) \in \mathbb{R}^{2n} | c \in \mathcal{C}, d \in \mathcal{D}\}$ en konveks delmængde af $\mathbb{R}^{2n}$. Skrives vektorerne i $\mathbb{R}^{2n}$ som (x,y) med $x \in \mathbb{R}^n$ og $y \in \mathbb{R}^n$, defineres der en lineær afbildning $F : \mathbb{R}^{2n} \to \mathbb{R}^n$ ved

$$F(x,y) := \alpha x + \beta y\,.$$

Det er forhåbentligt let at indse, at $F(\mathcal{C} \times \mathcal{D}) = \alpha\mathcal{C} + \beta\mathcal{D}$, så fra 5.4 4) fås, at $\alpha\mathcal{C} + \beta\mathcal{D}$ er konveks.

Vort næste begreb er det konvekse hylster af en mængde $\mathcal{M} \in \mathbb{R}^n$. Hvis $\mathcal{M} \subseteq \mathbb{R}^2$ er en begrænset delmængde af planen kan man forstille sig, at man slår en snor stramt om alle punkterne; det der ligger indenfor det område som snoren omslutter er det konvekse hylster.
Begrundelsen for eksistensen af begrebet det konvekse hylster er baseret på 2 ting. For det første er $\mathbb{R}^n$ en konveks mængde, og for det andet, er vilkårlige fællesmængder af konvekse mængder en konveks mængde ifølge 5.4 2). Lader vi derfor $\mathcal{M} \subseteq \mathbb{R}^n$, så findes der konvekse delmængder $\mathcal{C}$ af $\mathbb{R}^n$ så $\mathcal{C} \supseteq \mathcal{M}$, og fællesmængden af alle sådanne konvekse mængder er igen en konveks mængde som indeholder $\mathcal{M}$. Da denne fællesmængde er indeholdt i enhver anden konveks mængde som indeholder $\mathcal{M}$, må dette være den mindste konvekse delmængde af $\mathbb{R}^n$ som indeholder $\mathcal{M}$, og vi kan derfor definere denne mængde som det konvekse hylster af $\mathcal{M}$.

Definition 5.6. Lad $\mathcal{M} \subseteq \mathbb{R}^n$. Ved det konvekse hylster — $\mathrm{conv}(\mathcal{M})$ — af $\mathcal{M}$ forstås den mindste konvekse delmængde af $\mathbb{R}^n$, som indeholder $\mathcal{M}$.

I argumenterne umiddelbart over definitionen er der argumenteret for at $\mathrm{conv}(\mathcal{M})$ er fællesmængden af alle konvekse delmængder af $\mathbb{R}^n$ som indeholder $\mathcal{M}$, men denne karakterisation er ligesom definitionen ikke anvendelig, når man skal afgøre om et givet punkt $x \in \mathbb{R}^n$ ligger i $\mathrm{conv}(\mathcal{M})$ eller ej. For at klare dette problem skaffer vi os matematiske hjælpeværktøjer analoge til Sætning 2.13.

Vi starter med et lemma — en hjælpesætning — som ikke er dyb, ej heller svær at vise, men som ikke desto mindre er ganske nyttig at have i sin håndbagage.

Lemma 5.7. *Lad $\mathcal{M} \subseteq \mathbb{R}^n$, $x = \alpha_0 x_0 + \cdots + \alpha_k x_k$ og $y = \beta_0 y_0 + \cdots + \beta_m y_m$ konvekse kombinationer af elementer $x_0, \ldots, x_k, y_0, \ldots, y_m$ fra $\mathcal{M}$. For $\lambda \in [0,1]$ gælder $(1-\lambda)x + \lambda y$ er en konveks kombination af elementer fra $\mathcal{M}$.*

Bevis. Lad $z = (1-\lambda)x + \lambda y = (1-\lambda)\alpha_0 x_0 + \cdots + (1-\lambda)\alpha_k x_k + \lambda\beta_0 y_0 + \cdots + \lambda\beta_m y_m$, så er z en linearkombination af elementer fra $\mathcal{M}$ hvor alle koefficienterne er ikke–negative. På den anden side er summen af koefficienterne 1, så z er en konveks kombination af elementer fra $\mathcal{M}$. □

Vi kan så vise, at $\mathrm{conv}(\mathcal{M})$ netop består af alle konvekse kombinationer af elementer fra $\mathcal{M}$, eller i formelsprog:

Sætning 5.8. *Lad $\mathcal{M} \subseteq \mathbb{R}^n$, da gælder*

$$\operatorname{conv}(\mathcal{M}) = \left\{ \sum_{i=0}^{k} \lambda_i m_i \middle| k \in \mathbb{N}_0, \lambda_i \geq 0, \lambda_0 + \cdots + \lambda_k = 1, m_i \in \mathcal{M} \right\}.$$

Bevis. Da $\mathcal{M} \subseteq \operatorname{conv}(\mathcal{M})$ og $\operatorname{conv}(\mathcal{M})$ er konveks, må alle konvekse kombinationer af elementer fra $\mathcal{M}$ ligge i $\operatorname{conv}(\mathcal{M})$ ifølge Definition 5.2, så

$$\operatorname{conv}(\mathcal{M}) \supseteq \left\{ \sum_{i=0}^{k} \lambda_i m_i \middle| k \in \mathbb{N}_0, \lambda_i \geq 0, \lambda_0 + \cdots + \lambda_k = 1, m_i \in \mathcal{M} \right\}.$$

På den anden side ses det at højresiden ovenfor indeholder $\mathcal{M}$ (for $k = 0$), og fra Lemma 5.7 følger det at højresiden er konveks, så højresiden må indeholde $\operatorname{conv}(\mathcal{M})$, da denne sidste er indeholdt i enhver konveks mængde som indeholder $\mathcal{M}$. $\square$

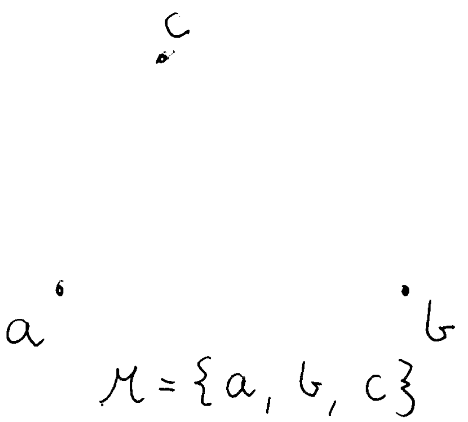
Med denne karakterisering af $\operatorname{conv}(\mathcal{M})$ kunne man tro at man fik brug for meget lange konvekse kombinationer, eller mere præcist, at der ikke ville være nogen overgrænse for længden "$k + 1$" af de konvekse kombinationer der skal anvendes for at fremstille $\operatorname{conv}(\mathcal{M})$. Det viser sig imidlertid at $k \leq n$ rækker, dvs. vi behøver højst konvekse kombinationer af $(n + 1)$ elementer.

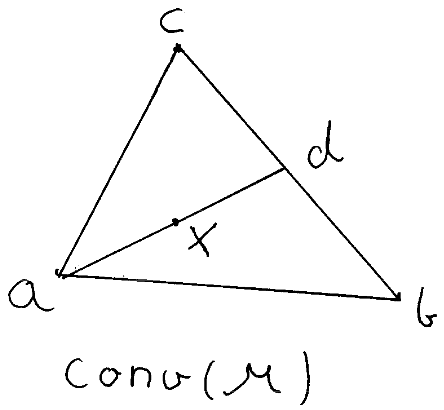
Sætning 5.9 (Carathéodory). *Lad $\mathcal{M} \subseteq \mathbb{R}^n$ og $x \in \operatorname{conv}(\mathcal{M})$, da findes $k \leq n$, $m_0, \ldots, m_k \in \mathcal{M}$, $\alpha_0, \ldots, \alpha_k \in \mathbb{R}$, $\alpha_i \geq 0$, $\alpha_0 + \cdots + \alpha_k = 1$, så $x = \alpha_0 m_0 + \cdots + \alpha_k m_k$.*

Argumentet i beviset er baseret på at et sæt af punkter $x_0, \ldots, x_k$ med mere end $n+1$ elementer er affint afhængigt.

Inden beviset bliver præsenteret, er det måske værd at kigge på et eksempel i 2 dimensioner.

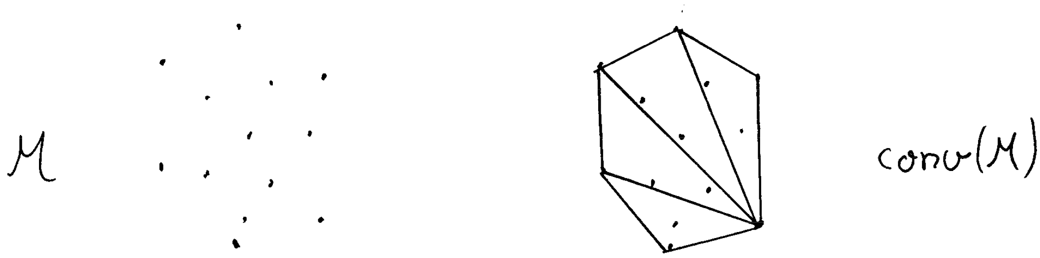
Eksempel 5.10. Betragt 3 punkter a, b, c i $\mathbb{R}^2$ som ikke ligger på linie.





og lad x være et indre punkt i trekanten med hjørnerne a, b, c; da kan x skrives som en konveks kombination af punkterne a, b og c. For at indse dette bemærkes det at punktet d er en konveks kombination af c og b, eftersom $d \in [c, b]$. Da d er valgt som

skæringspunktet mellem linierne $\ell(a, x)$ og $\ell(c, b)$ ligger x på stykket $[a, d]$, og er derfor en konveks kombination af a og d. I henhold til det "trivielle" Lemma 5.7, er x så en konveks kombination af a, b og c.
Tager man nu en endelig mængde $\mathcal{M}$



ses det ved induktion, at ethvert punkt i conv $(\mathcal{M})$ ligger i en trekant, hvis hjørner ligger i $\mathcal{M}$. Med baggrund i ovenstående argument for trekanten som det konvekse hylster af sine hjørner, ses det da at ethvert punkt i conv $(\mathcal{M})$ er en konveks kombination af højst 3 punkter fra $\mathcal{M}$.

Vi tillader os endnu en digression inden beviset for Sætning 5.9, idet jeg godt vil gøre opmærksom på, at det principale argument i nedenstående bevis bruges i mange andre sammenhænge, herunder specielt i forbindelse med beviset for at hvis et lineært program har en såkaldt tilladt løsning så har det også en tilladt basisløsning.

Bevis for Sætning 5.9 Lad $x \in \operatorname{conv}(\mathcal{M})$, da kan x skrives som en konveks kombination af $x = \alpha_0 m_0 + \cdots + \alpha_k m_k$. Blandt de mulige sådanne fremstillinger må der være en (eller eventuelt flere) som benytter færrest mulige elementer i summen. Lad nu $(k+1)$ være det mindst mulige antal summander i en konveks kombination $x = \alpha_0 m_0 + \cdots + \alpha_k m_k$, så må alle $\alpha_i > 0$. Antag nu, at $k + 1 \geq n + 2$, så må punkterne $m_0, \ldots, m_k$ være affint afhængige, og der må findes $\lambda_0, \ldots, \lambda_k$ som ikke alle er nul, så

$$\lambda_0 m_0 + \cdots + \lambda_k m_k = 0 \quad \text{og} \quad \lambda_0 + \cdots + \lambda_k = 0\,.$$

Da ikke alle $\lambda_i = 0$ og $\lambda_0 + \cdots + \lambda_k = 0$, må der være både positive λ_j ($\lambda_j > 0$) og negative λ_j ($\lambda_j < 0$). Dette vil vi udnytte til at konstruere en konveks kombination af $m_0, \ldots, m_k$ som fremstiller x men ikke benytter alle $m_0, \ldots, m_k$. Til den ende betragter vi for $s > 0$ den affine kombination

$$\begin{aligned}(\alpha_0 + s\lambda_0)m_0 + \cdots + (\alpha_k + s\lambda_k)m_k &= \alpha_0 m_0 + \cdots + \alpha_k m_k + 0 = x\\ (\alpha_0 + s\lambda_0) + \cdots + (\alpha_k + s\lambda_k) &= 1 + 0 = 1\,.\end{aligned}$$

Hvis $\lambda_i \geq 0$, så er $\alpha_i + s\lambda_i \geq \alpha_i > 0$ for alle $s > 0$. Hvis $\lambda_i < 0$ er $\alpha_i + s\lambda_i \geq 0$ når $0 < s < -\frac{\alpha_i}{\lambda_i}$.

Dvs. for $\hat{s} = \min\left\{\frac{-\alpha_i}{\lambda_i} \middle| \lambda_i < 0\right\}$ fås alle $\alpha_i + \hat{s}\lambda_i \geq 0$, men for i_0 med $\hat{s} = \frac{-\alpha_{i_0}}{\lambda_{i_0}}$ fås

$$\alpha_{i_0} + \hat{s}\lambda_{i_0} = 0\,,$$

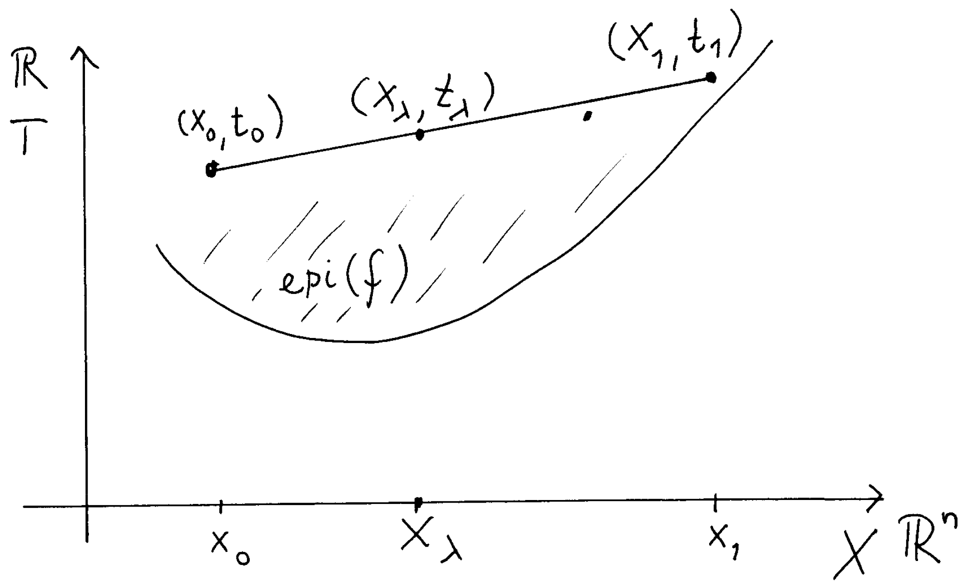
så
$$(\alpha_0 + \hat{s}\lambda_0)m_0 + \cdots + (\alpha_k + \hat{s}\lambda_k)m_k = x$$
er en konveks kombination af elementer fra $\mathcal{M}$ der fremstiller x og højst benytter k elementer. Altså er antagelsen "$k+1 \geq n+2$" forkert og vi kan slutte, at alle $x \in \operatorname{conv}(\mathcal{M})$ kan skrives som en konveks kombination af højst $n+1$ elementer fra $\mathcal{M}$. $\square$

Ligesom vi i afsnit 0 kort benyttede konveksitietsteorien til at udlede simple resultater om konvekse funktioner, vil vi her prøve at samle lidt teori om disse funktioner op, ud fra de fundne resultater.

Som direkte anvendelse af Sætning 5.3 fås

Sætning 5.11 (Jensens Ulighed, Generelt). *Lad $\mathcal{D} \subseteq \mathbb{R}^n$ være en konveks mængde og $f : \mathcal{D} \to \mathbb{R}$ en reel funktion. Funktionen er konveks hvis og kun hvis der gælder*
$$\begin{aligned}
&\forall k \in \mathbb{N}\ \forall \alpha_0, \ldots, \alpha_k \in \mathbb{R}\,,\ \alpha_i \geq 0\,,\ \alpha_0 + \cdots + \alpha_k = 1 \\
&\forall d_0, \ldots, d_k \in \mathcal{D} : \\
&f(\alpha_0 d_0 + \cdots + \alpha_k d_k) \leq \alpha_0 f(d_0) + \cdots + \alpha_k f(d_k)\,.
\end{aligned}$$

Bevis. Hvis f er konveks er epi(f) konveks, og da punkterne $(d_0, f(d_0)), \ldots, (d_k, f(d_k))$ alle ligger i epi(f) fås
$$(\alpha_0 d_0 + \cdots + \alpha_k d_k, \alpha_0 f(d_0) + \cdots + \alpha_k f(d_k) \in \operatorname{epi}(f)\,,$$
eller med andre ord
$$f(\alpha_0 d_0 + \cdots + \alpha_k d_k) \leq \alpha_0 f(d_0) + \cdots + \alpha_k f(d_k)\,.$$
Antager vi på den anden side at Jensens generelle ulighed gælder for f, så skal vi ifølge Sætning 5.3 kun vise, at for alle par af punkter (x_0, t_0) og (x_1, t_1) fra epi(f) er $[(x_0, t_0), (x_1, t_1)] \subseteq \operatorname{epi}(f)$.



Tegningen viser argumentet som i bogstaver forløber således:

Lad $0 \leq \lambda \leq 1$ og $(x_\lambda, t_\lambda) = (1-\lambda)(x_0, t_0) + \lambda(x_1, t_1) = ((1-\lambda)x_0 + \lambda x_1, (1-\lambda)t_0 + \lambda t_1)$. Vi skal da indse, at $t_\lambda \geq f(x_\lambda)$. Fra antagelsen om Jensens uligheds gyldighed fås

$$\begin{aligned} f(x_\lambda) &= f((1-\lambda)x_0 + \lambda x_1) \leq (1-\lambda)f(x_0) + \lambda f(x_1) \\ &\leq (1-\lambda)t_0 + \lambda t_1 = t_\lambda\,. \end{aligned}$$

□

Det næste resultat er en anvendelse af Sætning 5.4 3) på supremumdannelse af konvekse funktioner. Det viser sig at et sådant supremum er en konveks funktion igen.

Sætning 5.12. *Lad $\mathcal{D} \subseteq \mathbb{R}^n$ være en konveks mængde og $(f_i)_{i\in I}$ en familie af konvekse funktioner $f_i : \mathcal{D} \to \mathbb{R}$. Hvis det for alle d i $\mathcal{D}$ gælder at $\sup_{i\in I} f_i(d) < \infty$, defineres der en konveks funktion $f : \mathcal{D} \to \mathbb{R}$ ved $f(d) = \sup_{i\in I} f_i(d)$.*

Bevis. Man kan indse at $\mathrm{epi}(f)$ er konveks ved at vise at $\mathrm{epi}(f) = \bigcap_{i\in I} \mathrm{epi}(f_i)$, og dernæst udnytte konveksiteten af alle mængderne $\mathrm{epi}(f_i)$ samt Sætning 5.4 3).

Betingelsen $\sup f_i(d) < \infty$ betyder, at f er veldefineret som en reel funktion på $\mathcal{D}$ ved $f(d) = \sup_{i\in I} f_i(d)$. Lad da $(d, t) \in \mathrm{epi}(f)$, så fås

$$t \geq f(d) \geq f_i(d) \text{ for alle } i \in I \implies (d, t) \in \bigcap_{i\in I} \mathrm{epi}(f_i)$$

og $\mathrm{epi}(f) \subseteq \bigcap_{i\in I} \mathrm{epi}(f_i)$. Antages omvendt $(d, t) \in \bigcap_{i\in I} \mathrm{epi}(f_i)$ fås

$$\forall i \in I : t \geq f_i(d) \implies t \geq \sup f_i(d) = f(d)\,,$$

så $(d, t) \in \mathrm{epi}(f)$ og $\bigcap_{i\in I} \mathrm{epi}(f_i) \subseteq \mathrm{epi}(f)$. Altså er $\mathrm{epi}(f) = \bigcap_{i\in I} \mathrm{epi}(f_i)$ og dermed konveks, så $f = \sup_{i\in I} f_i$ er en konveks funktion. □

Den sidste sætning i dette afsnit er en af de lettere at vise, men ikke desto mindre meget god at huske, da den er anvendelig i mange sammenhænge.

Sætning 5.13.

1) *En affin funktion $F : \mathbb{R}^n \to \mathbb{R}$ er konveks.*

2) *Hvis $F_1, \ldots, F_k : \mathbb{R}^n \to \mathbb{R}$ er affine, så er $G : \mathbb{R}^n \to \mathbb{R}$*

$$G(x) := \max_{1\leq i\leq k} F_i(x)$$

konveks.

3) Hvis $F : \mathbb{R}^n \to \mathbb{R}^m$ er affin og $f : \mathbb{R}^m \to \mathbb{R}$ er konveks, så er $f \circ F$ konveks.

Bevis.

Ad 1) En affin funktion $F : \mathbb{R}^n \to \mathbb{R}$ opfylder Jensens Ulighed, med *lighedstegn* endda.

Ad 2) Betingelsen i Sætning 5.12 — $\sup_{i \in I} f_i(x) < \infty$ — er klart opfyldt, da der kun er endligt mange affine funktioner $F_1, \ldots, F_k$ i familien.

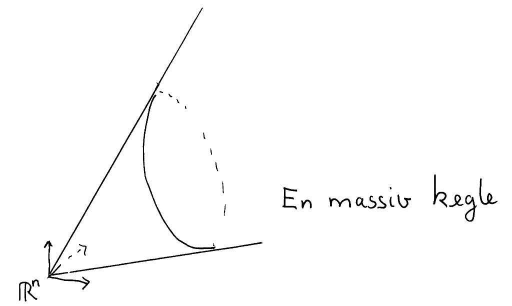
Ad 3) Jensens Ulighed ses klart at være opfyldt, idet affiniteten af F giver lighedstegnet herunder og konveksiteten af f det efterføgende ulighedstegn.

$$\begin{aligned} f(F((1-\lambda)x_0 + \lambda x_1)) &= f((1-\lambda)F(x_0) + \lambda F(x_1)) \\ &\leq (1-\lambda)f(F(x_0)) + \lambda f(F(x_1))\,, \end{aligned}$$

så $f \circ F$ er konveks. $\square$

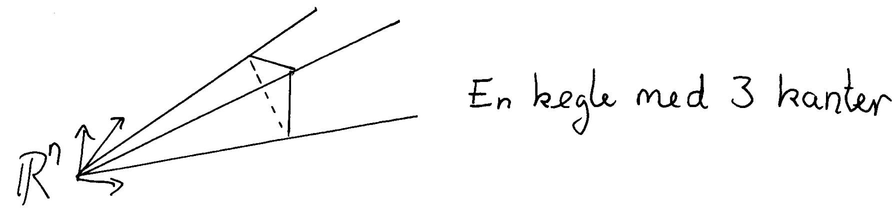
6 Konvekse kegler og recessionskeglen

Emnet for dette afsnit kan man forestille sig som et uendeligt, men massivt, kræmmerhus.



afsat med spidsen i origo, 0.

Billedet af kræmmerhuset antyder at overfladen er glat og krummet. Dette er imidlertid ikke tilfældet således som det næste eksempel antyder



Her er antydet en konveks kegle som er fællesmængden af 3 halvrum.

Strukturen af den teoretiske gennemgang af emnet konvekse kegler, ligner dem der blev anvendt ved behandlingen af affine underrum (Kapitel 2) og konvekse mængder (Kapitel 5), så vi starter med definitionen af den type af linearkombinationer som er karakteriserende for konvekse kegler.

Definition 6.1. Lad $k \in \mathbb{N}$, $x_1, \ldots, x_k \in \mathbb{R}^n$, $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$. Hvis det for alle $i \in \{1, \ldots, k\}$ gælder, at $\lambda_i \geq 0$ kaldes linearkombinationen $\lambda_1 x_1 + \cdots + \lambda_k x_k$ en positiv linearkombination.

Definition 6.2. Lad $\mathcal{K}$ være en delmængde af $\mathbb{R}^n$. Hvis 0–origo ligger i $\mathcal{K}$ og det for enhver positiv linearkombination $\lambda_1 x_1 + \cdots + \lambda_k x_k$ af elementer fra $\mathcal{K}$ gælder at denne er et element i $\mathcal{K}$, kaldes $\mathcal{K}$ en konveks kegle.

Denne definition er lidt abstrakt i forhold til det geometriske billede der indledte dette afsnit, så lad os prøve at etablere forbindelsen uformelt før vi introducerer endnu et abstrakt resultat. For det første er konvekse kegler ikke tomme da origo altid er med. Dernæst ses det at origo i sig selv er en konveks kegle som slet ikke er af form som dem der er tegnet tidligere, men det er til gengæld den eneste begrænsede konvekse kegle i $\mathbb{R}^n$. Hvis vi i definitionen holder os til $k = 1$ og 2 ser vi for $k = 1$, at hvis $\mathcal{K}$ er en konveks kegle, så vil for alle $x \in \mathcal{K}$ mængden $\{\lambda x | \lambda \geq 0\}$ være indeholdt i $\mathcal{K}$. For $k = 2$, $x_1, x_2 \in \mathcal{K}$ og $0 \leq \lambda \leq 1$ fås $(1-\lambda)x_1 + \lambda x_2 \in \mathcal{K}$, så en konveks kegle er konveks og indeholder alle positive multipla af enhver af sine vektorer.

Vi er nu godt på vej til nedenstående karakterisering af konvekse kegler. Bemærk i øvrigt de nære forbindelser til aksiomerne for lineære underrum.

Sætning 6.3. *Lad $\mathcal{K}$ være en delmængde af $\mathbb{R}^n$, da er $\mathcal{K}$ en konveks kegle hvis og kun hvis nedenstående 3 betingelser alle er opfyldte.*

KK1 $\quad 0 \in \mathcal{K}$.

KK2 $\quad \forall x_1, x_2 \in \mathcal{K} : x_1 + x_2 \in \mathcal{K}$.

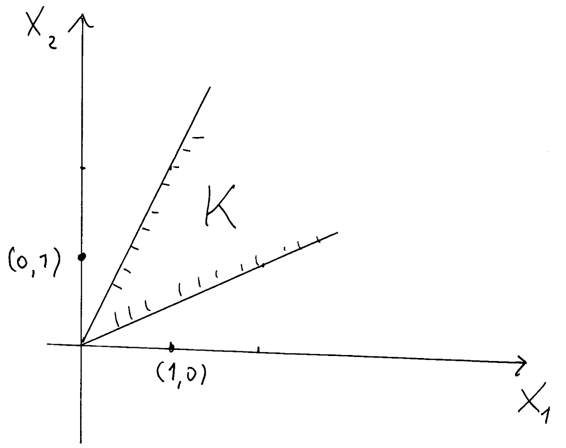
KK3 $\quad \forall \lambda \geq 0 \, \forall x \in \mathcal{K} : \lambda x \in \mathcal{K}$.

Bevis. Det følger umiddelbart af Definition 6.2 at betingelserne alle er nødvendige, dvs. at enhver konveks kegle opfylder betingelserne KK1, KK2 og KK3. Vi skal så blot indse, at betingelserne er tilstrækkelige. Lad $\mathcal{K}$ være en delmængde af $\mathbb{R}^n$ der opfylder de nævnte betingelser, og lad os nu indse, at $\mathcal{K}$ så er en konveks kegle. Vi har umiddelbart fra KK1 at $0 \in \mathcal{K}$. Lad nu $\lambda_1 x_1 + \cdots + \lambda_k x_k$ være en positiv kombination af elementer fra $\mathcal{K}$. I lighed med beviserne for Sætningerne 2.3 og 5.3 kan man ved induktion efter k vise, at sådanne positive kombinationer af elementer fra $\mathcal{K}$ er i $\mathcal{K}$. Jeg skriver ikke alle detaljer, men gør blot opmærksom på, at induktionsstarten er lidt anderledes her. På den anden side følger induktionsstarten direkte fra betingelsen KK3. Induktionstrinnet kommer fra KK3 og KK2, som antydet herunder.

KK3 $\implies \lambda_{k+1} x_{k+1} \in \mathcal{K}$

$\lambda_1 x_1 + \cdots + \lambda_k x_k \in K$ (induktionsantagelsen), og $\lambda_{k+1} x_{k+1} \in \mathcal{K}$ giver ved anvendelse af KK2 at $\lambda_1 x_1 + \cdots + \lambda_k x_k + \lambda_{k+1} x_{k+1} \in \mathcal{K}$. ☐

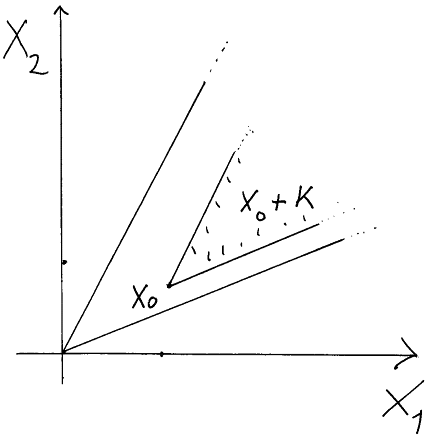
Eksempel 6.4. Lad $\mathcal{K} = \{(x_1, x_2) \in \mathbb{R}^2 | \frac{1}{2} x_1 \leq x_2 \leq 2x_1\}$.



Det ses at $\mathcal{K}$ er fællesmængden af 2 halvrum hvis rande er lineære underrum

$$\mathcal{K} = J((\frac{1}{2}, -1), 0) \cap J((-2, 1), 0)\,.$$

Betingelserne KK1 og KK3 ses umiddelbart at være gyldige. Betingelsen KK2 ses geometrisk lettest at være gyldig fra nedenstående tegning:



$$x_0 = (x_1^0, x_2^0) \text{ vilkårlig i } \mathcal{K}$$
$$x_0 + \mathcal{K} \subseteq \mathcal{K} \implies \text{KK2 gyldig}$$

Konvekse kegler har nogle fundamentale egenskaber som omtales i følgende sætning:

Sætning 6.5.

1) Lad $\mathcal{K} \subseteq \mathbb{R}^n$, $\mathcal{H} \subseteq \mathbb{R}^n$ være konvekse kegler og $\alpha, \beta \in \mathbb{R}$, da er $\alpha\mathcal{K} + \beta\mathcal{H}$ en konveks kegle i $\mathbb{R}^n$.

2) Lad $\mathcal{K} \subseteq \mathbb{R}^n$ og $\mathcal{H} \subseteq \mathbb{R}^m$ være konvekse kegler, da er $\mathcal{K} \times \mathcal{H} \subseteq \mathbb{R}^n \times \mathbb{R}^m$ en konveks kegle.

3) Lad $(\mathcal{K}_i)_{i\in I}$ være en familie af konvekse kegler i $\mathbb{R}^n$, så er $\bigcap_{i\in I} \mathcal{K}_i$ en konveks kegle.

4) Lad $\mathcal{K} \subseteq \mathbb{R}^n$ være en konveks kegle og $F : \mathbb{R}^n \to \mathbb{R}^m$ en lineær *afbildning, så er $F(\mathcal{K})$ en konveks kegle i $\mathbb{R}^m$.*

5) Lad $\mathcal{H} \subseteq \mathbb{R}^m$ være en konveks kegle og $F : \mathbb{R}^n \to \mathbb{R}^m$ en lineær *afbildning, så er $F^{-1}(\mathcal{H})$ en konveks kegle i $\mathbb{R}^n$.*

6) Lad $\mathcal{K} \subseteq \mathbb{R}^n$ være en konveks kegle, så er afslutningen $\overline{\mathcal{K}}$ også en konveks kegle.

Bevis. De fleste argumenter i beviserne for påstandene 1) – 6) er meget nært beslægtede med dem der blev givet i forbindelse med beviset for Sætning 5.4, så gennemgangen her bliver lidt kortfattet, dog med undtagelse af 1), som kan hjælpe læseren lidt i gang.

Ad 1) Det ses, at $0 = \alpha \cdot 0 + \beta \cdot 0$, så $0 \in \alpha\mathcal{K} + \beta\mathcal{H}$, så KK1, OK. For $x_1 = \alpha k_1 + \beta h_1$ og $x_2 = \alpha k_2 + \beta h_2$ i $\alpha\mathcal{K} + \beta\mathcal{H}$ fås $x_1 + x_2 = \alpha(k_1 + k_2) + \beta(h_1 + h_2) \in \alpha\mathcal{K} + \beta\mathcal{H}$, da KK2 gælder for $\mathcal{K}$ og $\mathcal{H}$. Heraf ses så at KK2 gælder for $\alpha\mathcal{K} + \beta\mathcal{H}$. For $\lambda \geq 0$ og $x = \alpha k + \beta h \in \alpha\mathcal{K} + \beta\mathcal{H}$ fås $\lambda x = \alpha(\lambda k) + \beta(\lambda h)$. Nu er $\lambda k \in \mathcal{K}$ og $\lambda h \in \mathcal{H}$ da $\mathcal{K}$ og $\mathcal{H}$ opfylder KK3, og det ses at $\alpha\mathcal{K} + \beta\mathcal{H}$ også opfylder KK3. □

Ad 2) Vi identificerer $\mathbb{R}^{n+m}$ med $\mathbb{R}^n \times \mathbb{R}^m$ ved at nummerere koordinaterne i $\mathbb{R}^{n+m}$ som anført herunder

$$(z_1, \ldots, z_{n+m}) = (x_1, \ldots, x_n, y_1, \ldots, y_m)\,,$$

dvs. $z_i = x_i$, $1 \leq i \leq n$, $z_{n+j} = y_j$, $1 \leq j \leq m$. Med denne notation kan vi f.eks. bevise KK2. Lad $z^1 \in \mathcal{K} \times \mathcal{H}$ og $z^2 \in \mathcal{K} \times \mathcal{H}$ med $z^1 = (k_1, h_1), z^2 = (k_2, h_2)$, $z^1 = (k_1^1, \ldots, k_n^1, h_1^1, \ldots, h_m^1)$, $z^2 = (k_1^2, \ldots, k_n^2, h_1^2, \ldots, h_m^2)$, så er

$$\begin{aligned} z^1 + z^2 &= (k_1^1 + k_1^2, \ldots, k_n^1 + k_n^2, h_1^1 + h_1^2, \ldots, h_m^1 + h_m^2) \\ &= (k_1 + k_2, h_1 + h_2) \in \mathcal{K} \times \mathcal{H}\,. \end{aligned}$$

Ad 3) Overlades helt til læseren som øvelse.

Ad 4) Her er det vigtigt at afbildningen F er lineær og ikke blot affin, da kravet KK1; $0 \in F(K)$ ellers ikke kan opfyldes uden videre. Da en lineær afbildning opfylder $F(0) = 0$ og $F(x + y) = F(x) + F(y)$ samt $F(\lambda x) = \lambda F(x)$, er det forhåbentligt klart at $F(\mathcal{K})$ opfylder KK–kravene når $\mathcal{K}$ gør det.

Ad 5) Dette er en anelse mere subtilt end 4) men ikke meget. Abtag f.eks. at $h_1, h_2 \in \mathcal{H}$ og $x_1 \in F^{-1}(h_1)$, $x_2 \in F^{-1}(h_2)$, så er $F(x_1 + x_2) = h_1 + h_2 \in \mathcal{H}$, og dermed $x_1 + x_2 \in F^{-1}(h_1 + h_2) \subseteq F^{-1}(\mathcal{H})$, og $F^{-1}(\mathcal{H})$ opfylder KK2.

Af 6) Det er klart at $0 \in \overline{\mathcal{K}}$, da $0 \in \mathcal{K} \subseteq \overline{\mathcal{K}}$. Lad nu $x_1, x_2 \in \overline{\mathcal{K}}$, så findes følger $(k_m^1)_{m\in\mathbb{N}}$ og $(k_m^2)_{m\in\mathbb{N}}$ med k_1^1, k_m^2 i K, og $k_m^1 \to x_1$, $k_m^2 \to x_2$ for $m \to \infty$. Da $\mathcal{K}$ opfylder KK2 ligger $k_m^1 + k_m^2$ i $\mathcal{K}$; men $k_m^1 + k_m^2 \to x_1 + x_2$ for $m \to \infty$, så $x_1 + x_2 \in \overline{\mathcal{K}}$, og $\overline{\mathcal{K}}$ opfylder KK2. Med hensyn til KK3 vælges til $x \in \overline{\mathcal{K}}$ en følge $(k_m)_{m\in\mathbb{N}}$ med $k_m \in \mathcal{K}$, $k_m \to x$ for $m \to \infty$. For $\lambda \geq 0$ fås, da $\mathcal{K}$ opfylder KK3, at $\lambda k_m \in \mathcal{K}$, men $\lambda k_m \to \lambda x$ for $m \to \infty$, så $\lambda x \in \overline{\mathcal{K}}$, og $\overline{\mathcal{K}}$ er en konveks kegle.

Bemærkning 6.6. Der er ingen udsagn om det indre af en konveks kegle, og grunden er, at denne mængde sjældent er en konveks kegle. Eksempel 6.4 illustrerer situationen godt. Her har keglen et ikke–tomt indre, men 0 er ikke et indre punkt. Hvis 0 er et indre punkt i en kegle i $\mathbb{R}^2$ siger intuitionen (forhåbentligt), at denne kegle må være hele $\mathbb{R}^2$. Dette gælder også i $\mathbb{R}^n$. Vis dette; dvs. vis, at for en konveks kegle $\mathcal{K} \subseteq \mathbb{R}^n$ som har 0 som indre punkt, må der gælde, at $\mathcal{K} = \mathbb{R}^n$.

I analogi med introduktionen af begreberne aff$(\mathcal{M})$ og conv$(\mathcal{M})$ giver Sætning 6.5 3) anledning til at der eksisterer en mindste konvekse kegle som indeholder en given mængde $\mathcal{M}$. Som i de tidligere tilfælde startes argumentet med at sige, at $\mathbb{R}^n$ er en konveks kegle, så for en vilkårlig mængde $\mathcal{M} \subseteq \mathbb{R}^n$ findes der konvekse kegler som indeholder $\mathcal{M}$. Fællesmængden af alle disse er en konveks kegle som indeholder $\mathcal{M}$, og som er indeholdt i enhver konveks kegler der indeholder $\mathcal{M}$. Altså er denne fællesmængde den mindste konvekse kegle i $\mathbb{R}^n$ som indeholder $\mathcal{M}$, og vi kan definere.

Definition 6.7. Lad $\mathcal{M} \subseteq \mathbb{R}^n$. Ved den indhyldende konvekse kegle for $\mathcal{M}$ – cone$(\mathcal{M})$ – forstås den mindste konvekse kegle i $\mathbb{R}^n$ som indeholder $\mathcal{M}$.

I analogi med Sætning 2.13 og Sætning 5.8 kommer der nu i Sætning 6.8 en anden karakterisering af cone$(\mathcal{M})$, som i konkrete situationer kan bruges til at afgøre om et element x i $\mathbb{R}^n$ ligger i cone$(\mathcal{M})$.

Sætning 6.8. *Lad $\mathcal{M} \subseteq \mathbb{R}^n$, da gælder*

$$\operatorname{cone}(\mathcal{M}) = \{0\} \cup \left\{ \sum_{i=1}^{k} \lambda_i m_i \middle| k \in \mathbb{N}\,,\ \lambda_i \geq 0\,,\ m_i \in \mathcal{M} \right\}.$$

Bevis. "$\mathcal{V} \subseteq \operatorname{cone}(\mathcal{M})$"

Lad $\mathcal{V} := \{0\} \cup \left\{ \sum_{i=1}^{k} \lambda_i m_i \middle| k \in \mathbb{N}\,,\ \lambda_i \geq 0\,,\ m_i \in \mathcal{M} \right\}$. Det følger af KK1 at $0 \in \operatorname{cone}(\mathcal{M})$, så for at vise at $\mathcal{V} \subseteq \operatorname{cone}(\mathcal{M})$ er det nok at vise, at $\sum_{i=1}^{k} \lambda_i m_i \in \operatorname{cone}(\mathcal{M})$, når $\lambda_i \geq 0$, $m_i \in \mathcal{M}$. Men $m_i \in \mathcal{M} \subseteq \operatorname{cone}(\mathcal{M})$, så udtrykket $\sum_{i=1}^{k} \lambda_i m_i$ er en positiv linear kombination af elementer fra den konvekse kegle cone$(\mathcal{M})$, og summen ligger derfor ifølge Definition 6.1 i cone$(\mathcal{M})$.

"$\operatorname{cone}(\mathcal{M}) \subseteq \mathcal{V}$"

For $k = 1$ og $\lambda_1 = 1$ ses det at $\mathcal{M} \subseteq \mathcal{V}$, så vi mangler blot at vise, at $\mathcal{V}$ er en konveks kegle for at indse, at cone$(\mathcal{M}) \subseteq \mathcal{V}$. Efter sin definition indeholder $\mathcal{V}$, 0; så $\mathcal{V}$ opfylder KK1. Lad $x_1, x_2 \in \mathcal{V}$,

$$x_1 = \sum_{i=1}^{h} \lambda_i m_i \qquad \lambda_i \geq 0\,,\ m_i \in \mathcal{M} \text{ og et } \ h \in \mathbb{N}$$
$$x_2 = \sum_{j=1}^{k} \mu_j n_j \qquad \mu_j \geq 0\,,\ n_j \in \mathcal{M} \text{ og et } \ k \in \mathbb{N},$$

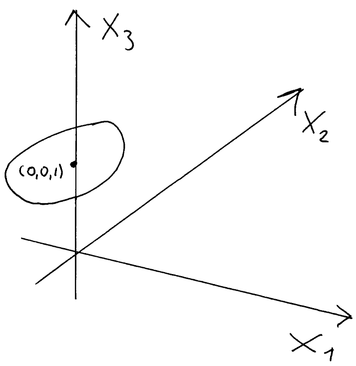
så er $x_1 + x_2 = \sum_{i=1}^{h} \lambda_i m_i + \sum_{j=1}^{k} \mu_j n_j \in \mathcal{V}$, da dette er en positiv kombination af $h + k$ elementer fra $\mathcal{M}$. Heraf følger, at $\mathcal{V}$ tilfredsstiller KK2. At også KK3 gælder, ses umiddelbart af identiteten herunder

$$\lambda\left(\sum_{j=1}^{k}\lambda_j m_j\right)=\sum_{j=1}^{k}(\lambda\lambda_j)m_j\in\mathcal{V}\,.$$

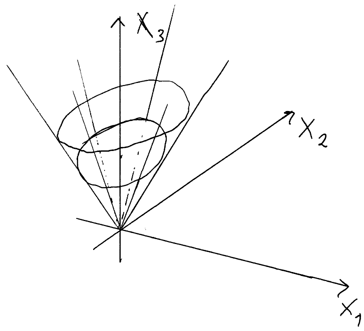
Følgelig er $\mathcal{V}$ en konveks kegle som indeholder $\mathcal{M}$ og dermed også cone $(\mathcal{M})$. Sætningen følger. □

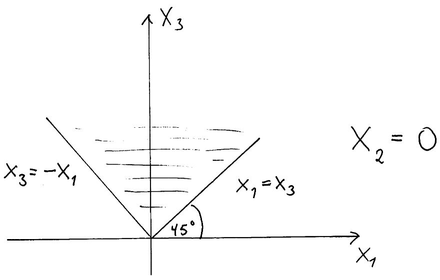
Inden vi fortsætter med det analoge resultat til Carathéodorys Sætning, er det måske rimeligt at regne et eksempel.

Eksempel 6.9. Lad $\mathcal{M} := \{(x_1,x_2,x_3)\in\mathbb{R}^3 | x_3 = 1\,,\, x_1^2+x_2^2=1\}$



En cirkel i planen $x_3 = 1$ med radius 1 og centrum i $(0,0,1)$.
Keglen der frembringes af denne mængde ses på tegningen lige herunder. Den næeste tegning viser snittet mellem keglen og x_1x_3-planen,

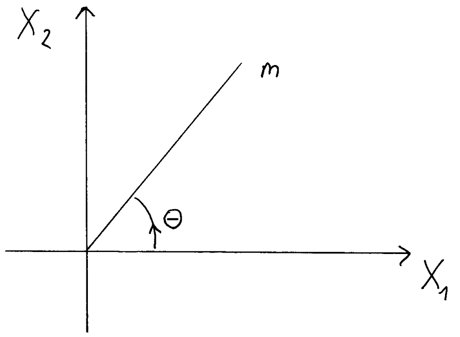




og når den roteres om x_3–aksen fremkommer cone $(\mathcal{M})$. Algebraisk kan keglen beskrives ved

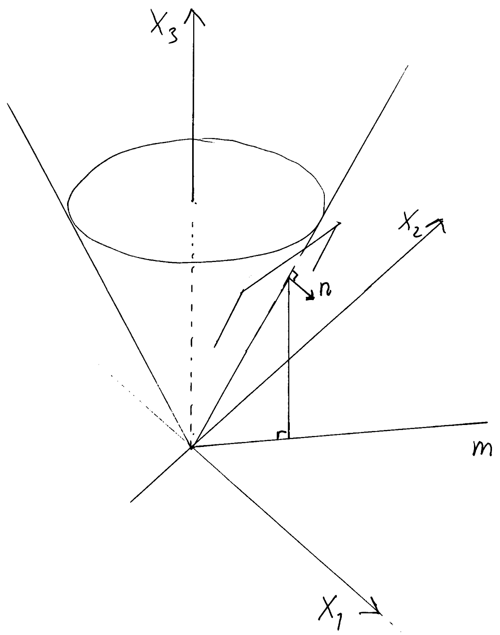
$$\text{cone}\,(\mathcal{M}) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 | x_3 \geq \sqrt{x_1^2 + x_2^2}\}$$

For enhver vinkel $\theta \in [0, 2\pi[$ tegnes halvlinien m ud fra 0 i x_1x_2–planen, som danner vinklen θ med den positive del af x_1–aksen



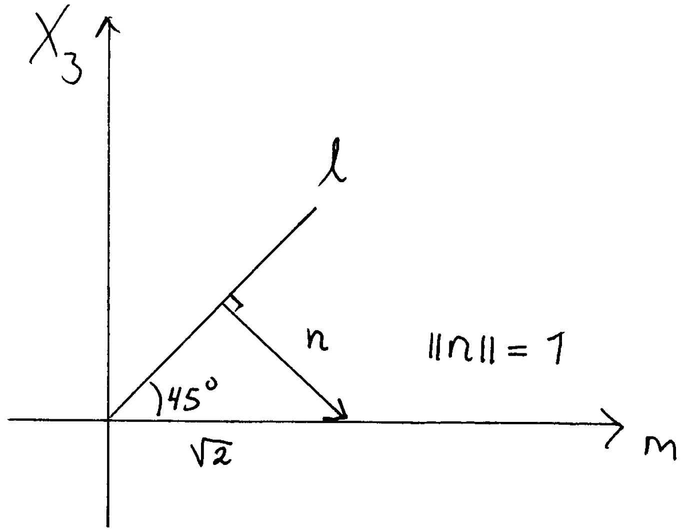
$$m = \{(t\cos\theta, t\sin\theta, 0)\,|t \geq 0\}\,.$$

Lodret over m har keglen cone $(\mathcal{M})$ en randlinie ℓ som vist på tegningen herunder



Det ses at der er en tangentplan til keglen som indeholder ℓ. Normalvektoren til denne tangentplan ligger i planen udspændt af ℓ og x_3–aksen, så tegningen gør det muligt at bestemme en sådan normalvektor af længde 1, kaldet n som

$$n = \left(\frac{\sqrt{2}}{2}\cos\theta, \frac{\sqrt{2}}{2}\sin\theta\,, -\frac{\sqrt{2}}{2}\right)\,.$$



$$n = \left(\frac{\sqrt{2}}{2}\cos\theta, \frac{\sqrt{2}}{2}\sin\theta, -\frac{\sqrt{2}}{2}\right)$$

$$\ell = \left\{ t\left(\frac{\sqrt{2}}{2}\cos\theta, \frac{\sqrt{2}}{2}\sin\theta, \frac{\sqrt{2}}{2}\right) \middle| t \geq 0 \right\}$$

.

Hvis dette er for meget geometri kan man i stedet tænke på den konvekse kegles rand som grafen for funktionen $F : \mathbb{R}^2 \to \mathbb{R}$ givet ved $F(x_1, x_2) = \sqrt{x_1^2 + x_2^2}$. Denne funktion er kontinuert på hele $\mathbb{R}^2$ og differentiabel i $\mathbb{R}^2 \setminus \{(0,0)\}$. I punktet $(\cos\theta, \sin\theta, F(\cos\theta, \sin\theta)) = (\cos\theta, \sin\theta, 1)$ har grafen normalvektoren, eller rettere tangentplanen til grafen, gennem punktet $(\cos\theta, \sin\theta, 1)$ har normalvektoren

$$\left(\frac{\partial F}{\partial x_1}(\cos\theta, \sin\theta), \frac{\partial F}{\partial x_2}(\cos\theta, \sin\theta), -1\right) = (\cos\theta, \sin\theta, -1),$$

som er ensrettet med vektoren n. Vi ser, at for alle $\theta \in [0, 2\pi[$ er cone $(\mathcal{M})$ indeholdt i halvrummet $J((\cos\theta, \sin\theta, -1), 0)$ og cone $(\mathcal{M}) = \bigcap_{\theta\in[0,2\pi[} J((\cos\theta, \sin\theta, -1), 0)$.

Sætning 6.10. *Lad $\mathcal{M}$ være en ikke–tom delmængde af $\mathbb{R}^n$ og $x \in$ cone $(\mathcal{M})$, da findes $k \in \{1, 2, \ldots, n\}$, $m_1, \ldots, m_k$ elementer i $\mathcal{M}$ og $\lambda_1 \geq 0, \ldots, \lambda_k \geq 0$, så $x = \lambda_1 m_1 + \cdots + \lambda_k m_k$.*

Bevis. Fra Sætning 6.8 vides det, at x kan fremstilles som en positiv kombination af $\alpha_1 m_1 + \cdots + \alpha_k m_k$ af elementer fra $\mathcal{M}$. Lad k være det mindste antal summander der fremkommer i en sådan fremstilling af x og antag, at $k \geq n+1$. Da k er mindst mulig må alle $\alpha_i > 0$ for $1 \leq i \leq k$, og da $k \geq n+1$, er sættet $m_1, \ldots, m_k$ lineært afhængigt.

Der findes derfor $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ som ikke alle er nul således at

$$\lambda_1 m_1 + \cdots + \lambda_k m_k = 0\,.$$

Hvis alle $\lambda_i \geq 0$, erstattes λ_i med $-\lambda_i$ (alle), og vi kan derfor tillade os at antage, at mindst et $\lambda_i < 0$. I analogi med beviset for Sætning 5.9 betragtes nu $s > 0$,

$$x = (\alpha_1 + s\lambda_1)m_1 + \cdots + (\alpha_k + s\lambda_k)m_k$$

og $\alpha_i + s\lambda_i \geq \alpha_i > 0$ hvis $\lambda_i \geq 0$. Hvis $\lambda_i < 0$ fås $\alpha_i + s\lambda_i \geq 0$ for $0 < s \leq -\frac{\alpha_i}{\lambda_i}$. Sættes da $s_0 = \min\left\{-\frac{\alpha_i}{\lambda_i}\middle|\lambda_i < 0\right\}$ fås

$$x = (\alpha_1 + s_0\lambda_1)m_1 + \cdots + (\alpha_k + s_0\lambda_k)m_k\,,$$

en positiv kombination som benytter $(k-1)$ eller færre elementer fra $\mathcal{M}$. Antagelsen om at $k \geq n+1$ er derfor forkert og sætningen følger. □

Denne gennemgang af konvekse kegler afsluttes med en introduktion af nogle flere begreber, som vil dukke op i andre sammenhænge senere i kurset. Det første begreb er "Det største lineære underrum" indeholdt i en konveks kegle $\mathcal{K}$. Eksistensen af et sådant sikres af Sætning 6.11.

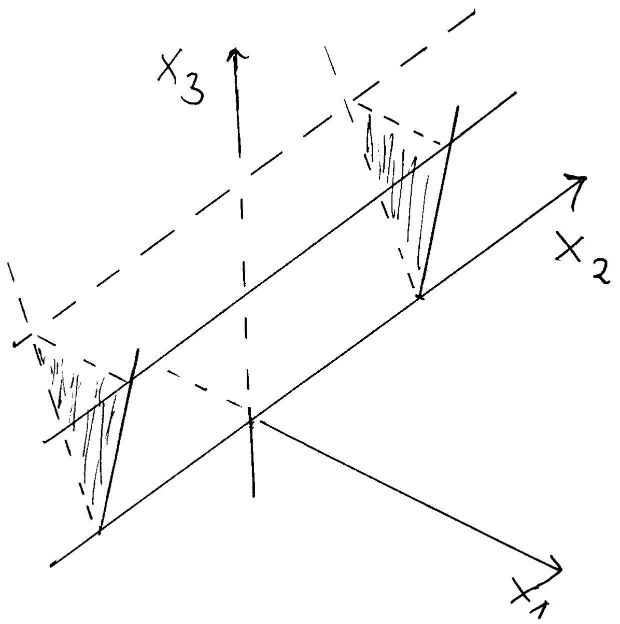
Sætning 6.11. *Lad $\mathcal{K} \subseteq \mathbb{R}^n$ være en konveks kegle, så er $\mathcal{K} \cap (-\mathcal{K})$ et lineært underrum af $\mathcal{K}$, og for ethvert andet lineært underrum $\mathcal{L}$ af $\mathbb{R}^n$, som opfylder $\mathcal{L} \subseteq \mathcal{K}$ gælder $\mathcal{L} \subseteq \mathcal{K} \cap (-\mathcal{K})$.*

Bevis. Ifølge Sætning 6.5 (1)–3)) er $\mathcal{K} \cap (-\mathcal{K})$ en konveks kegle så underrumsaksiomerne for lineære underrum er opfyldte: $\underline{0} \in \mathcal{K} \cap (-\mathcal{K})$; $x_1, x_2 \in \mathcal{K} \cap (-\mathcal{K}) \implies x_1 + x_2 \in \mathcal{K} \cap (-\mathcal{K})$ og $x \in \mathcal{K} \cap (-\mathcal{K})$, $\lambda \geq 0 \implies \lambda x \in \mathcal{K} \cap (-\mathcal{K})$. Lad da $x \in \mathcal{K} \cap (-\mathcal{K})$ og $\lambda < 0$, vi skal så blot vise at $\lambda x \in \mathcal{K} \cap (-\mathcal{K})$ for at indse, at $\mathcal{K} \cap (-\mathcal{K})$ er et lineært underrum. Da $x \in \mathcal{K} \cap (-\mathcal{K})$ er $-x \in \mathcal{K}$ og $-\lambda > 0$, så $\lambda x = (-\lambda)(-x) \in \mathcal{K}$. På den anden side er $x \in \mathcal{K}$, så $\lambda x = -((-\lambda)x) \in -\mathcal{K}$, så $\lambda x \in \mathcal{K} \cap (-\mathcal{K})$ for $\lambda < 0$ også, og $\mathcal{K} \cap (-\mathcal{K})$ er et lineært underrum. Det er forhåbentligt ikke svært at indse, at for et vilkårligt lineært underrum $\mathcal{L}$ af $\mathbb{R}^n$ gælder det at $-\mathcal{L} = \mathcal{L}$. Heraf ses, at hvis $\mathcal{L}$ er et lineært underrum og $\mathcal{L} \subseteq \mathcal{K}$, så fås $-\mathcal{L} = \mathcal{L} \subseteq \mathcal{K}$ og $\mathcal{L} \subseteq -\mathcal{K}$, dvs. $\mathcal{L} \subseteq \mathcal{K} \cap (-\mathcal{K})$, og sætningen følger. □

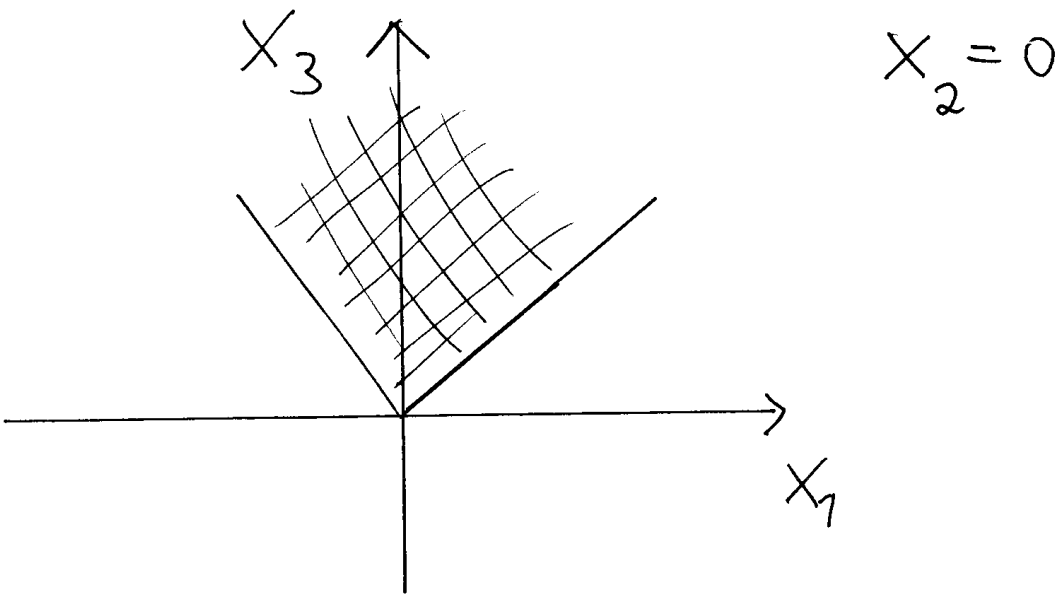
Eksempel 6.12. Lad $\mathcal{K} \subseteq \mathbb{R}^3$ være givet ved

$$\begin{aligned}\mathcal{K} &= J((1,0,-1),0) \cap J((-1,0,-1),0) \\ &= \{(x_1,x_2,x_3) \in \mathbb{R}^3 | x_1 - x_3 \leq 0 \text{ og } -x_1 - x_3 \leq 0\}\,.\end{aligned}$$

Altså en fællesmængde af 2 halvrum som har 0 som randpunkt. Her kan x_2 variere frit og begrænsningerne i x_1, x_3 er forsøgt antydet nedenfor



en uendelig lang og massiv bjælke der står på sin ene kant, med denne placeret over x_2–aksen.



Keglen skåret op ved $x_2 = 0$. Keglen fortsætter med samme snit for alle faste værdier af x_2.

Det ses, at

$$K \cap -K = \{(0, s, 0) | s \in \mathbb{R}\},$$

samt at

$$\begin{aligned} K &= (K \cap -K) + K \cap (K \cap -K)^{\perp} = \\ &= x_2 - \text{aksen } + \{(x_1, 0, x_3) | x_3 \geq x_1 \text{ og } x_3 \geq -x_1\}. \end{aligned}$$

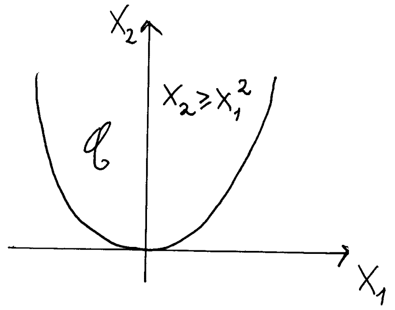
Dette sidste er et generelt fænomen som vi vil vende tilbage til i afsnittet om ubegrænsede konvekse mængder. Først introduceres begrebet “recessionskeglen for en konveks mængde”. Dette begreb er af stor betydning ved studiet af ubegrænsede konvekse mængder.

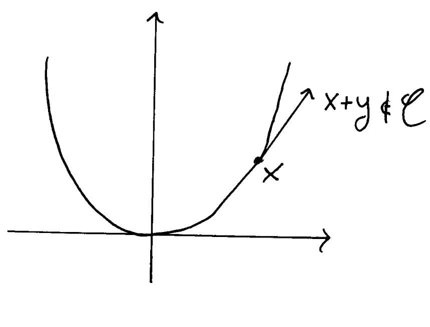
Definition 6.13. Lad $\mathcal{C}$ være en konveks delmængde af $\mathbb{R}^n$. Ved recessionskeglen — $\operatorname{recc}(\mathcal{C})$ — forstås delmængden af $\mathbb{R}^n$ givet ved

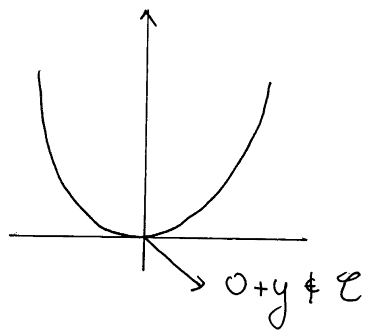
$$\operatorname{recc}(\mathcal{C}) := \{y \in \mathbb{R}^n | \forall x \in \mathcal{C} : x + y \in \mathcal{C}\} = \{y \in \mathbb{R}^n | y + \mathcal{C} \subseteq \mathcal{C}\}.$$

Bemærk, at $\operatorname{recc}(\emptyset) = \mathbb{R}^n$ og hvis $0 \in \mathcal{C}$ så gælder der $\operatorname{recc}(\mathcal{C}) \subseteq \mathcal{C}$. Et lille argument, som du selv kan lave, viser at $\operatorname{recc}(\mathcal{C})$ bliver afsluttet hvis $\mathcal{C}$ selv er afsluttet.

Eksempel 6.14.
a) Lad $\mathcal{C} = \{(x_1, x_2) \in \mathbb{R}^2 | x_2 \geq x_1^2\}$



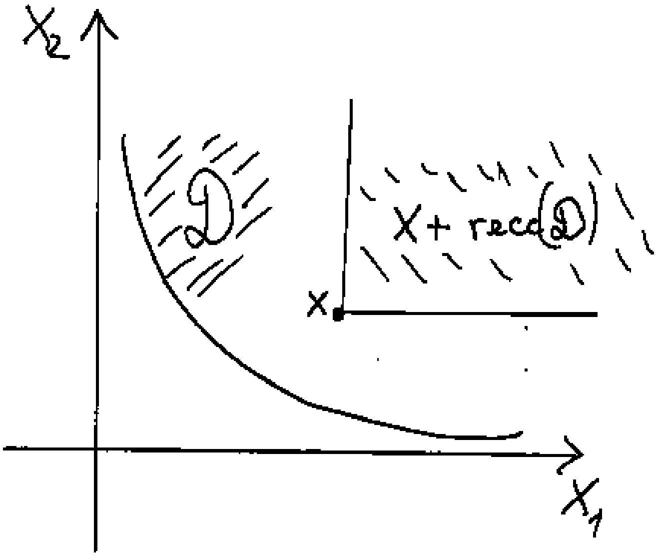


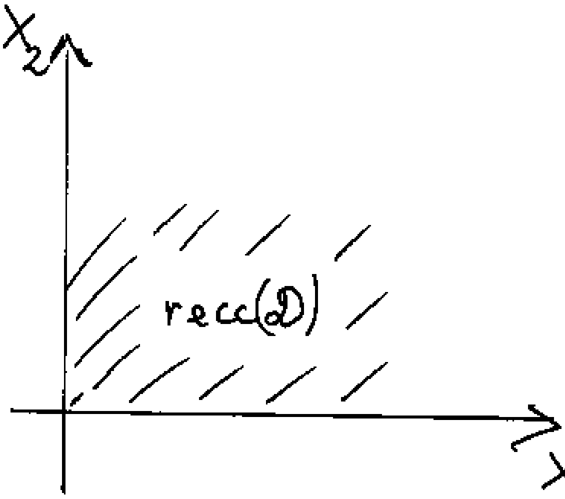


her er $\operatorname{recc}(\mathcal{C}) = \{(x_1, x_2) \Big| x_1 = 0\,, x_2 \geq 0\}$, dvs den ikke negative del af x_2–aksen.

Det ses let at $\operatorname{recc}(\mathcal{C})$ indeholder alle $(0, s)$, $s \geq 0$. På den anden side ses det nemt, at for $y = (y_1, y_2)$ med $y_1 \neq 0$ eller $y_2 < 0$ vil $y \notin \operatorname{recc}(\mathcal{C})$. Se f.eks. på ovenstående tegninger

b) Lad nu $\mathcal{D} = \{(x_1, x_2) \in \mathbb{R}^2 | x_1 > 0 \text{ og } x_1 x_2 \geq 1\}$





Vi ser at $\operatorname{recc}(\mathcal{D}) = \{(x_1, x_2) | x_1 \geq 0 \text{ og } x_2 \geq 0\}$.

Sætning 6.15. *Lad $\mathcal{C}$ være en konveks mængde i $\mathbb{R}^n$, da er* $\operatorname{recc}(\mathcal{C})$ *en konveks kegle i $\mathbb{R}^n$.*

Bevis. Det er klart at $0 \in \operatorname{recc}(\mathcal{C})$ så KK1 er opfyldt. Lad $y_1, y_2 \in \operatorname{recc}(\mathcal{C})$, så er for et vilkårligt x i $\mathcal{C}$ $x + (y_1 + y_2) = (x + y_1) + y_2$. Her er $x + y_1 \in \mathcal{C}$ da $y_1 \in \operatorname{recc}(\mathcal{C})$, og dernæst $(x + y_1) + y_2 \in \mathcal{C}$ da $y_2 \in \operatorname{recc}(\mathcal{C})$, så $y_1 + y_2 \in \operatorname{recc}(\mathcal{C})$ og KK2 følger. Lad nu $y \in \operatorname{recc}(\mathcal{C})$ og $\lambda \geq 0$. Hvis $\lambda = 0$ er $\lambda y = 0 \in \operatorname{recc}(\mathcal{C})$. Antag derfor at $\lambda > 0$, og lad $m \in \mathbb{N}$ være valgt så $m > \lambda$. Da KK2 er vist gyldig for $\operatorname{recc}(\mathcal{C})$ ses det, at $my \in \operatorname{recc}(\mathcal{C})$. Lad nu $x \in \mathcal{C}$. For at vise, at $x + \lambda y \in \mathcal{C}$, bemærkes det, at $x + my \in \mathcal{C}$, så da $\mathcal{C}$ er konveks fås for $0 \leq \alpha \leq 1$, at

$$(1 - \alpha)x + \alpha(x + my) \in \mathcal{C}\,,$$

men for $\alpha = \frac{\lambda}{m}$ fås da $x + \lambda y \in \mathcal{C}$, og KK3 holder for $\operatorname{recc}(\mathcal{C})$ som derfor er en konveks kegle. □

Opgave. Lad $K \subseteq \mathbb{R}^n$ være en konveks kegle.
Vis, at $K - K$ er et lineært underrum.
Vis, at $K - K = \operatorname{aff}(K)$.
Vis, at $K - K$ er det mindste lineære underrum som indeholder K; altså $K - K = \operatorname{span}(K)$.

7 Ubegrænsede konvekse mængder

Som bekendt kaldes en delmængde $\mathcal{B} \subseteq \mathbb{R}^n$ begrænset hvis der findes et $R > 0$ så $\mathcal{B} \subseteq K(0, R)$. Hvis mængden ikke er begrænset kaldes den ubegrænset. En ofte anvendt egenskab ved en ubegrænset delmængde $\mathcal{C}$ af $\mathbb{R}^n$ er, at for alle $k \in \mathbb{N}$ findes $x_k \in \mathcal{C} \cap (\mathbb{R}^n \setminus K(0, k))$, dvs. der findes en følge $(x_k)_{k\in\mathbb{N}}$, $x_k \in \mathcal{C}$ og $\|x_k\| \geq k$.

Vi starter dette kapitel med dets hovedresultatet som er Sætning. 7.1

Sætning 7.1. *Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være en ikke–tom afsluttet konveks mængde, da er $\mathcal{C}$ ubegrænset hvis og kun hvis* $\operatorname{recc}(\mathcal{C}) \neq \{0\}$.

Bevis. Hvis $\operatorname{recc}(\mathcal{C}) \neq 0$ findes $y \neq 0$ med $y \in \operatorname{recc}(\mathcal{C})$. Da $\operatorname{recc}(\mathcal{C})$ er en kegle fås for $x \in \mathcal{C}$ og $k \in \mathbb{N}$, at $ky \in \operatorname{recc}(\mathcal{C})$ og $x + ky \in \mathcal{C}$. Følgen $x_k = x + ky$ opfylder da $x_k \in \mathcal{C}$, og $\|x_k\| = \|x + ky\| \geq \|ky\| - \|x\| = k\|y\| - \|x\| \to +\infty$ for $k \to \infty$ så $\mathcal{C}$ er ubegrænset.

Antag nu at $\mathcal{C}$ er ubegrænset; da har den som sagt en følge $(x_k)_{k\in\mathbb{N}}$, $x_k \in \mathcal{C}$ så $\|x_k\| \geq k$. Heraf kan dannes en følge af enhedsvektorer $y_k := \frac{1}{\|x_k\|}x_k$. Denne følge har en konvergent delfølge da sfæren $S := \{x \in \mathbb{R}^n \mid \|x\| = 1\}$ er kompakt. Lad $(y_{k_i})_{i\in\mathbb{N}}$ være en sådan konvergent delfølge med grænsepunkt $y \in S$. Da skal vi se at $y \in \operatorname{recc}(\mathcal{C})$. Tages nemlig $x \in \mathcal{C}$ fås

$$
\begin{aligned}
x + y &= \lim_{i\to\infty}(x + y_{k_i}) = \lim_{i\to\infty}\left(x + \frac{1}{\|x_{k_i}\|}x_{k_i}\right) \\
&= \lim_{i\to\infty}\left(\left(\left(1 - \frac{1}{\|x_{k_i}\|}\right)x + \frac{1}{\|x_{k_i}\|}x_{k_i}\right) + \frac{1}{\|x_{k_i}\|}x\right) \\
&= \lim_{i\to\infty}\left(\left(1 - \frac{1}{\|x_{k_i}\|}\right)x + \frac{1}{\|x_{k_i}\|}x_{k_i}\right) \;;\; \text{da } \left\|\frac{x}{\|x_{k_i}\|}\right\| \to 0\,.
\end{aligned}
$$

Da $\mathcal{C}$ er konveks er $c_{k_i} := \left(1 - \frac{1}{\|x_{k_i}\|}\right)x + \frac{1}{\|x_{k_i}\|}x_{k_i} \in \mathcal{C}$, så $x + y = \lim c_{k_i}$ er i $\mathcal{C}$, da $\mathcal{C}$ er afsluttet. Heraf følger at $y \in \operatorname{recc}(\mathcal{C})$, $\|y\| = 1$, og $\operatorname{recc}(\mathcal{C}) \neq \{0\}$.

$\square$

I resten af dette afsnit bliver det vist hvorledes en ubegrænset afsluttet konveks mængde kan beskrives som en sum af et lineært underrum, samt en konveks mængde der ikke indeholder linier.

Lemma 7.2. *Lad $\mathcal{C}$ være en afsluttet konveks delmængde af $\mathbb{R}^n$ og lad $b \in \mathbb{R}^n \setminus \{0\}$.*

Af nedenstående 6 betingelser er 1), 2) samt 3) indbyrdes ensbetydende og 4), 5) samt 6) er indbyrdes ensbetydende.

1) $b \in \mathrm{recc}(\mathcal{C})$.

2) $\forall c \in \mathcal{C}\ \forall \lambda \in \mathbb{R}\lambda \geq 0: \quad c + \lambda b \in \mathcal{C}$.

3) $\exists c_0 \in \mathcal{C}\ \forall \lambda \in \mathbb{R}\lambda \geq 0: \quad c_0 + \lambda b \in \mathcal{C}$.

4) $b \in \mathrm{recc}(\mathcal{C}) \cap (-\mathrm{recc}(\mathcal{C}))$.

5) $\forall c \in \mathcal{C}\ \forall \lambda \in \mathbb{R}: \quad c + \lambda b \in \mathcal{C}$.

6) $\exists c_0 \in \mathcal{C}\ \forall \lambda \in \mathbb{R}: \quad c_0 + \lambda b \in \mathcal{C}$.

Bevis. 1) $\implies$ 2): Da $b \in \mathrm{recc}(\mathcal{C})$ er $\lambda b \in \mathrm{recc}(\mathcal{C})$ for $\lambda \geq 0$, så $c + \lambda b \in \mathcal{C}$ når $\lambda \geq 0$, og 2) følger.

2) $\implies$ 3): For $c_0 \in \mathcal{C}$, er 3) et specialtilfælde af 2).

3) $\implies$ 1): Lad $k \in \mathbb{N}$, så fås, idet $\mathcal{C}$ er konveks at nedenstående c_k er element i $\mathcal{C}$.

$$c_k := \frac{k-1}{k} c + \frac{1}{k}(c_0 + kb)\,,$$

Ved en simpel omskrivning ses det at:

$$c_k = c + b + \frac{1}{k}(c_0 - c)\,,$$

så $c_k \to c + b$ for $k \to \infty$, og da $\mathcal{C}$ er afsluttet ses det at $c + b \in \mathcal{C}$, så $b \in \mathrm{recc}(\mathcal{C})$.

Eftersom $(-1)(-1) = 1$ ses det umiddelbart at ækvivalenserne 1), 2) og 3) medfører at nedenstående 3 udsagn 7), 8), og 9) er ensbetydende:

7) $b \in -\mathrm{recc}(\mathcal{C})$.

8) $\forall c \in \mathcal{C}\ \forall \lambda \in \mathbb{R}\lambda \leq 0: \quad c + \lambda b \in \mathcal{C}$.

9) $\exists c_0 \in \mathcal{C}\ \forall \lambda \in \mathbb{R}\lambda \leq 0: \quad c_0 + \lambda b \in \mathcal{C}$.

Det ses at der gælder
1) $\wedge$ 7) $\Leftrightarrow$ 4); 2) $\wedge$ 8) $\Leftrightarrow$ 5); 3) $\wedge$ 9) $\Leftrightarrow$ 6);
og det er vist at betingelserne 4), 5) og 6) er enbetydende.

$\square$

Hovedindholdet af ovenstående lemma er altså, at en afsluttet konveks mængde $\mathcal{C}$ indeholder en linie med retningsvektor $b \neq 0$ hvis og kun hvis $b \in \mathrm{recc}(\mathcal{C}) \cap (-\mathrm{recc}(\mathcal{C}))$.

Fra lineær algebra vides det, at et lineært underrum $\mathcal{L} \subseteq \mathbb{R}^n$ har et ortogonalt komplement $\mathcal{L}^\perp$ som også er et lineært underrum. Det defineres ved

$$\begin{aligned} \mathcal{L}^\perp &= \{x \in \mathbb{R}^n \mid \forall \ell \in \mathcal{L} : \ell \cdot x = 0\} \\ &= \bigcap_{\ell \in \mathcal{L}} \ell^\perp = \text{ en fællesmængde af lineære underrum.} \end{aligned}$$

Som bekendt kan enhver vektor $x \in \mathbb{R}^n$ på netop én måde skrives som en sum $x = \ell + u$, hvor $\ell \in \mathcal{L}$ og $u \in \mathcal{L}^\perp$, og der gælder $\|x\|^2 = \|\ell\|^2 + \|u\|^2$. Endvidere er $(\mathcal{L}^\perp)^\perp = \mathcal{L}$.

Sætning 7.3. *Lad $\mathcal{C}$ være en lukket konveks mængde, da gælder*

1) $\mathcal{L} := \mathrm{recc}(\mathcal{C}) \cap (-\mathrm{recc}(\mathcal{C}))$ *er et lineært underrum.*

2) $\mathcal{C} \cap \mathcal{L}^{\perp}$ *er en lukket konveks mængde som ikke indeholder linier.*

3) $\mathcal{C} = \mathcal{L} + \mathcal{C} \cap \mathcal{L}^{\perp}$.

Bevis. Ad 1): Ifølge Sætning 6.15 er $\mathrm{recc}(\mathcal{C})$ en konveks kegle, og dermed følger 1) fra Sætning 6.11.

Ad 2): $\mathcal{C}$ er afsluttet og $\mathcal{L}^{\perp} = \bigcap_{\ell \in \mathcal{L} \setminus \{0\}} H(\ell, 0)$. En hyperplan $H(\ell, 0)$ er afsluttet som originalmængden til 0 ved den kontinuerte afbildning $x \to x \cdot \ell$. Fællesmængder af afsluttede mængder er afsluttede, så $\mathcal{L}^{\perp}$ er afsluttet og $\mathcal{C} \cap \mathcal{L}^{\perp}$ er afsluttet; endvidere er $\mathcal{L}^{\perp}$ et lineært underrum, og dermed en konveks mængde, så $\mathcal{C} \cap \mathcal{L}^{\perp}$ er konveks og afsluttet. Antag nu, at $\mathcal{C} \cap \mathcal{L}^{\perp}$ indeholder en linie $m := \{c_0 + \lambda b | \lambda \in \mathbb{R}\}$, hvor $b \neq 0$, $c_0 \in \mathcal{C}$, da giver Lemma 7.2 at $b \in \mathcal{L}$; på den anden side giver antagelsen $m \subseteq \mathcal{C} \cap \mathcal{L}^{\perp}$, at $c_0 \in \mathcal{C} \cap \mathcal{L}^{\perp}$ og $c_0 + b \in \mathcal{C} \cap \mathcal{L}^{\perp}$ dvs., da $\mathcal{L}^{\perp}$ er et lineært underrum at

$$b = (c_0 + b) - c_0 \in \mathcal{L}^{\perp} \text{ og } b \in \mathcal{L}\,,$$

så $\|b\|^2 = b \cdot b = 0$ i modstrid med antagelsen om, at $b \neq 0$, eller at m er en egentlig linie.

Ad 3): Da $\mathcal{L} \subseteq \mathrm{recc}(\mathcal{C})$ følger det at $\mathcal{C} + \mathcal{L} \subseteq \mathcal{C}$; på den anden side gælder $0 \in \mathcal{L}$, så $\mathcal{C} \subseteq \mathcal{C} + \mathcal{L}$, og $\mathcal{C} = \mathcal{C} + \mathcal{L}$. Enhver vektor $c \in \mathcal{C}$ kan dekomponeres som $c = \ell_c + n_c$, hvor $\ell_c \in \mathcal{L}$ og $n_c \in \mathcal{L}^{\perp}$ på entydig måde. Da $n_c = c - \ell_c \in \mathcal{C} + \mathcal{L}$ ses det, at $n_c \in \mathcal{C} \cap \mathcal{L}^{\perp}$, og $c = \ell_c + n_c \in \mathcal{L} + \mathcal{L}^{\perp} \cap \mathcal{C}$, altså $\mathcal{C} \subseteq \mathcal{L} + \mathcal{C} \cap \mathcal{L}^{\perp}$. På den anden side er $\mathcal{C} \cap \mathcal{L}^{\perp} \subseteq \mathcal{C}$, så $\mathcal{L} + \mathcal{C} \cap \mathcal{L}^{\perp} \subseteq \mathcal{L} + \mathcal{C} = \mathcal{C}$, og $\mathcal{L} + \mathcal{C} \cap \mathcal{L}^{\perp} = \mathcal{C}$. □

Bemærk at ifølge beviset for 3) er $\mathcal{C} \cap \mathcal{L}^{\perp}$ faktisk lig den ortogonale projektion af $\mathcal{C}$ på $\mathcal{L}^{\perp}$.

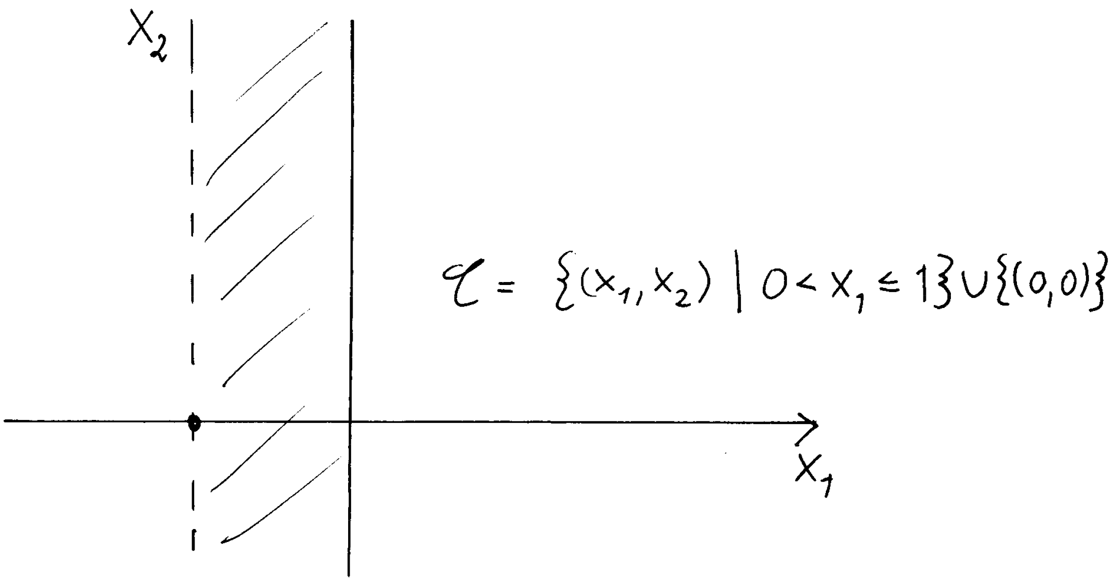
Vi fortsætter studiet af ubegrænsede konvekse mængder i afsnit 12.

Eksempel 7.4. Bemærk kravet om at $\mathcal{C}$ skal være *afsluttet* og ubegrænset for at $\mathrm{recc}(\mathcal{C}) \neq 0$.

Nedenstående eksempel viser hvorfor der kræves afsluttethed.

Lad $\mathcal{C} = \{(x_1, x_2) \in \mathbb{R}^2 \mid 0 < x_1 \leq 1\} \cup \{(0, 0)\}$

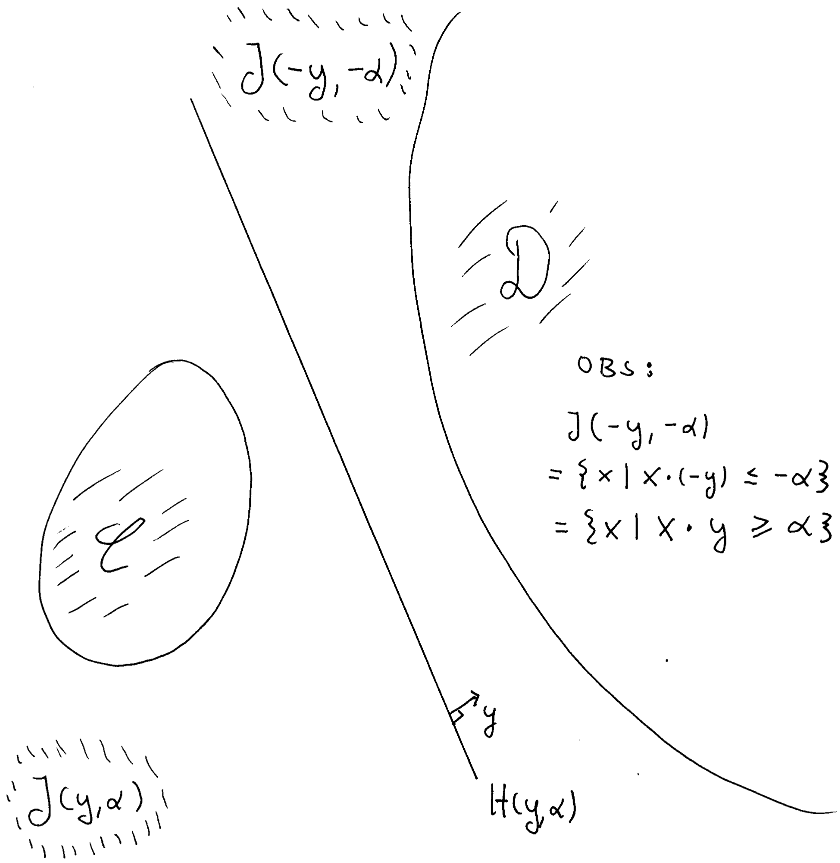
$\operatorname{recc}(\overline{\mathcal{C}}) = \{(0,t) \mid t \in \mathbb{R}\}$
$\operatorname{recc}(\mathcal{C}) = \{0\}$ da $(0,0) + (0,t) \in \mathcal{C} \iff t = 0\,.$



Generelt gælder der i hht. opgave 31 at $\operatorname{recc}(\mathcal{C}) \subseteq \operatorname{recc}(\overline{\mathcal{C}})$.

8 Separationssætningen, elementært

Indholdet af separationssætningen illustreres nemmest med en tegning.



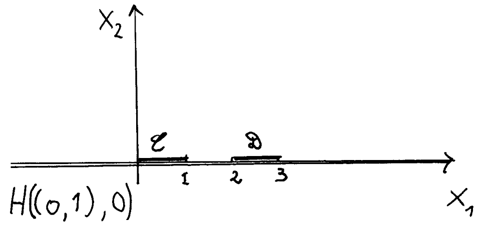
Tegningen forestiller 2 *disjunkte* konvekse delmængder $\mathcal{C}$ og $\mathcal{D}$ af $\mathbb{R}^n$. Som antydet på tegningen findes der da en hyperplan $H(y, \alpha)$, så $\mathcal{C} \subseteq J(y, \alpha)$ og $\mathcal{D} \subseteq J(-y, -\alpha)$.

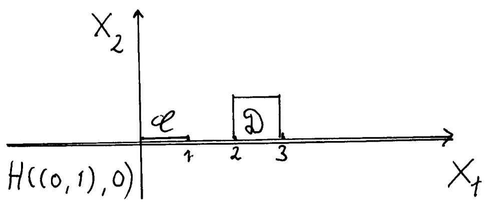
Vi når ikke frem til at vise et helt så generelt resultat i første omgang, idet vi her nøjes med at betragte det tilfælde, hvor der er en positiv afstand mellem $\mathcal{C}$ og $\mathcal{D}$, dvs. $\inf\{\|c - d\| \mid c \in \mathcal{C}, d \in \mathcal{D}\} > 0$.

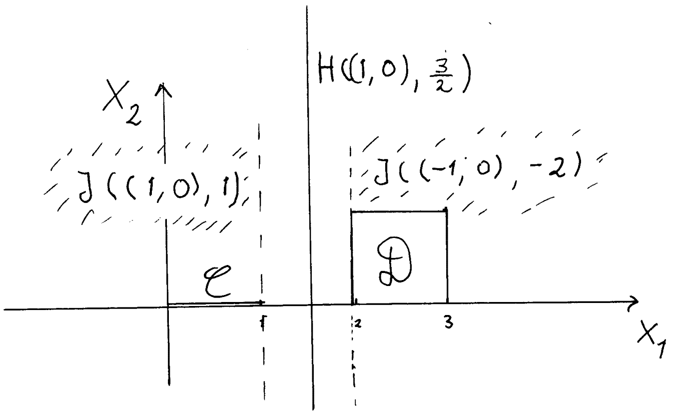
For at kunne tale om separation definerer vi 3 grader af separation som en hyperplan kan inducere mellem 2 konvekse mængder $\mathcal{C}$ og $\mathcal{D}$. Først anskueliggøres begreberne med tegninger i planen; men allerførst introduceres det svageste separationsbegreb helt abstrakt.

Definition 8.1. Lad $\mathcal{C}, \mathcal{D}$ være delmængder af $\mathbb{R}^n$, $y \in \mathbb{R}^n \setminus \{0\}$ og $\alpha \in \mathbb{R}$, da siges hyperplanen $H(y, \alpha)$ at separere $\mathcal{C}$ og $\mathcal{D}$ hvis $\mathcal{C} \subseteq J(y, \alpha)$ og $\mathcal{D} \subseteq J(-y, -\alpha)$.

Eksempel 8.2. Lad $\mathcal{C} = \{(s, 0) \mid 0 \leq s \leq 1\}$, $\mathcal{D} = \{(t, 0) \mid 2 \leq t \leq 3\}$ Disse delmængder af X_1–aksen er tegnet på den øverste tegning og man kan se at de separeres at førsteaksen selv, men separationen er ikke hvad vi vil kalde egentlig. På den næste tegning er det relative indre af det nye D disjunkt med førsteaksen, så her vil vi kalde separationen egentlig men ikke stærk. Den sidste tegning illustrerer stærk separation.







Lad $y = (0, 1)$, $\alpha = 0$, så er $H(y, \alpha) = X_1$–aksen

$$\begin{aligned}
\mathcal{C} \subseteq H(y, \alpha) \subseteq J(y, \alpha) \quad &= \{(x_1, x_2) \in \mathbb{R}^2 | x_2 \leq 0\}, \\
\mathcal{D} \subseteq H(y, \alpha) \subseteq J(-y, -\alpha) &= \{(x_1, x_2) \in \mathbb{R}^2 | -x_2 \leq 0\}, \\
&= \{(x_1, x_2) \in \mathbb{R}^2 | x_2 \geq 0\}.
\end{aligned}$$

Altså $H((0, 1), 0)$ separerer, men vi synes nok det er lidt udartet og ville have foretrukket separation ved f.eks. $H((1,0),\frac{3}{2}) = \{(x_1,x_2) \mid x_1 = \frac{3}{2}\}$.

Vi vil nu præcisere disse separationsbegreber.

Definition 8.3. Lad $\mathcal{C}$ og $\mathcal{D}$ være delmængder af $\mathbb{R}^n$ og $y \in \mathbb{R}^n \setminus \{0\}$, $\alpha \in \mathbb{R}$.

(i) Hyperplanen $H(y,\alpha)$ siges at separere $\mathcal{C}$ og $\mathcal{D}$ *egentligt*, hvis $\mathcal{C} \subseteq J(y,\alpha)$, $\mathcal{D} \subseteq J(-y,-\alpha)$ og $\mathcal{C} \cup \mathcal{D} \nsubseteq H(y,\alpha)$.

(ii) Hyperplanen $H(y,\alpha)$ siges at separere $\mathcal{C}$ og $\mathcal{D}$ *stærkt*, såfremt der findes $\gamma < \alpha$ og $\beta > \alpha$ så $\mathcal{C} \subseteq J(y,\gamma)$ og $\mathcal{D} \subseteq J(-y,-\beta)$.

Bemærk. I eksemplerne fra før hvor $\mathcal{C}$ og $\mathcal{D}$ begge er delmængder af x–aksen, separerer denne $\mathcal{C}$ og $\mathcal{D}$, men ikke egentligt. Derimod separerer x–aksen det næste par egentligt, men ikke stærkt. Det gør derimod $H((1,0),\frac{3}{2})$, idet $\mathcal{C} \subseteq J((1,0),1)$ og $\mathcal{D} \subseteq J((1,0),-2)$

Definition 8.4. Lad $\mathcal{C}$ og $\mathcal{D}$ være konvekse delmængder af $\mathbb{R}^n$, da defineres afstanden dist $(\mathcal{C},\mathcal{D})$ mellem $\mathcal{C}$ og $\mathcal{D}$ ved

(i) dist $(\mathcal{C},\mathcal{D}) = \infty$ hvis en af mængderne er tom

(ii) dist $(\mathcal{C},\mathcal{D}) = \inf\{\|c-d\| \mid c \in \mathcal{C}, d \in \mathcal{D}\}$ ellers.

Sætning 8.5. *Lad $\mathcal{C}$ og $\mathcal{D}$ være konvekse delmængder af $\mathbb{R}^n$. Hvis $0 <$ dist $(\mathcal{C},\mathcal{D}) < \infty$ kan $\mathcal{C}$ og $\mathcal{D}$ separeres stærkt.*

Bevis. Betragt den konvekse mængde $\mathcal{F} := \mathcal{D}-\mathcal{C} = \{d-c \mid c \in \mathcal{C}, d \in \mathcal{D}\}$ (Sætning 5.4.1). Da hverken $\mathcal{C}$ eller $\mathcal{D}$ er tomme, er $\mathcal{F}$ en ikke–tom konveks mængde, og der gælder for δ givet som $\delta =$ dist $(\mathcal{C},\mathcal{D})$, at $\delta > 0$ og for alle f i $\mathcal{F}$ er $\|f\| \geq \delta$.

Da $\delta = \inf\{\|d-c\| \mid c \in \mathcal{C}, d \in \mathcal{D}\}$ findes for $m \in \mathbb{N}$, $f_m = d_m - c_m$ i $\mathcal{F}$, så $\delta \leq \|f_m\| \leq \sqrt{\delta^2 + \frac{1}{m^2}} \leq \delta + 1$. Specielt er følgen $(f_m)_{m\in\mathbb{N}}$ begrænset så der findes en konvergent delfølge $(f_{m_i})_{i\in\mathbb{N}}$. Lad $y \in \mathbb{R}^n$ betegne grænseværdien for denne delfølge, så må der gælde $\|y\| = \delta$.
Næste punkt er så at vise, at for alle $f \in \mathcal{F}$ gælder det at $f \cdot y \geq \delta^2$. Hertil bemærkes det først at for alle $f \in \mathcal{F}$ gælder $\|f\| \geq \delta$; dvs. $\mathcal{F} \subseteq \{x \in \mathbb{R}^n \mid \|x\| \geq \delta\}$. Denne sidste mængde er originalmængden til det lukkede interval $[\delta,\infty[$ ved den kontinuerte funktion $x \to \|x\|$, og er dermed selv lukket. Følgelig er afslutningen $\overline{\mathcal{F}} \subseteq \{x \in \mathbb{R}^n \mid \|x\| \geq \delta\}$. Efter sin definition ($y = \lim_{i\to\infty} f_{m_i}$) er y element i $\overline{\mathcal{F}}$, og efter Sætning 5.4.6 er $\overline{\mathcal{F}}$ konveks, så for et vilkårligt $g \in \overline{\mathcal{F}}$ og $\alpha \in [0,1]$ gælder

$$(1-\alpha)y + \alpha g \in \overline{\mathcal{F}} \text{ samt } \|(1-\alpha)y + \alpha g\|^2 \geq \delta^2 = \|y\|^2 .$$

Udregning af $\|(1-\alpha)y + \alpha g\|^2$ giver da

$$\begin{aligned}(1-\alpha)^2\|y\|^2 + \alpha^2\|g\|^2 + 2\alpha(1-\alpha)y \cdot g &\geq \|y\|^2 \\ (\alpha^2 - 2\alpha)\|y\|^2 + \alpha^2\|g\|^2 + 2\alpha(1-\alpha)y \cdot g &\geq 0 .\end{aligned}$$

Betragtes nu kun $\alpha > 0$, kan uligheden ovenfor multipliceres med $\frac{1}{\alpha}$ og vi får

$$\forall \alpha \in]0,1] : (\alpha - 2)\|y\|^2 + \alpha\|g\|^2 + (2 - 2\alpha) y \cdot g \geq 0\,.$$

Lader vi så α gå mod 0, fås

$$-2\|y\|^2 + 2y \cdot g \geq 0 \text{ eller } y \cdot g \geq \delta^2\,.$$

Denne ulighed gælder for alle $g \in \overline{\mathcal{F}}$, altså specielt også for $g = d - c$, så vi får

$$\begin{aligned}
&\forall c \in \mathcal{C}\ \forall d \in \mathcal{D} : \qquad y \cdot (d - c) \geq \delta^2 \\
&\text{eller} \\
&\forall c \in \mathcal{C}\ \forall d \in \mathcal{D} : \qquad y \cdot d \geq y \cdot c + \delta^2\,.
\end{aligned}$$

Da $\mathcal{C}$ og $\mathcal{D}$ ikke er tomme findes c_0, d_0 i henholdsvis $\mathcal{C}$ og $\mathcal{D}$, og vi får

$$\forall d \in \mathcal{D} : \qquad\qquad y \cdot d \geq y \cdot c_0 + \delta^2\,.$$

Talmængden $\{y \cdot d | d \in \mathcal{D}\}$ er derfor nedadtil begrænset, og vi kan definere

$$\begin{aligned}
&\beta := \inf\{y \cdot d \mid d \in \mathcal{D}\} \in \mathbb{R} \\
&\beta \geq y \cdot c_0 + \delta^2\,.
\end{aligned} \tag{1}$$

Analogt fås for d_0

$$\forall c \in \mathcal{C} : \qquad\qquad y \cdot d_0 - \delta^2 \geq y \cdot c\,.$$

Talmængden $\{y \cdot c | c \in \mathcal{C}\}$ er derfor opadtil begrænset, og vi kan definere

$$\begin{aligned}
&\gamma := \sup\{y \cdot c | c \in \mathcal{C}\} \in \mathbb{R} \\
&\gamma \leq y \cdot d_0 - \delta^2\,.
\end{aligned} \tag{2}$$

Da endvidere c_0 og d_0 er tilfældige i henholdsvis $\mathcal{C}$ og $\mathcal{D}$, kan vi f.eks. lade d_0 variere over $\mathcal{D}$ og uligheden gælder som for d_0;

$$\gamma \leq y \cdot d_0 - \delta^2$$

giver

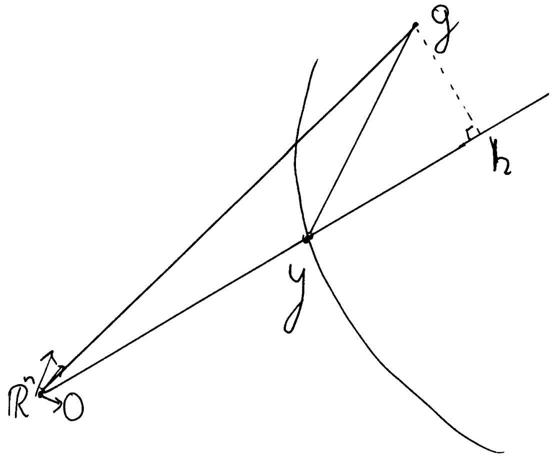
$$\gamma \leq \inf\{y \cdot d_0 | d_0 \in \mathcal{D}\} - \delta^2 = \beta - \delta^2\,. \tag{3}$$

Samler vi sammen fås fra (3): $\beta > \gamma$ da $\delta > 0$, og fra (1) $\mathcal{D} \subseteq J(-y, -\beta)$, og fra (2) $\mathcal{C} \subseteq J(y, \gamma)$. Da $y \neq 0$ ses det, at med $\alpha = \frac{1}{2}(\beta + \gamma)$ vil $H(y, \alpha)$ separere $\mathcal{C}$ og $\mathcal{D}$ stærkt. Sætningen er dermed bevist. □

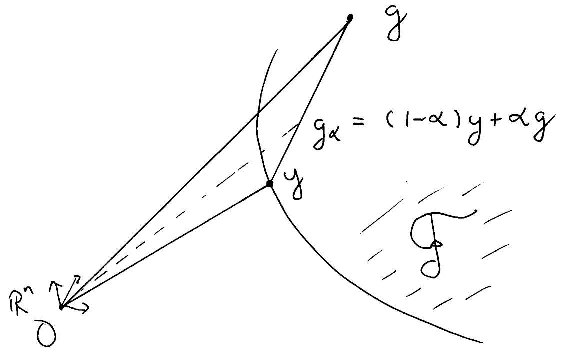
Opgave: Vis at hvis $\mathcal{C}$ og $\mathcal{D}$ er ikke tomme afsluttede disjunkte delmængder af $\mathbb{R}^n$ og mindst en af dem er kompakt så vil $0 < \operatorname{dist}(\mathcal{C}, \mathcal{D}) < \infty$. Fra opgaven ses det så at ikke tomme disjunkte afsluttede konvekse mængder kan separeres stærkt såfremt en af dem er kompakt. Eksempel 8.11 illustrerer både ovenstående opgave samt separationsresultatet.

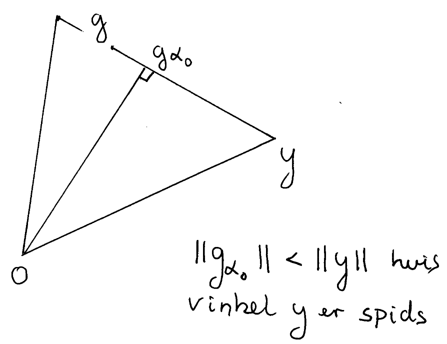
Bemærkning. Den fundamentale bestanddel i ovenstående bevis er egentlig geometrisk, og man bør derfor også prøve at forstå beviset geometrisk. Situationen er altså at vi har en afsluttet konveks mængde $\overline{\mathcal{F}}$ i $\mathbb{R}^n$, og et punkt $y \in \overline{\mathcal{F}}$ af minimal norm.

$$y \in \overline{\mathcal{F}}\,,\ \|y\| = \delta > 0\,,\ \forall g \in \overline{\mathcal{F}} : \|g\| \geq \delta$$



Man skal nu overbevise sig om, at $g \cdot y \geq \|y\|^2$. Da det for $g_\alpha = (1-\alpha)y + \alpha g$, $0 \leq \alpha \leq 1$ gælder at $g_\alpha \in \overline{\mathcal{F}}$ er $\|g_\alpha\| \geq \delta = \|y\|$. Betragter man nu den tegnede trekant ses det, at vinkel y må være stump. Var den nemlig spids ville fodpunktet for højden fra 0 give et punkt g_α i $\overline{\mathcal{F}}$ som lå tættere på 0





end y.

Når vinkel y er ret eller stump ser man fra den første tegning, at $g \cdot y \geq y \cdot y$ idet $g \cdot y = \|y\| \, \|g\| \cos(< g0y) = \|y\| \, \|h\| \geq \|y\|^2$.

Som en umiddelbar konsekvens af Sætning 8.5 har vi sætningerne 8.6 og 8.8.

Sætning 8.6. *Lad $\mathcal{C}$ være en ikke–tom afsluttet og konveks delmængde af $\mathbb{R}^n$, da gælder*

(i) Hvis $x \in \mathbb{R}^n \setminus \mathcal{C}$, så kan $\{x\}$ og $\mathcal{C}$ separeres stærkt.

(ii) $\mathcal{C}$ er fællesmængden af alle de afsluttede halvrum som indeholder $\mathcal{C}$;

$$\mathcal{C} = \bigcap_{J(y,\alpha) \supseteq \mathcal{C}} J(y, \alpha) \,.$$

Bevis. (i): Antag $x \notin \mathcal{C}$. Da $\mathcal{C}$ er afsluttet findes $r > 0$ så $K(x, r) \cap \mathcal{C} = \emptyset$, dvs. der er positiv afstand mellem $\{x\}$ og $\mathcal{C}$, så Sætning 8.5 kan anvendes.

(ii): Lad $\mathcal{D}$ betegne fællesmængden af alle de afsluttede halvrum som indeholder $\mathcal{C}$

$$\mathcal{D} := \bigcap_{J(y,\alpha) \supseteq \mathcal{C}} J(y, \alpha) \,.$$

Det følger af $\mathcal{D}$'s definition, at $\mathcal{D}$ indeholder $\mathcal{C}$. På den anden side følger det fra (i), at $\mathbb{R}^n \setminus \mathcal{C} \subseteq \mathbb{R}^n \setminus \mathcal{D}$, altså et punkt $x \notin \mathcal{C}$ ligger heller ikke i $\mathcal{D}$, men $\mathbb{R}^n \setminus \mathcal{C} \subseteq \mathbb{R}^n \setminus \mathcal{D}$ er ensbetydende med $\mathcal{D} \subseteq \mathcal{C}$, så $\mathcal{D} = \mathcal{C}$. □

Da fællesmængden af en familie af afsluttede mængder er afsluttet og $\mathbb{R}^n$ er en afsluttet mængde ses det, at det giver god mening at definere det afsluttede konvekse hylster — $\operatorname{cl}\operatorname{conv}(\mathcal{M})$ — af en mængde $\mathcal{M} \subseteq \mathbb{R}^n$ som fællesmængden af alle afsluttede konvekse delmængder af $\mathbb{R}^n$ som indeholder $\mathcal{M}$. Denne afsluttede konvekse mængde indeholder $\mathcal{M}$, og er klart den mindste sådanne, da den er indeholdt i enhver afsluttet konveks delmængde af $\mathbb{R}^n$, som indeholder $\mathcal{M}$.

Definition 8.7. Lad $\mathcal{M} \subseteq \mathbb{R}^n$. Ved $\operatorname{cl}\operatorname{conv}(\mathcal{M})$ forstås den mindste afsluttede konvekse delmængde af $\mathbb{R}^n$, som indeholder $\mathcal{M}$.

Spørgsmålet er naturligvis om $\operatorname{cl}\operatorname{conv}(\mathcal{M}) = \overline{\operatorname{conv}(\mathcal{M})}$, samt om der er andre måder $\operatorname{cl}\operatorname{conv}(\mathcal{M})$ kan karakteriseres på.

Sætning 8.8. *Lad $\mathcal{M} \subseteq \mathbb{R}^n$, da er $\operatorname{cl}\operatorname{conv}(\mathcal{M})$ fællesmængden af alle de afsluttede halvrum som indeholder $\mathcal{M}$*

$$\operatorname{cl}\operatorname{conv}(\mathcal{M}) = \bigcap_{J(y,\alpha) \supseteq \mathcal{M}} J(y, \alpha) \,.$$

Endvidere gælder der $\operatorname{cl}\operatorname{conv}(\mathcal{M}) = \overline{\operatorname{conv}(\mathcal{M})}$.

Bevis. Lad $\mathcal{D} = \bigcap_{J(y,\alpha)\supseteq\mathcal{M}} J(y,\alpha)$. Det er klart at $\mathcal{D}$ er afsluttet og konveks og indeholder $\mathcal{M}$, da hvert $J(y,\alpha)$ har disse egenskaber, dvs. $\mathcal{D} \supseteq \operatorname{cl}\operatorname{conv}(\mathcal{M})$. Af Sætning8.6 (i) følger det, at hvis $x \notin \operatorname{cl}\operatorname{conv}(\mathcal{M})$ findes $J(y,\alpha)$, så $x \notin J(y,\alpha)$, men $J(y,\alpha) \supseteq \operatorname{cl}\operatorname{conv}(\mathcal{M}) \supseteq \mathcal{M}$, dvs. $x \notin \mathcal{D}$, så $\mathcal{D} \subseteq \operatorname{cl}\operatorname{conv}(\mathcal{M})$, og det er vist at $\mathcal{D} = \operatorname{cl}\operatorname{conv}(\mathcal{M})$.

Med hensyn til den sidste påstand "$\operatorname{cl}\operatorname{conv}(\mathcal{M}) = \overline{\operatorname{conv}(\mathcal{M})}$" ses det fra Sætning 5.4.6, at $\overline{\operatorname{conv}(\mathcal{M})}$ er konveks, afsluttet og indeholder $\mathcal{M}$ så $\overline{\operatorname{conv}(\mathcal{M})} \supseteq \operatorname{cl}\operatorname{conv}(\mathcal{M})$. På den anden side er $\operatorname{cl}\operatorname{conv}(\mathcal{M})$ konveks, så $\operatorname{cl}\operatorname{conv}(\mathcal{M}) \supseteq \operatorname{conv}(\mathcal{M})$, men $\operatorname{cl}\operatorname{conv}(\mathcal{M})$ er også afsluttet, så $\operatorname{cl}\operatorname{conv}(\mathcal{M}) \supseteq \overline{\operatorname{conv}(\mathcal{M})}$, og sætningen følger. □

Kapitlet sluttes med et specialtilfælde af Sætning 8.6 (i), idet der nemlig er en ekstra finesse når $\mathcal{C}$ er en konveks kegle i stedet for blot en konveks mængde.

Definition 8.9. Lad $\mathcal{K} \subseteq \mathbb{R}^n$ være en konveks kegle. Ved *normalkeglen* til $\mathcal{K}$, $\operatorname{normc}(\mathcal{K})$ forstås mængden af vektorer som har ikke–positivt skalarprodukt med alle vektorer fra $\mathcal{K}$.

$$\begin{aligned}\operatorname{normc}(\mathcal{K}) &= \{y \in \mathbb{R}^n \mid \forall k \in \mathcal{K} :\ k \cdot y \le 0\}\\ &= \cap_{k\in\mathcal{K}} J(k,0)\,.\end{aligned}$$

Sætning 8.10. *Lad $\mathcal{K} \subseteq \mathbb{R}^n$ være en konveks kegle, så er* $\operatorname{normc}(\mathcal{K})$ *en afsluttet konveks kegle, og* $\operatorname{normc}(\operatorname{normc}(\mathcal{K}))$ *er afslutningen af $\mathcal{K}$.*

Bevis. Da $\operatorname{normc}(\mathcal{K}) = \bigcap_{k\in\mathcal{K}} J(k,0)$, og alle $J(k,0)$ er afsluttede konvekse kegler (overvej!), så er $\operatorname{normc}(\mathcal{K})$ en afsluttet konveks kegle (Sætning 6.6.3). Efter Definition 8.9 følger det, at $\mathcal{K} \subseteq \operatorname{normc}(\operatorname{normc}(\mathcal{K}))$, som er afsluttet og dermed $\overline{\mathcal{K}} \subseteq \operatorname{normc}(\operatorname{normc}(\mathcal{K}))$. Antag nu, at $x \notin \overline{\mathcal{K}}$, da kan x og $\overline{\mathcal{K}}$ separeres stærkt, så der må findes $H(y,\alpha)$ som separerer x og $\overline{\mathcal{K}}$ stærkt, dvs. der findes $\gamma < \alpha < \beta$ så

$$\overline{\mathcal{K}} \subseteq J(y,\gamma) \ \text{ og } \ y \cdot x \ge \beta > \alpha > \gamma\,.$$

Eftersom $0 \in \overline{\mathcal{K}}$ fås $y \cdot 0 \le \gamma$, dvs. $\gamma \ge 0$. Vort første mål er at vise, at $y \in \operatorname{normc}(\mathcal{K})$, dvs. $\mathcal{K} \subseteq J(y,0)$. Antag derfor, at $y \notin \operatorname{normc}(\mathcal{K})$, så findes $k \in \mathcal{K}$ så $y \cdot k > 0$. På den anden side fås $\forall m \in \mathbb{N}$: $mk \in \mathcal{K}$, så $mk \in \overline{\mathcal{K}}$ og $mk \in J(y,\gamma)$, dvs. $y \cdot (mk) \le \gamma$, og $y \cdot k \le \frac{1}{m}\gamma$ for alle $m \in \mathbb{N}$. Altså fås $y \cdot k \le 0$ i modstrid med at $y \cdot k$ er antaget positiv. Altså er $y \in \operatorname{normc}(\mathcal{K})$, og da $\beta > \gamma \ge 0$ og $y \cdot x \ge \beta > 0$ fås $x \notin \operatorname{normc}(\mathcal{K})$, så $\mathbb{R}^n \setminus \overline{\mathcal{K}} \subseteq \mathbb{R}^n \setminus \operatorname{normc}(\operatorname{normc}(\mathcal{K}))$, dvs.

$$\operatorname{normc}(\operatorname{normc}(\mathcal{K})) \subseteq \overline{\mathcal{K}} \ \text{ og } \ \overline{\mathcal{K}} = \operatorname{normc}(\operatorname{normc}(\mathcal{K}))\,.$$

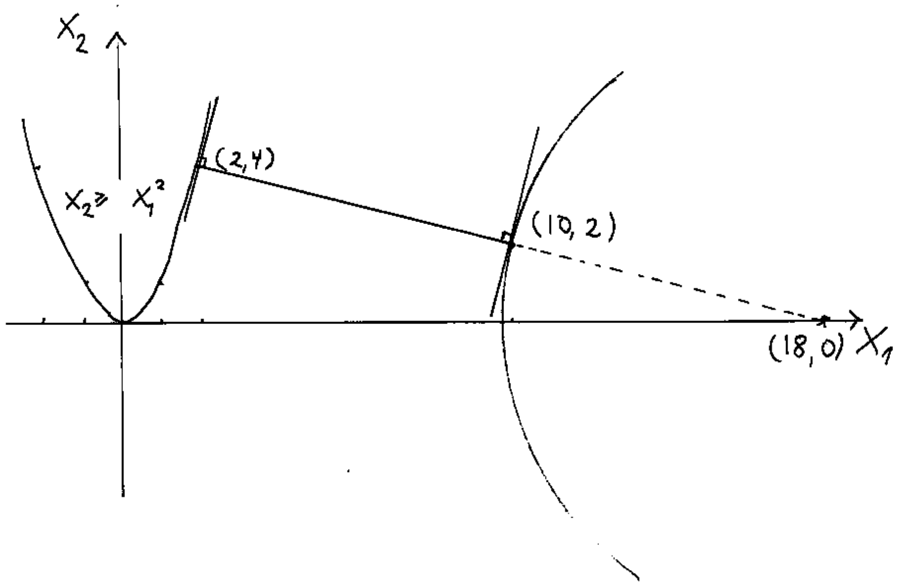
□

Ovenstående sætning er det første eksempel på dualitetsteori, dvs. vi betragter en delmængde $\mathcal{K}$ af $\mathbb{R}^n$, derefter betragter vi de linearformer (dvs. lineære afbildninger af $\mathbb{R}^n$ ind i $\mathbb{R}$) som har en særlig egenskab på $\mathcal{K}$. I dette tilfælde altså $\operatorname{normc}(\mathcal{K})$. Da de lineære afbildninger af $\mathbb{R}^n$ ind i $\mathbb{R}$ kan identificeres med $\mathbb{R}^n$ via $y \in \mathbb{R}^n \leftrightarrow f_y = \langle y, \cdot\rangle \in L(\mathbb{R}^n, \mathbb{R})$, kan vi gentage processen og nu opfatte linearformerne som punkter i $\mathbb{R}^n$, og punkterne i $\mathbb{R}^n$ som linearformer på linearformerne. Hvis dette er volapyk skal du ikke fortvivle; men koncentrere dig om det konkrete indhold $\operatorname{normc}(\operatorname{normc}(\mathcal{K})) = \overline{\mathcal{K}}$ for konvekse kegler.

Eksempel 8.11. I beviset for Sætning 8.5 findes den vektor y som giver separationen lidt indirekte og ikke så klart gennemskueligt. I virkeligheden drejer det sig om at finde to punkter, et i $\overline{\mathcal{C}}$, og et i $\overline{\mathcal{D}}$, så afstanden mellem disse er mindst mulig. Vektoren y er da forbindelsesvektoren mellem disse punkter. Denne procedure er ikke altid praktisabel, men hvis den ene af mængderne er kompakt og den anden er afsluttet (og begge stadig konvekse, naturligvis) kan det lade sig gøre. Jeg vil illustrere dette med et eksempel.

Lad

$$\begin{aligned}\mathcal{C} &= \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 \geq x_1^2\}\,,\\ \mathcal{D} &= \{(x_1, x_2) \in \mathbb{R}^2 \mid (x_1 - 18)^2 + x_2^2 \leq 68\}\end{aligned}$$



Tegningen antyder at de to punkter der er nærmest hinanden har koordinatsættene (2,4) og (10,2). Man kan overbevise sig om rigtigheden af dette ved at at vise, at forbindelseslinien er vinkelret på tangenterne til henholdsvis parablen og cirklen i de respektive punkter:

$$\begin{aligned}&(2,4) \in \text{ randen af } \mathcal{C} \text{ ses ved indsætning}\\ &(10,2) \in \text{randen af } \mathcal{D} \text{ ses ved indsætning}\\ &y := (10,2) - (2,4) = (8,-2)\,.\end{aligned}$$

Tangenten til $\mathcal{C}$ i (2,4) har retningsvektor $(1, 2 \cdot 2) = (1,4)$. Tangenten til $\mathcal{D}$ i $(10,2)$ har retningsvektoren $((18,0) - (10,2))^{\perp} = (8,-2)^{\perp} = (2,8)$,
og for $y = (10,2) - (2,4) = (8,-2)$ fås

$$\begin{aligned}\forall c \in \mathcal{C}: &\qquad c \cdot y \leq (2,4) \cdot (8,-2) = 8\\ \forall d \in \mathcal{D}: &\qquad d \cdot y \geq (10,2) \cdot (8,-2) = 76\end{aligned}$$

dvs.

$$\begin{aligned}\mathcal{C} &\subseteq J((8,-2),8)\\ \mathcal{D} &\subseteq J((-8,2),-76)\,.\end{aligned}$$

Bemærk, at hyperplanerne $H((8,-2),8)$ og $H((-8,2),-76)$ lige netop rører henholdsvis $\mathcal{C}$ og $\mathcal{D}$. Man siger da, at hyperplanerne er støtte–hyperplaner, men mere om det senere. ($\mathcal{C} \subseteq J(y,\alpha)$ og $\mathcal{C} \cap H(y,\alpha) \neq \emptyset \iff H(y,\alpha)$ er støtte–hyperplan for $\mathcal{C}$.)

9 Farkas alternativ og dualitet

I Kapitel 8 så vi på begrebet normalkegle til en konveks kegle $\mathcal{K}$, og definerede normalkeglen til $\mathcal{K}$ som $\operatorname{normc}(\mathcal{K}) = \{y \in \mathbb{R}^n \mid \forall x \in \mathcal{K} : y \cdot x \leq 0\}$. Her er der tale om at vi betragter $y \in \mathbb{R}^n$ som den lineære afbildning af $\mathbb{R}^n$ ind i $\mathbb{R}$ som er givet ved $x \to y \cdot x$. Herved bliver $\operatorname{normc}(\mathcal{K})$ en mængde af lineære afbildninger fra $\mathbb{R}^n$ ind i $\mathbb{R}$, som kun antager ikke–positive værdier på $\mathcal{K}$. Denne leg udvides i emnet lineær programmering til nogle praktisk anvendelige metoder, men her vil vi kun lege med matematikken og de begreber der ligger nærmest for. Det første begreb vi støder på er polaren til en mængde.

Definition 9.1. Lad $\mathcal{M} \subseteq \mathbb{R}^n$. Ved polaren $\operatorname{polar}(\mathcal{M})$ til $\mathcal{M}$ forstås $\mathbb{R}^n$ hvis $\mathcal{M} = \emptyset$ og ellers mængden givet ved

$$\operatorname{polar}(\mathcal{M}) = \{y \in \mathbb{R}^n \mid \forall x \in \mathcal{M} : y \cdot x \leq 1\}.$$

Bemærk at $\operatorname{polar}(0) = \mathbb{R}^n$ og $\operatorname{polar}(\mathbb{R}^n) = 0$.

Sætning 9.2. *Hvis $\mathcal{K}$ er en konveks kegle, er normalkeglen til $\mathcal{K}$ lig polaren til $\mathcal{K}$, dvs.*

$$\{y \in \mathbb{R}^n \mid \forall x \in \mathcal{K} : y \cdot x \leq 0\} = \{y \in \mathbb{R}^n \mid \forall x \in \mathcal{K} : y \cdot x \leq 1\}.$$

Bevis. Da $y \cdot x \leq 0 \implies y \cdot x \leq 1$ fås at venstresiden er en delmængde af højresiden, altså at normalkeglen er en delmængde af polaren til $\mathcal{K}$. Lad da y ligge i polaren til $\mathcal{K}$ og lad $x \in \mathcal{K}$, da fås for alle $N \in \mathbb{N}$ at $Nx \in \mathcal{K}$ så $y \cdot (Nx) \leq 1$, og $y \cdot x \leq \frac{1}{N}$. Heraf ses, at $y \cdot x \leq 0$, og y er element i normalkeglen. $\square$

Vi starter med den såkaldte bipolarsætning.

Sætning 9.3. *Lad $\mathcal{M} \subseteq \mathbb{R}^n$ være ikke–tom, da er* $\operatorname{polar}(\operatorname{polar}(\mathcal{M})) = \operatorname{cl}\operatorname{conv}(\mathcal{M} \cup 0)$.

Bevis. Da $\operatorname{polar}(\mathcal{M}) = \{y \in \mathbb{R}^n \mid \forall m \in \mathcal{M} : m \cdot y \leq 1\}$ ses det, at for $y \in \operatorname{polar}(\mathcal{M})$ er $y \cdot m \leq 1$ ($m \in \mathcal{M}$) så $\mathcal{M} \subseteq \operatorname{polar}(\operatorname{polar}(\mathcal{M}))$. Ligeledes er det klart at 0 ligger i polaren til alle mængder, så $0 \in \operatorname{polar}(\operatorname{polar}(\mathcal{M}))$. Endelig ses det at $\operatorname{polar}(\operatorname{polar}(\mathcal{M})) = \bigcap_{y \in \operatorname{polar}(\mathcal{M})} J(y, 1)$, så $\operatorname{polar}(\operatorname{polar}(\mathcal{M}))$ 0 er fællesmængden af en familie af afsluttede halvrum, og $\operatorname{polar}(\operatorname{polar}(\mathcal{M}))$ er dermed både afsluttet og konveks. Alt i alt er $\operatorname{polar}(\operatorname{polar}(\mathcal{M}))$ en afsluttet konveks delmængde af $\mathbb{R}^n$ som omfatter $\mathcal{M} \cup \{0\}$, og derfor gælder $\operatorname{polar}(\operatorname{polar}(\mathcal{M})) \supseteq \operatorname{cl}\operatorname{conv}(\mathcal{M} \cup \{0\})$. For at vise den modsatte inklusion vælges $x \notin \operatorname{cl}\operatorname{conv}(\mathcal{M} \cup \{0\})$, og det vises, at $x \notin \operatorname{polar}(\operatorname{polar}(\mathcal{M}))$. Til den ende bemærkes det først, at når $x \notin \operatorname{cl}\operatorname{conv}(\mathcal{M} \cup \{0\})$, så kan $\{x\}$ og $\operatorname{cl}\operatorname{conv}(\mathcal{M} \cup \{0\})$ separeres stærkt (Sætning 8.6). Der findes derfor $y \in \mathbb{R}^n \setminus \{0\}$ og reelle tal $\gamma < \alpha < \beta$, så $\mathrm{cl}\,\mathrm{conv}\,(\mathcal{M} \cup \{0\}) \subseteq J(y,\gamma)$ og $y \cdot x \geq \beta$. Da $0 \in J(y,\gamma)$ fås $0 \leq \gamma < \alpha < \beta$, så ved at erstattet y med $\hat{y}$ givet ved $\hat{y} = \frac{1}{\alpha} y$ fås $J(y,\gamma) = J(\frac{1}{\alpha}y, \frac{1}{\alpha}\gamma) = J(\hat{y}, \frac{\gamma}{\alpha}) \subseteq J(\hat{y}, 1)$, da $1 > \frac{\gamma}{\alpha}$.

Heraf ses at $\mathcal{M} \subseteq J(\hat{y},1)$, så $\forall m \in \mathcal{M} : m \cdot \hat{y} \leq 1$, og dermed $\hat{y} \in \mathrm{polar}\,(\mathcal{M})$. På den anden side giver $\alpha < \beta$ at $\hat{y} \cdot x \geq \frac{1}{\alpha} \cdot \beta > 1$ så $x \notin \mathrm{polar}\,(\mathrm{polar}\,(\mathcal{M}))$, da $x \notin \mathrm{polar}\,(\{\hat{y}\})$ og $\mathrm{polar}\,(\{\hat{y}\}) \supset \mathrm{polar}\,(\mathrm{polar}\,(\mathcal{M}))$, eftersom $\hat{y} \in \mathrm{polar}\,(\mathcal{M})$. Sætningen følger. □

Herefter følger et specialtilfælde af bipolarsætningen kaldet Farkas Lemma. I den form det præsenteres her handler sætningen om en karakterisering af $\mathrm{cone}\,(\{a_1, \ldots, a_m\})$, altså keglen frembragt af endeligt mange vektorer fra $\mathbb{R}^n$. Om lidt skal vi se at resultatet også kan læses anderledes.

Sætning 9.4 (Farkas Lemma). *Lad $m \in \mathbb{N}$ og $a_1, \ldots, a_m \in \mathbb{R}^n$. For $a \in \mathbb{R}^n$ gælder der*

$$a \in \mathrm{cone}\,(\{a_1, \ldots, a_m\})$$

hvis og kun hvis

$$\forall x \in \mathbb{R}^n : a_1 \cdot x \leq 0, \ldots, a_m \cdot x \leq 0 \implies a \cdot x \leq 0.$$

Bevis. Beviset er baseret på at $\mathrm{cone}\,(\{a_1, \ldots, a_m\})$ er lukket. Dette resultat kan godt vises (Opgave 1.39) uden særlige hjælpemidler, men kommer forholdsvis let, når vi kender strukturen af polyedre noget bedre. Vi starter altså med at basere beviset på at $\mathrm{cone}\,(\{a_1, \ldots, a_m\})$ er afsluttet konveks og indeholder 0. Lad $A = \{\lambda_i a_i | 1 \leq i \leq m,\ \lambda_i \geq 0\}$, altså alle ikke–negative multipla af vektorene a_i, $1 \leq i \leq m$. Da er $0 \in A$ og $\mathrm{conv}\,(A) = \mathrm{cone}\,(\{a_1, \ldots, a_m\})$, så $\mathrm{cone}\,(\{a_1, \ldots, a_m\}) = \mathrm{polar}\,(\mathrm{polar}\,(A)) = \{a \in \mathbb{R}^n \mid \forall x \in \mathrm{polar}\,(A) : a \cdot x \leq 0\}$. Da $\mathrm{polar}\,(A) = \{x \in \mathbb{R}^n | a_i \cdot x \leq 0, 1 \leq i = m\}$ følger sætningen. □

I kapiteloverskriften står der “Farkas Alternativ” og i Sætning 9.4 “Farkas Lemma”, det virker måske forvirrende, men årsagen er at man ved brug af “matrix-sprog” kan omforme indholdet af 9.4 så indholdet tilsyneladende er et andet. Denne anden version præsenteres umiddelbart herunder i 9.5

Sætning 9.5 (Farkas Alternativ). *Lad A være en $m \times n$ matrix og $c \in \mathbb{R}^n$, da er netop et af nedenstående 2 udsagn sandt.*

1) Der findes $y \in \mathbb{R}^m$ med $y_i \geq 0$, $1 \leq i \leq m$ så $A^t y = c$.

2) Der findes $x \in \mathbb{R}^n$ så $(Ax)_i \geq 0$, $1 \leq i \leq m$ og $c \cdot x < 0$.

Bevis. Indholdet af sætningen kan kort formuleres som: 1) $\iff$ non 2). Defineres vektorer $a_i \in \mathbb{R}^n$ for $1 \leq i \leq m$ ved $a_i = {}_iA$, så er udsagnet 1) endbetydende med $c \in \mathrm{cone}\,(\{a_1, \ldots, a_m\})$. Vi undersøger nu hvad udsagnet *non 2)* siger ved direkte negering af udsagnet 2).

$$\text{non 2)} \iff \forall x \in \mathbb{R}^n \text{ med egenskaben } (Ax)_i \geq 0 \text{ for alle } i \in \{1, \ldots, m\}$$
$$\text{gælder } c \cdot x \geq 0\,.$$

En simpel overvejelse, *baseret på en erstatning af x med $-x$*, viser så at ovenstående udsagn er ensbetydende med:

$$\text{non 2)} \iff \forall x \in \mathbb{R}^n \text{ med egenskaben } (Ax)_i \leq 0 \text{ for alle } i \in \{1, \dots, m\}$$
$$\text{gælder } c \cdot x \leq 0\,.$$

Idet $(Ax)_i = {}_iA \cdot x = a_i \cdot x$ fås

$$\text{non 2)} \iff (\forall x \in \mathbb{R}^n : a_1 \cdot x \leq 0, \dots, a_m \cdot x \leq 0 \implies c \cdot x \leq 0)\,.$$

I følge Sætning 9.4 fås da non 2) $\iff$ 1), og sætningen er vist. $\square$

9.6 Lidt om dualitetsteori

Som bekendt er enhver lineær afbildning $f : \mathbb{R}^n \to \mathbb{R}$ givet ved en vektor $y \in \mathbb{R}^n$, så $f(x) = y \cdot x$, $(x \in \mathbb{R}^n)$. En lineær afbildning $f : \mathbb{R}^n \to \mathbb{R}$ kaldes også en linearform, så vi kan altså identificere et $y \in \mathbb{R}^n$ med en linearform $f_y : \mathbb{R}^n \to \mathbb{R}$ ved $f_y(x) = y \cdot x$. Rummet af linearformer $L(\mathbb{R}^n, \mathbb{R})$ normeres ved hjælp af normen på $\mathbb{R}^n$, og den såkaldte afbildningsnorm ved

$$\forall f \in L(\mathbb{R}^n, \mathbb{R}) \text{ defineres } \|f\| := \sup\{|f(x)| \mid \|x\| \leq 1\,,\, x \in \mathbb{R}^n\}\,.$$

Idet $f = f_y$ for et y fås

$$\|f_y\| = \sup\{|y \cdot x| \mid \|x\| \leq 1\} \leq \|y\| \text{ p.g.a. Cauchy-Schwarz}\,.$$

Hvis $y = 0$ fås $f_y = 0$ og $\|f_y\| = 0$. Hvis $y \neq 0$ er for $x = \frac{1}{\|y\|}y$, $\|x\| = 1$ og $f_y(x) = \|y\|$, så $\|f_y\| \geq \|y\|$, og dermed $\|f_y\| = \|y\|$. AND SO WHAT? Vi har blot fundet at afbildningen $\mathbb{R}^n \ni y \to f_y \in L(\mathbb{R}^n, \mathbb{R})$ er en surjektiv isometri, men dette får skam betydning senere i livet. Det er ikke altid det går sådan. Havde man f.eks. udstyret $\mathbb{R}^n$ med p–normen for $1 \leq p \leq \infty$, $(p \neq 2)$ som defineres nedenfor, ville man få noget helt andet. Lad $x \in \mathbb{R}^n$. For $p = \infty : \|x\|_\infty := \max\{|x_i| \mid 1 \leq i \leq n\}$, og for $1 \leq p < \infty$: $\|x\|_p := (|x_1|^p + \cdots + |x_n|^p)^{1/p}$. Inden vi vender os til hvad man får bemærkes det først at når $|||\ \ |||$ er en vilkårlig norm på $\mathbb{R}^n$, defineres den "duale" norm $|||\cdot|||^*$ på $L(\mathbb{R}^n, \mathbb{R})$ ved for $f \in L(\mathbb{R}^n, \mathbb{R})$ at definere

$$|||f|||^* := \sup\{|f(x)| \mid |||x||| \leq 1\}\,.$$

Udregningerne ovenfor viser nu, at $\|\ \ \|_2^* = \|\ \ \|_2$ når $L(\mathbb{R}^n, \mathbb{R})$ identificeres med $\mathbb{R}^n$, eller helt formelt, for f_y gælder $\|f_y\|_2^* = \|y\|_2$.

Helt generelt kan man vise, at der gælder for tal $p, q \in [1, \infty]$ med $\frac{1}{p} + \frac{1}{q} = 1$, $(\frac{1}{\infty} = 0)$, at

$$\|f_y\|_p^* = \|y\|_q\,.$$

Et eksempel på dette fænomen kan udtrykkes lidt anderledes ved hjælp af polarer. Lad $y \in \mathbb{R}^n$

$$\begin{aligned} \|f_y\|_p^* \leq 1 &\iff \forall x \in \mathbb{R}^n \,,\ \|x\|_p \leq 1 : |y \cdot x| \leq 1 \\ &\iff \forall x \in \mathbb{R}^n \,,\ \|x\|_p \leq 1 : y \cdot x \leq 1 \quad (\| - x\|_p = \|x\|_p) \\ &\iff y \in \operatorname{polar}\left(\{x \in \mathbb{R}^n \mid \|x\|_p \leq 1\}\right). \end{aligned}$$

Når dette knyttes sammen med $\|f_y\|_p^* = \|y\|_q$ fås altså:

For $p, q \in [1, \infty]$ med $\frac{1}{p} + \frac{1}{q} = 1$ gælder

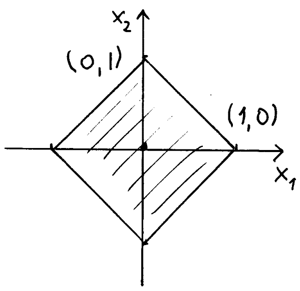
$$\operatorname{polar}\left(\{x \in \mathbb{R}^n \mid \|x\|_p \leq 1\}\right) = \{y \in \mathbb{R}^n \mid \|y\|_q \leq 1\}.$$

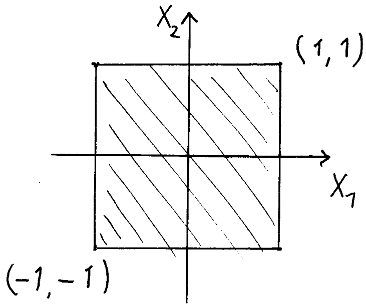
Se opgaverne 3.2, 4 første skridt.

I tråd med opgave 3.2 lader vi nu $n = 2$, og prøver af regne konkret på $\|x\|_p$. Som antydet ovenfor kan det være rimeligt at studere enhedskuglerne B_p for de forskellige normer $\| \quad \|_p$

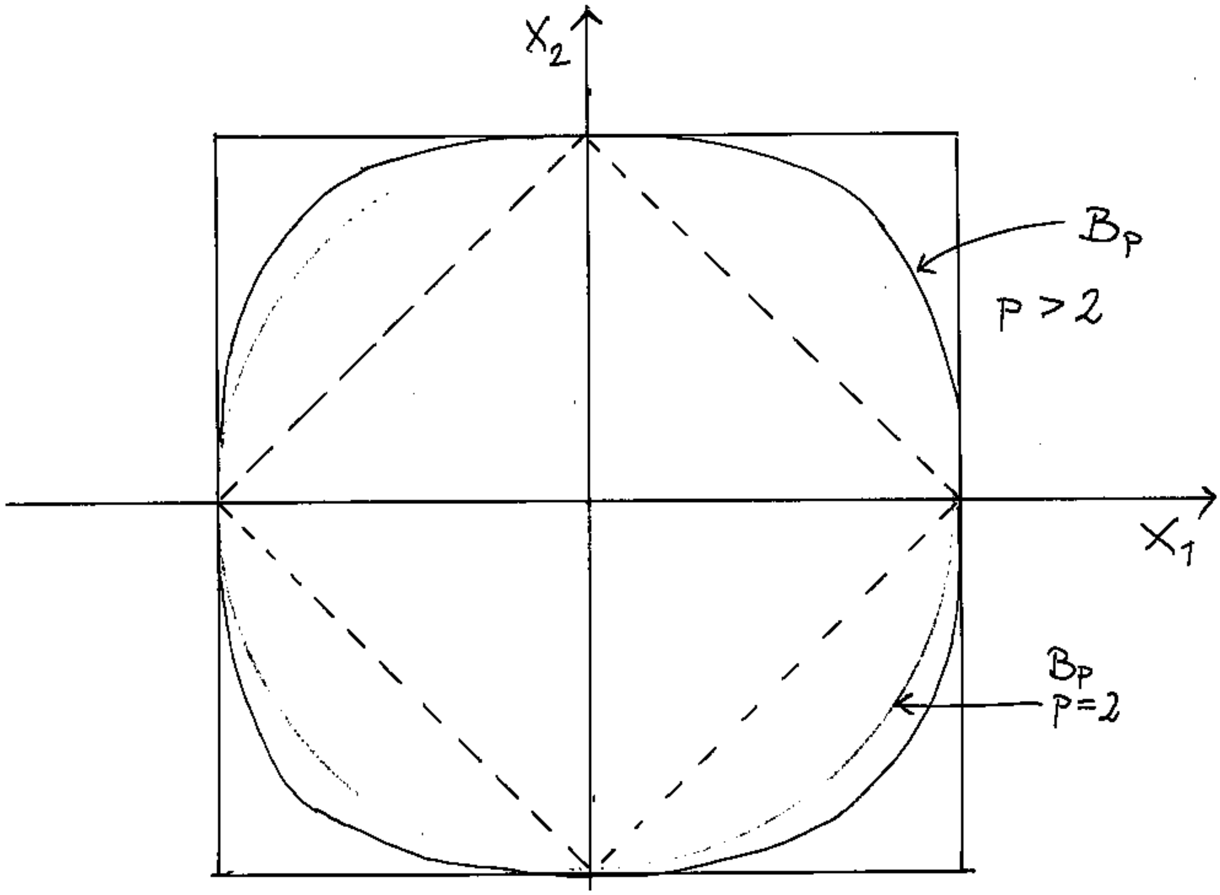
$$B_p := \{x \in \mathbb{R}^2 \mid \|x\|_p \leq 1\}$$

For $p = 1$ og $p = \infty$ fås $B_1 = \{(x_1, x_2) \mid |x_1| + |x_2| \leq 1\}$ og $B_\infty = \{(x_1, x_2) \mid \max\{|x_1|, |x_2|\} \leq 1\}$





Disse “kugler har flade sider. Kun den “rigtige” kugle for $p = 2$ er “kuglerund”, de andre ligger ind imellem, og kuglerne vokser med p



Når p bevæger sig fra 1 til ∞ deformeres med B_p gradvist fra det ene kvadrat til det andet. Dette er udnyttet af Piet Hein ved konstruktionen af den såkaldte superellipse. Her er akserne dog normalt ikke lige lange. Jeg vil tro at superellipsens p ligger i intervallet $[\frac{5}{2}, 3]$, men jeg ved det ikke.

10 Dimension og relativ topologi

Lad os forestille os en konveks mængde som er et liniestykke, da siger vor intuition os at den eneste rimelige definition af dimensionen af denne mængde må være 1, uanset at liniestykket ligger i $\mathbb{R}^n$. Analogt må dimensionen af det konvekse hylster af 3 punkter i $\mathbb{R}^n$ der ikke ligger på linie være 2, og man kunne fortsætte med 4 punkter der ikke ligger i en plan, dvs 4 affint uafhængige punkter udspænder en konveks mængde af dimension 3. I alle tilfældene ses det at den naturlige definition af dimensionen af en konveks mængde er at den må være den samme som dimensionen af det affine hylster af den givne konvekse mængde. Dette afspejles i den generelle definition herunder.

Definition 10.1. Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være konveks. Dimensionen af $\mathcal{C}$ - $\dim(\mathcal{C})$ - defineres ved $\dim(\mathcal{C}) := \dim(\operatorname{aff}(\mathcal{C}))$.

Antag $\mathcal{C} \neq \emptyset$ og lad $k = \dim(\operatorname{aff}(\mathcal{C}))$, da ses det fra Sætning 2.15 samt Definition 2.16, at $\mathcal{C}$ må indeholde et sæt $(c_0, \ldots, c_k)$ af $k+1$ punkter som er affint uafhængige og udspænder $\operatorname{aff}(\mathcal{C})$, dvs.

$$\operatorname{aff}(\mathcal{C}) = \{\alpha_0 c_0 + \cdots + \alpha_k c_k | \alpha_0 + \cdots + \alpha_k = 1\}\,.$$

Heraf kan man se, at dimensionen af $\mathcal{C}$ også kan måles lokalt, dvs. indenfor $\mathcal{C}$ selv, som det tal der er bestemt som: "det maksimale antal affint uafhængige punkter der kan udtages fra $\mathcal{C}$ minus 1". Dette kan formuleres lidt mere formelt som

Sætning 10.2. *Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være konveks og ikke–tom. Dimensionen $k = \dim(\mathcal{C})$ er karakteriseret ved at der findes en delmængde $\{c_0, \ldots, c_k\}$ af $\mathcal{C}$ bestående af $k+1$ punkter som er affint uafhængige, men enhver delmængde fra $\mathcal{C}$ som indeholder $k+2$ elementer eller flere er affint afhængig.*

Konvekse delmængder i $\mathbb{R}^n$ som ikke har fuld dimension (n), kan ikke have indre punkter, da en hyperplan $H(y, \alpha)$, dvs. et affint underrum af dimension $(n-1)$ ikke har indre punkter. (Se Kapitel 4 aspekt 4).

Vi starter derfor den topologiske undersøgelse af konvekse mængder med en undersøgelse af konvekse delmængder af dimension n.

Sætning 10.3. *Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være konveks og $\dim(\mathcal{C}) = n$, da er det indre $\overset{\circ}{\mathcal{C}} \neq \emptyset$.*

Bevis. Eftersom $\dim(\mathcal{C}) = n$ findes der et sæt $(c_0, \ldots, c_n)$ bestående af $n+1$ affint uafhængige punkter fra $\mathcal{C}$. Vi kan derfor definere en affin isomorfi, dvs. en bijektiv affin afbildning $F : \mathbb{R}^n \to \mathbb{R}^n$ ved

$$F(0) = c_0\,,\ F(e_j) = c_j\,,\ 1 \leq j \leq n\,.$$

I følge Sætning 2.10 findes der en regulær $n \times n$ matrix $A \in \mathbf{M}(n,n)$, så $F(x) = Ax + c_0$ for $x \in \mathbb{R}^n$. Den inverse afbildning $F^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ er også affin, idet den er givet ved $F^{-1}(y) = A^{-1}y - A^{-1}c_0$. Ifølge Sætning 1.5 er F^{-1} kontinuert, hvilket medfører at originalmængden til en åben mængde $\mathcal{V}$ for F^{-1} er åben. Skrevet i symboler giver dette: $(F^{-1})^{-1}(\mathcal{V})$ er åben og dermed for enhver åben mængde $\mathcal{V} \subseteq \mathbb{R}^n$ er $F(\mathcal{V}) = (F^{-1})^{-1}(\mathcal{V})$ en åben mængde. Dette kan udnyttes til at vise, at $\mathcal{C}$ har indre punkter. Lad nemlig $\mathcal{V} \subseteq \mathbb{R}^n$ være givet ved

$$\mathcal{V} = \{x \in \mathbb{R}^n | x_1 > 0, \ldots, x_n > 0\,,\, x_1 + \cdots + x_n < 1\}\,.$$

Så er $\mathcal{V}$ åben da $\mathcal{V}$ er fællesmængden af $(n+1)$ åbne halvrum. Yderligere er $\mathcal{V} \neq \emptyset$, da $T := \left(\frac{1}{n+1}, \ldots, \frac{1}{n+1}\right) \in \mathcal{V}$ $\left(\frac{1}{n+1} + \cdots + \frac{1}{n+1} = \frac{n}{n+1} < 1\right)$.

En simpel udregning nedenfor viser at $F(\mathcal{V}) \subseteq \mathcal{C}$, så $\overset{\circ}{\mathcal{C}} \supseteq F(\mathcal{V}) \neq \emptyset$:
For

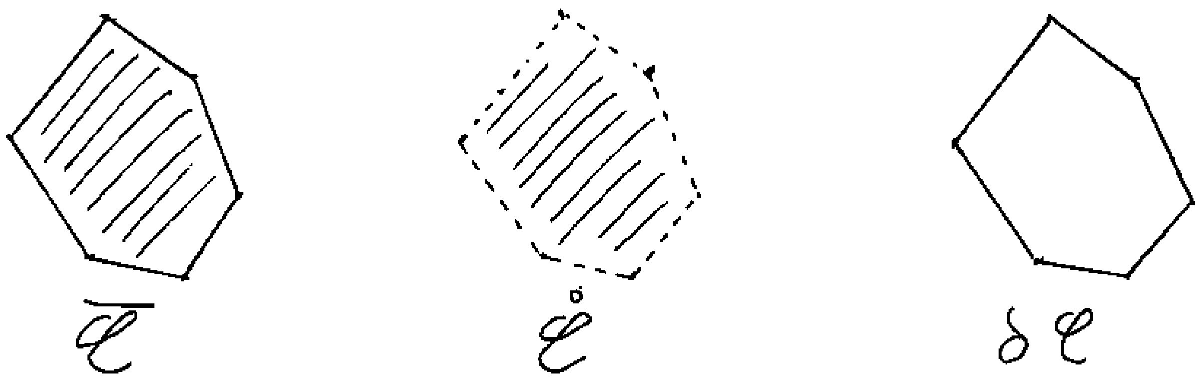
$$x = (x_1, \ldots, x_n) \in \mathbb{R}^n \text{ er } x = (1 - x_1 - \cdots - x_n)0 + x_1 e_1 + \cdots + x_n e_n\,,$$

en affin kombination af $(0, e_1, \ldots, e_n)$, så

$$F(x) = (1 - x_1 - x_2, \ldots, -x_n)c_0 + x_1 c_1 + \cdots + x_n c_n\,.$$

Når $x \in \mathcal{V}$ bliver $F(x)$ derfor en konveks kombination af $(c_0, c_1, \ldots, c_n)$, så $F(\mathcal{V}) \subseteq \mathcal{C}$, og det indre af $\mathcal{C}$ er ikke–tomt idet $\overset{\circ}{\mathcal{C}}$ indeholder den åbne mængde $F(\mathcal{V})$. $\square$

Det næste resultat er en af klassikerne indenfor konveksitetsteori. Det udtaler "geometrisk", at når det indre af en konveks mængde $\mathcal{C}$ ikke er tomt, så er randen af $\mathcal{C}$ en skal, som antydet på tegningen

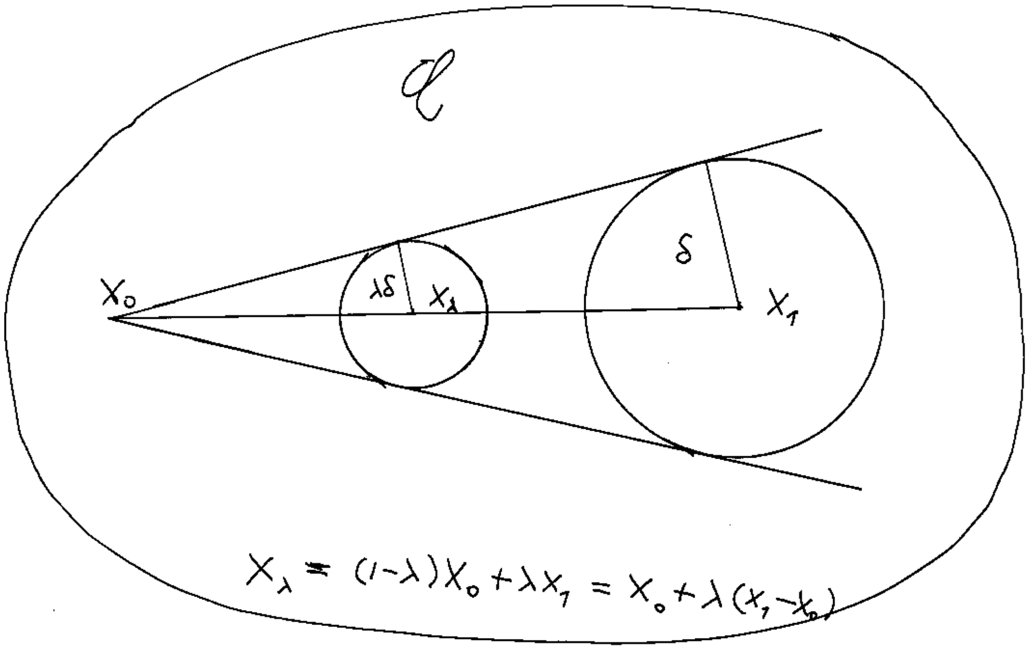


Betydningen af disse vage udsagn fremgår af Sætning 10.6.

Som eksempel på mængder der ikke har sådanne pæne egenskaber, kan man betragte $\mathbb{Q} \subseteq \mathbb{R}$. Her er $\overline{\mathbb{Q}} = \mathbb{R}$, og $\overset{\circ}{\mathbb{Q}} = \emptyset$, så $\partial(\mathbb{Q}) = \mathbb{R}$.

Lemma 10.4. *Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være konveks med ikke–tomt indre. For $x_0 \in \overline{\mathcal{C}}$ og $x_1 \in \overset{\circ}{\mathcal{C}}$ gælder $]x_0, x_1] \subseteq \overset{\circ}{\mathcal{C}}$.*

Bevis. Antag først, at $x_0 \in \mathcal{C}$ og $\delta > 0$ er valgt, så $K(x_1, \delta) \subseteq \mathcal{C}$. For $0 < \lambda < 1$ lad $x_\lambda = (1-\lambda)x_0 + \lambda x_1 = x_0 + \lambda(x_1 - x_0)$, og det ses, at $\|x_\lambda - x_0\| = \lambda \|x_1 - x_0\|$. Betragt nu tegningen



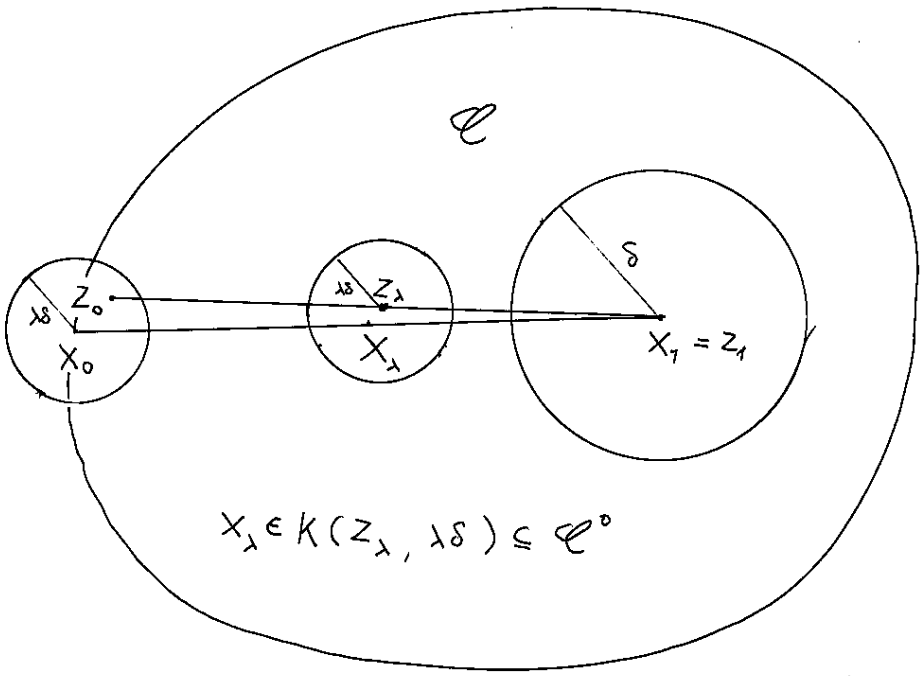
Det ses ved ligedannethed i forholdet $\lambda : 1$, at $K(x_\lambda, \lambda\delta) \subseteq \mathcal{C}$, så $x_\lambda \in \mathring{\mathcal{C}}$. Såfremt man føler sig utryg ved dette argument, kan argumentet danne basis for et algebraisk argument, som man kan skrive ned i alle detaljer. Lad da $y \in K(x_\lambda, \lambda\delta)$, da skal jeg vise at $y \in \mathcal{C}$. Tegningen viser at jeg skal finde z som

$$z = x_1 + \frac{1}{\lambda}(y - x_\lambda).$$

Heraf ses, at $\|x_1 - z\| = \frac{1}{\lambda}\|y - x_\lambda\| < \frac{1}{\lambda}(\lambda\delta) = \delta$ så $z \in K(x_1, \delta)$, og dermed $z \in \mathcal{C}$. Tager jeg nu den konvekse kombination $(1-\lambda)x_0 + \lambda z$ fås et element i $\mathcal{C}$, men

$$(1-\lambda)x_0 + \lambda z = (1-\lambda)x_0 + \lambda x_1 + (y - x_\lambda) = x_\lambda + (y - x_\lambda) = y.$$

Hermed er påstanden i lemmaet bevist når $x_0 \in \mathcal{C}$. Antag nu $x_0 \in \overline{\mathcal{C}} \setminus \mathcal{C}$ som antydet på tegningen herunder, og lad igen $0 < \lambda < 1$, $x_\lambda = (1-\lambda)x_0 + \lambda x_1 = x_0 + \lambda(x_1 - x_0)$.

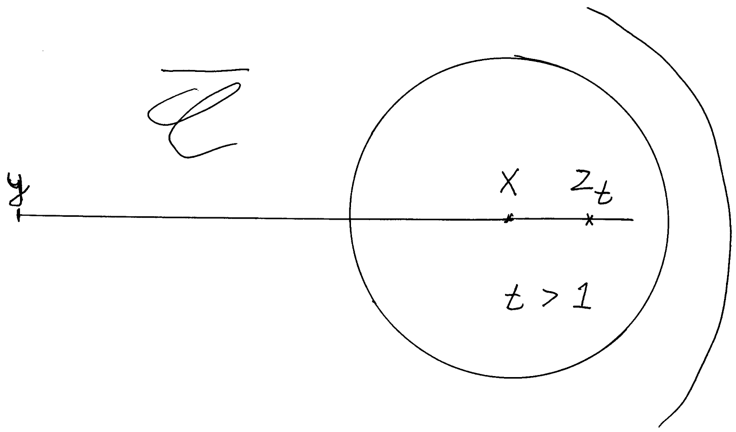


Da $x_0 \in \overline{\mathcal{C}}$ findes $z_0 \in K(x_0, \lambda\delta) \cap \mathcal{C}$, og for $z_\lambda := (1-\lambda)z_0 + \lambda x_1$ giver resultatet ovenfor, at $K(z_\lambda, \lambda\delta) \subseteq \mathcal{C}$. Da $\|x_\lambda - z_\lambda\| = \|(1-\lambda)(z_0 - x_0)\| \leq \|z_0 - x_0\| < \lambda\delta$, får vi $x_\lambda \in K(z_\lambda, \lambda\delta) \subseteq \mathcal{C}$, men også $x_\lambda \in \mathring{\mathcal{C}}$, da $K(z_\lambda, \lambda\delta)$ er en åben mængde indeholdt i $\mathcal{C}$, og dermed indeholdt i $\mathring{\mathcal{C}}$. $\square$

Sætning 10.5. *Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være konveks med ikke–tomt indre, da gælder $\overline{\mathcal{C}} = \overline{\mathring{\mathcal{C}}}$ og $\mathring{\overline{\mathcal{C}}} = \mathring{\mathcal{C}}$.*

Bevis. Da $\mathring{\mathcal{C}} \subseteq \mathcal{C}$ er $\overline{\mathring{\mathcal{C}}} \subseteq \overline{\mathcal{C}}$. Lad nu $x \in \overline{\mathcal{C}}$, $y \in \mathring{\mathcal{C}}$, og definer $z_n := (1 - \frac{1}{n})x + \frac{1}{n}y$. Ifølge Lemma 10.4 er $z_n \in \mathring{\mathcal{C}}$, men $z_n \to x$ for $n \to \infty$, så $x \in \overline{\mathring{\mathcal{C}}}$ og $\overline{\mathring{\mathcal{C}}} = \overline{\mathcal{C}}$.

Da $\mathcal{C} \subseteq \overline{\mathcal{C}}$ er $\mathring{\mathcal{C}} \subseteq \mathring{\overline{\mathcal{C}}}$. Lad nu $x \in \mathring{\overline{\mathcal{C}}}$ og $\delta > 0$ være valgt så $K(x, \delta) \subseteq \overline{\mathcal{C}}$ og vælg y tilfældigt i $\mathring{\mathcal{C}}$. Betragt da liniestykket $[y, x] = \{(1-t)y + tx \mid 0 \leq t \leq 1\}$ og lad z_t betegne punkterne herpå, $z_t := (1-t)y + tx$, $z_0 = y$ og $z_1 = x$.



Da $K(x,\delta) \subseteq \overline{\mathcal{C}}$ kan liniestykket fortsættes ud over x $(t > 1)$, således at $z_t \in K(x,\delta) \subseteq \overline{\mathcal{C}}$, når t ikke er alt for meget større end 1. Igen giver tegningen fundamentet for en eksakt udregning

$$\|z_t - x\| = \|(1-t)y + (t-1)x\| = |t-1|\,\|x-y\|\,,$$

så $z_t \in K(x,\delta)$ når $|t-1| < \frac{\delta}{\|x-y\|}$, dvs. for $t = \frac{\delta}{2\|x-y\|}$ er $z_t \in \overline{\mathcal{C}}$, og da $x \in [y, z_t[$, er $x \in \overset{\circ}{\mathcal{C}}$ ifølge Lemma 10.4.

Bemærkning: En lidt mere "moden" måde at præsentere det foregående argument — efter tegningen — ville være at sige: Lad z betegne et punkt på forlængelsen af liniestykket $[y,x]$ ud over x. Sålænge $\|x-z\| < \delta$ ligger z i $\overline{\mathcal{C}}$, og dermed fås fra Lemma 10.4, idet $x \in [y, z[$ at $x \in \overset{\circ}{\mathcal{C}}$. □

Der kan drages endnu nogle konklusioner af Lemma 10.4, men denne gang vedrørende snit mellem konvekse mængder og linier i $\mathbb{R}^n$.

Sætning 10.6. *Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være en afsluttet konveks mængde, og lad ℓ betegne en linie med retningsvektor $b \neq 0$ gennem et punkt c fra $\mathcal{C}$ – $\ell = \{c + tb | t \in \mathbb{R}\}$ Der gælder da:*

(i) $I := \{t \in \mathbb{R} | c + tb \in \mathcal{C}\}$ *er et afsluttet interval i* $\mathbb{R}^n$ *og* $0 \in I$.

(ii) I *er opadtil begrænset hvis og kun hvis* $b \notin \operatorname{recc}(\mathcal{C})$.

(iii) I *er nedadtil begrænset hvis og kun hvis* $-b \notin \operatorname{recc}\mathcal{C}$.

(iv) Hvis I *er et begrænset interval* $[\alpha, \beta]$ *og* $c \in \overset{\circ}{\mathcal{C}}$ *gælder*

$$\begin{array}{ll} c + tb \notin \mathcal{C} & \text{for } -\infty < t < \alpha \text{ eller } \beta < t < \infty \\ c + tb \in \delta(\mathcal{C}) & \text{for } t = \alpha \text{ eller } t = \beta \\ c + tb \in \overset{\circ}{\mathcal{C}} & \text{for } \alpha < t < \beta\,. \end{array}$$

Bevis.

Ad (i): Lad $F : \mathbb{R} \to \mathbb{R}^n$ være givet ved $F(t) = c + tb$, da er F affin og kontinuert og $I = F^{-1}(\mathcal{C})$, som derfor er afsluttet og konveks. Ifølge en opgave er de konvekse delmængder af $\mathbb{R}$ netop intervallerne. Da $F(0) = c \in \mathcal{C}$, er $0 \in I$, og (i) følger.

Ad (ii): Hvis $b \in \mathrm{recc}(\mathcal{C})$, er $tb \in \mathrm{recc}(\mathcal{C})$ for $t \geq 0$, da $\mathrm{recc}(\mathcal{C})$ er en konveks kegle, men så er $c + tb \in \mathcal{C}$ for $t \geq 0$ og $I \supseteq [0, \infty[$.

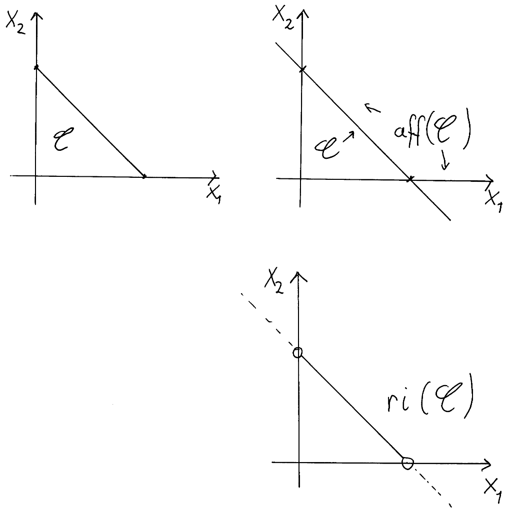
Hvis I ikke er opadtil begrænset indehholder I intervallet $[0, \infty[$og fra Lemma 7.2 fås da at $b \in \mathrm{recc}(\mathcal{C})$, og (ii) er vist.

Ad (iii): Helt analogt bevis som beviset for (ii). Prøv selv!

Ad (iv): Dette er en direkte konsekvens af Lemma 10.4. □

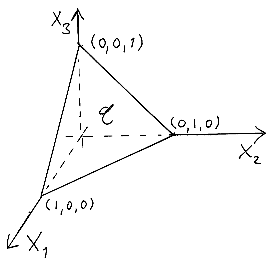
Vi skal nu undersøge tilfældet hvor $\dim(\mathcal{C}) < n$, så er det indre af $\mathcal{C}$ tomt, og vi har et problem med at diskutere $\mathcal{C}$'s topologiske egenskaber. Problemet er imidlertid ikke så stort, da $\mathcal{C}$ er indeholdt i $\mathrm{aff}(\mathcal{C})$, og i forhold til dette lavere–dimensionale rum, giver det god mening at tale om (relativt) indre punkter.

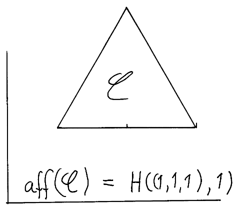
Eksempel 10.7. Lad $\mathcal{C} = \{(x_1, x_2) \in \mathbb{R}^2 | 0 \leq x_1 \, , \, 0 \leq x_2 \, , \, x_1 + x_2 = 1\}$



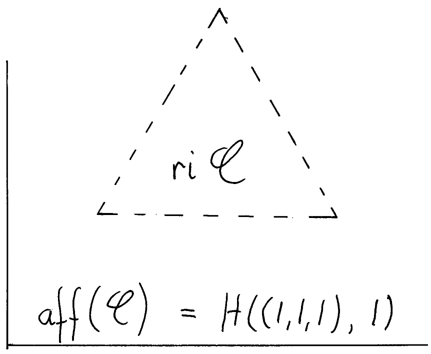
Det indre af $\mathcal{C}$ med hensyn til $\mathrm{aff}\,\mathcal{C}$.

Dette kaldes det relative indre af $\mathcal{C}$ og betegnes $\mathrm{ri}\,(\mathcal{C})$ og er her givet som $\{(x_1, x_2) \in \mathbb{R}^2 | x_1 > 0 \, , \, x_2 > 0 \, , \, x_1 + x_2 = 1\}$. Vi kigger nu på et tilsvarende eksempel i 3 dimensioner. Lad $\mathcal{C} = \{(x_1, x_2, x_3) \in \mathbb{R}^3 | x_1 \geq 0 \, , \, x_2 \geq 0 \, , \, x_3 \geq 0 \, , \, x_1 + x_2 + x_3 = 1\}$.





$\mathrm{aff}(\mathcal{C}) = H((1,1,1),1) = \{(x_1,x_2,x_3)|x_1+x_2+x_3=1\}$.
Relativt til $\mathrm{aff}(\mathcal{C})$ ser situationen således ud:



Forhåbentligt har disse eksempler givet dig en intuition med hensyn til hvorledes man skal forstå eller opfatte begrebet, relativt indre. Det bliver defineret helt præcist i Definition 10.8. Her vil jeg blot ende den intuitive gennemgang med at argumentere for, at enhver ikke–tom konveks mængde har et ikke–tomt relativt indre. Hvis $\dim(\mathcal{C}) = k$, er $\mathrm{aff}(\mathcal{C}) = x_0 + \mathcal{U}$, hvor $\mathcal{U}$ er et k–dimensionalt lineært underrrum af $\mathbb{R}^n$. For $\mathcal{D} := \mathcal{C} - x_0$ fås da, at $\mathcal{D} \subseteq \mathcal{U}$, $\dim \mathcal{D} = \dim \mathcal{U} = k$, men $\mathcal{U}$ har en ortonormalbasis, $f_1, \ldots, f_k$ og kan dermed via afbildningen $\mathbb{R}^k \ni (x_1, \ldots, x_k) \to x_1 f_1 + \cdots + x_k f_k \in \mathcal{U}$ opfattes som værende $\mathbb{R}^k$.

Betragtes $\mathcal{D}$ som konveks delmængde af $\mathbb{R}^k$, viser Sætning 10.3 at $\mathring{\mathcal{D}} \neq \emptyset$, dvs. at som delmængde af $\mathbb{R}^n$, er det relative indre af $\mathcal{D}$ ikke–tomt, og dermed er det relative indre af $\mathcal{C}$ ikke–tomt.

Definition 10.8. Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være en ikke–tom konveks mængde. Ved det relative indre af $\mathcal{C}$ — $\mathrm{ri}\,(\mathcal{C})$ — forstås det indre af $\mathcal{C}$ opfattet som delmængde af det metriske rum $(\mathrm{aff}(\mathcal{C}), d)$, hvor $d(a,b) = \|a-b\|$. Hvis $\mathcal{C} = \emptyset$ er det relative indre af $\mathcal{C}$ den tomme mængde.

Sætning 10.9. *Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være konveks og ikke–tom, da er* $\mathrm{ri}\,(\mathcal{C})$ *ikke–tom.*

Bevis. Lad $k = \dim(\mathcal{C})$. Hvis $k = 0$ er $\mathcal{C} = \{c\}$ et enkelt punkt, så $\operatorname{aff}(\mathcal{C}) = \{c\} = \mathcal{C}$, og $\operatorname{ri}(\mathcal{C}) = \overset{\circ}{\operatorname{aff}(\mathcal{C})}$ som delmængde af $\operatorname{aff}(\mathcal{C})$, dvs. $\operatorname{ri}(\mathcal{C}) = \overset{\circ}{\operatorname{aff}(\mathcal{C})} = \{c\} \neq \emptyset$. Hvis $k \geq 1$, er $\operatorname{aff}(\mathcal{C}) = z_0 + \mathcal{U}$, hvor $\mathcal{U}$ er et lineært underrum af $\mathbb{R}^n$ af dimension $k \geq 1$. Derfor har $\mathcal{U}$ en ortonormalbasis $(f_1, \dots, f_k)$. Vi kan derfor definere en isometri F af $\mathbb{R}^k$ på $\operatorname{aff}(\mathcal{C})$ ved

$$F(x_1, \dots, x_k) = z_0 + x_1 f_1 + \dots + x_k f_k \,.$$

Det er klart at F er surjektiv, og isometriegenskaben følger af, at for $x = (x_1, \dots, x_k)$, $y = (y_1, \dots, y_k) \in \mathbb{R}^k$ er

$$\|F(x) - F(y)\| = \left\| \sum_{i=1}^{k} (x_i - y_i) f_i \right\| = \left(\sum_{i=1}^{k} (x_i - y_i)^2 \right)^{1/2} = \|x - y\| \,.$$

Heraf følger at de metriske rum $(\mathbb{R}^k, d)$ og $(\operatorname{aff}(\mathcal{C}), d)$ har egenskaben

$$\forall \mathcal{M} \subseteq \operatorname{aff}(\mathcal{C}) : \mathcal{M} \text{ er åben i } (\operatorname{aff}(\mathcal{C}), d) \Leftrightarrow F^{-1}(\mathcal{M}) \text{er åben i } \mathbb{R}^k.$$

Specielt for $\mathcal{C}$ fås da $\operatorname{ri}(\mathcal{C}) = F(\overset{\circ}{F^{-1}(\mathcal{C})})$. Da F er affin, følger det fra Sætning 5.4.4 at $F^{-1}(\mathcal{C})$ er konveks, og vi skal så blot overbevise os om, at $\dim(F^{-1}(\mathcal{C})) = k$ for at kunne anvende Sætning 10.3 til at indse, at $\operatorname{ri}(\mathcal{C}) \neq \emptyset$. Til den ende betragter vi et affint uafhængigt sæt $(c_0, c_1, \dots, c_k)$ bestående af $k + 1$ elementer fra $\mathcal{C}$. Sættet $d_0 = F^{-1}(c_0), \dots, d_i = F^{-1}(c_i)$, $d_k = F^{-1}(c_k)$ ligger i $\mathcal{D} := F^{-1}(\mathcal{C})$, og må være affint uafhængigt, da F er en affin isomorfi af $\mathbb{R}^k$ på $\operatorname{aff}(\mathcal{C})$. Eftersom $\mathcal{D} = F^{-1}(\mathcal{C})$ er en delmængde af $\mathbb{R}^k$ og indeholder et affint uafhægigt sæt bestående af $k + 1$ vektorer må $\mathcal{D}$ have dimension k i $\mathbb{R}^k$ ifølge Sætning 10.3. Heraf ses så at $\operatorname{ri}(\mathcal{C}) = F(\overset{\circ}{\mathcal{D}}) \neq \emptyset$. □

Det kan synes lidt tungt at gå hele vejen rundt om $F : \mathbb{R}^k \to \operatorname{aff}(\mathcal{C}) \subseteq \mathbb{R}^n$ for at etablere dette resultat. Grunden til at jeg har valgt denne fremstilling er, at jeg har erfaring for, at nogle studerende føler sig utrygge ved de intuitive argumenter der blev præsenteret før Sætning 10.9. Når denne utryghed har forladt dig efter at du har arbejdet med dette bevis, vil jeg anbefale dig at tænke geometrisk således som indledningen til emnet — Eksempel 10.7 — lægger op til. Uden beviser vil jeg derfor tillade mig at nævne følgende resultater:

Lemma 10.10 (Lemma 10.4 generaliseret). *Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være konveks.*
For $x_0 \in \overline{\mathcal{C}}$ og $x_1 \in \operatorname{ri}(\mathcal{C})$:$]x_0, x_1] \subseteq \operatorname{ri}(\mathcal{C})$.

Sætning 10.11 (Sætning 10.5 generaliseret). *Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være konveks, da gælder*

$$\operatorname{ri}(\overline{\mathcal{C}}) = \operatorname{ri}(\mathcal{C}) \,, \quad \overline{\operatorname{ri}(\mathcal{C})} = \overline{\mathcal{C}} \,.$$

Sætning 10.12 (Konsekvens af Sætning 10.6). *Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være konveks. Lad $x_0 \in \operatorname{ri}(\mathcal{C})$ og $x_1 \in \mathcal{C}$. Hvis $x_1 - x_0 \notin \operatorname{recc}(\mathcal{C})$ findes $\beta \geq 1$, så $x_0 + \lambda(x_1 - x_0) \in \operatorname{ri}(\mathcal{C})$ for $0 \leq \lambda < \beta$, og $x_0 + \lambda(x_1 - x_0) \notin \overline{\mathcal{C}}$ for $\lambda > \beta$ samt $x_0 + \beta(x_1 - x_0) \in \overline{\mathcal{C}}$.*

11 En vanskeligere separationssætning

Dette afsnit indeholder kun ét resultat.

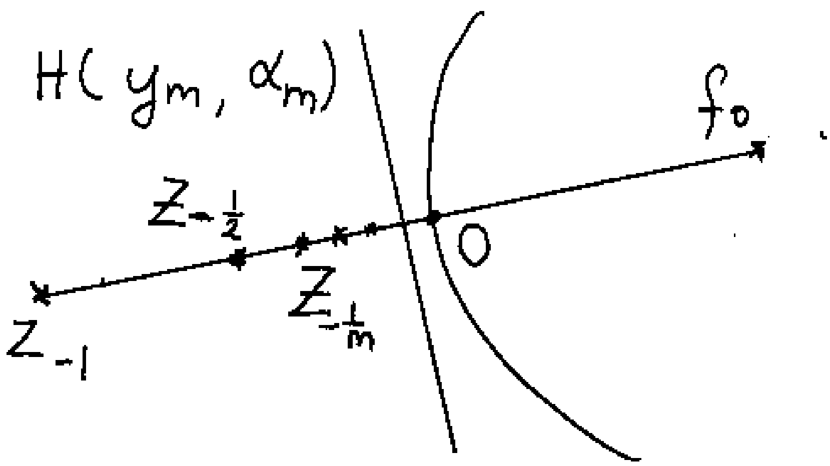
Sætning 11.1. *Lad $\mathcal{C}$ og $\mathcal{D}$ være disjunkte konvekse ikke–tomme delmængder af* $\mathbb{R}^n$*, da kan $\mathcal{C}$ og $\mathcal{D}$ separeres ved en hyperplan.*

Dette er ikke den bedst mulige separationssætning man kan få, men den bedst mulige kræver meget mere teori og abstrakt tænkning, så vi nøjes med det næstbedste, som da også er godt nok i mange tilfælde. Står man nu en dag og mangler den bedste vare er det så heldigt, at man gerne må benytte den, selv om man ikke har set beviset, og så er skaden jo til at overse. Den bedste sætning er

Sætning 11.2. *Lad $\mathcal{C}$ og $\mathcal{D}$ være konvekse ikke–tomme delmængder af* $\mathbb{R}^n$*, da kan $\mathcal{C}$ og $\mathcal{D}$* separeres egentligt *ved en hyperplan hvis og kun hvis* $\mathrm{ri}\,(\mathcal{C}) \cap \mathrm{ri}\,(\mathcal{D}) = \emptyset$.

Bevis. (Sætning 11.1) Som i beviset for Sætning 8.5 — den første separationssætning — betragter vi $\mathcal{F} := \mathcal{D} - \mathcal{C} = \{d - c \mid c \in \mathcal{C}\,,\, d \in \mathcal{D}\}$. Forudsætningen siger, at $0 \notin \mathcal{F}$. Hvis $0 \notin \overline{\mathcal{F}}$ er $\mathrm{dist}(\mathcal{C}, \mathcal{D}) > 0$, og Sætning 8.5 kan anvendes til at opnå, endog, stærk separation. Vi kan derfor antage, at $0 \in \overline{\mathcal{F}}$ (og $0 \notin \mathcal{F}$). Lad $f_0 \in \mathrm{ri}\,(\mathcal{F})$ (Sætning 10.9) og definer for $t \in \mathbb{R}, 0 < t \leq 1$, $z_t = (1-t)\cdot 0 + tf_0 = tf_0$. Da er $z_t \in \mathrm{ri}\,(\mathcal{F})$ for $0 < t \leq 1$ (Sætning 10.12). Endvidere ses det at for $t = 0$ er $z_0 = 0 \notin \mathrm{ri}\,(\mathcal{F})$, da $\mathrm{ri}\,(\mathcal{F}) \subseteq \mathcal{F}$.

Som konsekvens af Lemma 10.10 eller Sætning 10.12 fås da, at for $t < 0$ er $z_t \notin \overline{\mathcal{F}}$ og Sætning 8.5 medfører så at for $m \in \mathbb{N}$, at $z_{-\frac{1}{m}}$ kan separeres stærkt fra $\mathcal{F}$.



Lad $y_m \in \mathbb{R}^n$ og α_m være fundne, så $H(y_m, \alpha_m)$ separerer $z_{-\frac{1}{m}}$ og $\mathcal{F}$ stærkt. Et øjebliks overvejelse viser, at for $s \in \mathbb{R}_+$(dvs.$s > 0$) er

$$J(sy_m, s\alpha_m) = \{x \in \mathbb{R}^n | s(y_m \cdot x) \leq s\alpha_m\} = J(y_m, \alpha_m)\,.$$

Det vil således ikke skade at antage at $\|y_m\| = 1$, thi hvis $\|y_m\| \neq 1$, kan vi i stedet for y_m benytte $\frac{1}{\|y_m\|}y_m$, og i stedet for α_m, $\frac{1}{\|y_m\|}\alpha_m$.

Følgen $(y_m)_{m\in\mathbb{N}}$ er indeholdt i enhedssfæren $\mathcal{S} = \{y \in \mathbb{R}^n \mid \|y\| = 1\}$. Mængden $\mathcal{S}$ er klart begrænset og lukket så $\mathcal{S}$ er kompakt, og $(y_m)_{m\in\mathbb{N}}$ har en konvergent delfølge $(y_{m_i})_{i\in\mathbb{N}}$, $y_{m_i} \to y$ for $i \to \infty$.

Eftersom $H(y_{m_i}, \alpha_{m_i})$ separerer $z_{-\frac{1}{m_i}}$ og $\overline{\mathcal{F}}$ stærkt, og $0 \in \overline{\mathcal{F}}$, fås for $f \in \overline{\mathcal{F}}$

$$-\frac{1}{m_i}f_0 \cdot y_{m_i} = z_{-\frac{1}{m_i}} \cdot y_{m_i} < \alpha_{m_i} < y_{m_i} \cdot f\,.$$

Venstresiden vurderes ved

$$|-\frac{1}{m_i}f_0 \cdot y_{m_i}| \leq \frac{1}{m_i}\|f_0\|\,\|y_{m_i}\|\,,$$

så

$$(\|y_{m_i}\| = 1)\,,\, -\frac{1}{m_i}\|f_0\| < \alpha_{m_i} < y_{m_i} \cdot f\,.$$

For $f = 0$ fås $\alpha_{m_i} < 0$.

Følgelig fås $\alpha_{m_i} \to 0$ for $i \to \infty$, og $y \cdot f = \lim_{i\to\infty} y_{m_i} \cdot f \geq \lim \alpha_{m_i} = 0$.

Idet $\mathcal{F} = \mathcal{D} - \mathcal{C}$, fås for $d \in \mathcal{D}$ og $c \in \mathcal{C}$, $y \cdot d \geq y \cdot c$. Vælges nu $c_0 \in \mathcal{C}$ og $d_0 \in \mathcal{D}$ ses det, at $\{y \cdot c | c \in \mathcal{C}\}$ er opadtil begrænset af $y \cdot d_0$, og $\{y \cdot d | d \in \mathcal{D}\}$ er nedadtil begrænset af $y \cdot c_0$. Det giver derfor mening at definere $\gamma = \sup\{y \cdot c | c \in \mathcal{C}\}$ og $\beta = \inf\{y \cdot d | d \in \mathcal{D}\}$, og vi har $\gamma \leq \beta$ samt $\mathcal{C} \subseteq J(y, \gamma)$ og $\mathcal{D} \subseteq J(-y, -\beta)$, og dermed separerer $H(y, \gamma)$ mængderne $\mathcal{C}$ og $\mathcal{D}$. □

Bemærkning 11.3.

Vi kan komme lidt tættere på det optimale resultat, idet vi får det gratis hvis $\dim(\mathcal{D} - \mathcal{C}) = n$. I dette tilfælde kan $\mathcal{D}$ og $\mathcal{C}$ ikke begge være indeholdt i en hyperplan $H(y, \alpha)$, da $\mathcal{D} - \mathcal{C}$ så ville være indeholdt i det lineære underrum $\{y\}^\perp$ i modstrid med antagelsen om, at $\dim(\mathcal{D} - \mathcal{C}) = n$. Altså fås egentlig separation.

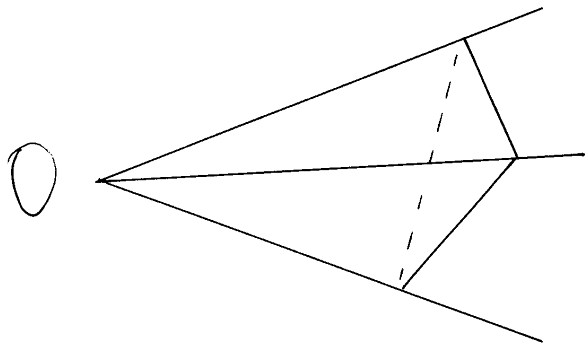
Som konsekvens af Sætning 11.2 kan man se, at hvis $x \in \overline{\mathcal{C}} \setminus \mathrm{ri}\,(\mathcal{C})$, så kan de konvekse mængder $\{x\}$ og $\mathrm{ri}\,(\mathcal{C})$ separeres egentligt, dvs. der findes $H(y, \alpha)$ så $x \cdot y \leq \alpha$, og for $c \in \mathrm{ri}\,(\mathcal{C})$ gælder $c \cdot y \geq \alpha$, og der findes $c_0 \in \mathrm{ri}\,(\mathcal{C})$ med $c_0 \cdot y > \alpha$. Da nu $x \in \overline{\mathcal{C}} = \overline{(\mathrm{ri}\,\mathcal{C})}$ er $x \cdot y \geq \alpha$ så $x \in H(y, \alpha)$ og $\mathcal{C} \subseteq J(-y, -\alpha)$, men $\mathcal{C} \nsubseteq H(y, \alpha)$. Der findes altså en støttehyperplan $H(y, \alpha)$ gennem hvert punkt $x \in \overline{\mathcal{C}} \setminus \mathrm{ri}\,(\mathcal{C})$, således at $\mathcal{C} \nsubseteq H(y, \alpha)$.

12 Ekstremalpunkter, ekstremale retninger og polyedre

Når man betragter en konveks polygon i planen er det nok klart, at den er det konvekse hylster af sine hjørner. Ligeledes er det klart, at "hvis ordet ekstremalpunkt" skal tolkes efter pålydende, så må det være hjørnerne der ér ekstremalpunkterne for den konvekse polygon.

Herunder vil vi give en definition på begrebet ekstremalpunkt, for så umiddelbart efter, at vise, at en vilkårlig kompakt konveks mængde er det afsluttede konvekse hylster af sine ekstremalpunkter.

Der gælder en lignende sætning for konvekse kegler, men bortset fra 0 – origo, har en konveks kegle ingen ekstremalpunkter ($a = \frac{1}{2}(\frac{1}{2}a) + \frac{1}{2}(\frac{3}{2}a)$). Tegningen herunder



illustrerer derimod, at en afsluttet konveks kegle har *"ekstremale retninger"*; i det viste tilfælde er der 3 ekstremale retninger, og det er klart, at keglen netop er den konvekse kegle frembragt af 3 vektorer, én for hver af kanterne. Dette resultat vil blive generaliseret nærmest ordret i Sætning 12.6.

Ovenstående 2 resultater kombineres til et generelt resultat om lukkede konvekse mængder. Nu er de ubegrænsede konvekse mængder ikke alle konvekse kegler, f.eks. er $\{(x_1, x_2) \in \mathbb{R}^2 | x_2 \geq x_1^2\}$ ubegrænset og konveks, men langt fra at være en kegle. Det er alligevel muligt at få information om de ubegrænsede lukkede konvekse mængder. Lad $\mathcal{C}$ være en sådan, da dekomponeres $\mathcal{C}$ først ved hjælp af Sætning 7.3. Lad $\mathcal{L} = \text{recc}(\mathcal{C}) \cap (-\text{recc}(\mathcal{C}))$ – det største lineære underrum indeholdt i $\mathcal{C}$, så er

1. $\mathcal{C} = \mathcal{L} + \mathcal{L}^\perp \cap \mathcal{C}$; $\mathcal{D} := \mathcal{L}^\perp \cap \mathcal{C}$.

2. (Sætning 12.7). En afsluttet konveks mængde $\mathcal{D}$ uden linier opfylder
   $\mathcal{D} = \text{conv}\,(\text{ext}(\mathcal{D})) + \text{recc}(\mathcal{D})$.

3. (Sætning 12.8). En afsluttet konveks mængde $\mathcal{D}$ uden linier, er det konvekse hylster af summen af sine ekstremalpunkter og sine ekstremale retninger.

Kapitlet afsluttes med at anvende Sætning 12.8 på polyedre, for derefter konkret at beskrive en metode til at finde de ekstremale retninger og ekstremalpunkterne for et polyeder uden linier.

Definition 12.1. Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være en konveks mængde. Et punkt $c \in \mathcal{C}$ siges at være et ekstremalpunkt hvis $\mathcal{C} \setminus \{c\}$ er konveks.
Mængden af samtlige ekstremalpunkter i $\mathcal{C}$ betegnes $\mathrm{ext}(\mathcal{C})$.

Bemærkning 12.2. Ifølge opgave 1.15 er $c \in \mathcal{C}$ et ekstremalpunkt i den konvekse mængde $\mathcal{C}$ hvis og kun hvis der gælder

$$\forall c_0, c_1 \in \mathcal{C} : c \in]c_0, c_1[\implies c_0 = c_1 = c\,.$$

Definition 12.3. Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være en konveks mængde.

- En halvlinie h ud fra 0 med retningsvektor $b \neq 0$, $h = \{\lambda b | \lambda \geq 0\}$, kaldes en retning for $\mathcal{C}$ hvis $\mathcal{C} + h \subseteq \mathcal{C}$, eller ækvivalent hvis $h \subseteq \mathrm{recc}(\mathcal{C})$.
- En retning $h = \{\lambda b | \lambda \geq 0\}$ for en konveks mængde $\mathcal{C}$ kaldes ekstremal hvis der gælder:
$$\forall d, e \in \mathrm{recc}(\mathcal{C}) : d + e \in h \Rightarrow d \in h \text{ og } e \in h\,.$$
- Mængden af ekstremale retninger for $\mathcal{C}$ betegnes $\mathrm{exray}(\mathcal{C})$.

Bemærkninger 12.4. Eftersom en halvlinie ud fra 0 er en retning for $\mathcal{C}$ hvis og kun hvis retningsvektoren tilhører $\mathrm{recc}(\mathcal{C})$ ses det at $\mathrm{exray}(\mathcal{C}) = \mathrm{exray}(\mathrm{recc}(\mathcal{C}))$.
Bemærk iøvrigt at hvis $\mathcal{C}$ indeholder en linie $\ell = \{\lambda b | \lambda \in \mathbb{R}\}$ så kan $\ell_+ := \{\lambda b | \lambda \geq 0\}$ ikke være en ekstremal retning da $b = 2b + (-b)$.

I den resterende del af kapitlet vil vi kun diskutere ekstremale retninger for konvekse mængder uden linier.

Sætning 12.5. *Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være konveks og kompakt, så er $\mathcal{C}$ det konvekse hylster af sine ekstremalpunkter.*

Bevis. Sætningens påstand holder klart for $\mathcal{C} = \emptyset$. Beviset føres som et induktionsbevis med induktionsudsagn $\mathcal{U}_k$ for $k \geq 0$, $(\dim(\emptyset) := -1)$.

$\mathcal{U}_k$: Enhver konveks kompakt mængde $\mathcal{C}$ af dimension højst k indeholdt i et talrum $\mathbb{R}^\ell$ er det konvekse hylster af sine ekstremalpunkter.

Induktionen starter i $k = 0$. En konveks mængde af dimension 0 er et punkt x_0; $(\text{aff}(\mathcal{C}) = \{x_0\} \implies \mathcal{C} = \{x_0\})$, $\text{ext}(\{x_0\}) = \{x_0\}$ så $\mathcal{U}_0$ er sand.

Betragt nu en konveks og kompakt delmængde $\mathcal{C}$ af $\mathbb{R}^n$ af dimension $(k+1)$, og antag, at $\mathcal{U}_k$ er sand $(k \geq 0)$. Det vises først at $\mathcal{C}$ har ekstremalpunkter. Da $k+1 \geq 1$, findes der mindst 2 punkter $x_0, x_1 \in \mathcal{C}$, $x_0 \neq x_1$. Lad $y := x_1 - x_0$, så er $y \neq 0$, og $f_y(\mathcal{C}) = \{y \cdot x | x \in \mathcal{C}\}$ er et kompakt interval i $\mathbb{R}$; $[\alpha, \beta]$, $\beta = \max_{x \in \mathcal{C}} y \cdot x$, $\alpha = \min_{x \in \mathcal{C}} y \cdot x$, da f_y er kontinuert, og $\mathcal{C}$ er kompakt og konveks. Eftersom $y \cdot x_1 - y \cdot x_0 = \|y\|^2 > 0$ er $\alpha < \beta$, så $\mathcal{D} := f_y^{-1}(\beta) \cap \mathcal{C}$ er en konveks og kompakt delmængde af $\mathcal{C}$ (Sætning 5.4.5), og $\mathcal{D} \neq \mathcal{C}$. Nu kan $\mathcal{D}$ ikke have dimension $k+1$, thi f_y vil, da den er en affin funktion, være konstant på $\text{aff}(\mathcal{D})$ med værdi β, men f_y er ikke konstant på $\mathcal{C} \subseteq \text{aff}(\mathcal{C})$, så $\text{aff}(\mathcal{D}) \underset{\neq}{\subset} \text{aff}(\mathcal{C})$. Altså er $\dim(\text{aff}(\mathcal{D})) \leq k$ og $\dim(\mathcal{D}) \leq k$. Da udsagnet $\mathcal{U}_k$ er antaget at være sandt har $\mathcal{D}$ ekstramalpunkter, og $\mathcal{D} = \text{conv}\,(\text{ext}(\mathcal{D}))$.

Lad d være ekstremalpunkt i $\mathcal{D}$, så viser de næste linier, at d også er ekstremalpunkt i $\mathcal{C}$, og dermed, at $\text{ext}(\mathcal{C})$ ikke er tom. Antag $c_0, c_1 \in \mathcal{C}$ og $d \in]c_0, c_1[$, $d = (1-\lambda)c_0 + \lambda c_1$, så er $y \cdot c_0 \leq y \cdot d = \beta$, og $y \cdot c_1 \leq y \cdot d = \beta$, og $(1-\lambda) y \cdot c_0 + \lambda y \cdot c_1 = y \cdot d = \beta$. Da $0 < \lambda < 1$, må $y \cdot c_0 = y \cdot c_1 = \beta$, så $c_0, c_1 \in \mathcal{D}$, og da d er ekstremal i $\mathcal{D}$, er $c_0 = c_1 = d$. Heraf følger, at $\mathcal{C} \setminus \{d\}$ er konveks, så d er et ekstremalpunkt i $\mathcal{C}$. Da $\text{ext}(\mathcal{C}) \neq \emptyset$, kan vi nu definere en ikke–tom konveks delmængde $\mathcal{E}$ af $\mathcal{C}$ ved $\mathcal{E} := \text{conv}\,(\text{ext}(\mathcal{C}))$. Antag nu at $c_0 \in \mathcal{C} \setminus \mathcal{E}$. Da $c_0 \notin \mathcal{E}$, kan c_0 og $\mathcal{E}$ separeres egentligt ved hjælp af en hyperplan $H(y, \alpha)$, dvs. $\mathcal{E} \subseteq J(y, \alpha)$, $c_0 \cdot y \geq \alpha$. Lad som før $\beta = \max\{f_y(c) | c \in \mathcal{C}\}$ og $\mathcal{D} = \mathcal{C} \cap f_y^{-1}(\beta)$. Som tidligere i beviset ses det at denne mængde $\mathcal{D}$ har ekstremalpunkter og at disse også er ekstremalpunkter i $\mathcal{C}$ så induktionsantagelsen giver at $\mathcal{D} \subseteq \mathcal{E}$. Vi kan yderligere slutte at $\beta = \alpha$, men så må $c_0 \cdot y = \beta$ og dermed $c_0 \in \mathcal{D} \subseteq \mathcal{E}$, og antagelsen har ført til en modstrid. □

Vi fortsætter umiddelbart med det analoge resultat for afsluttede konvekse kegler. Det er imidlertid åbenbart at resultatet ikke kan være sandt for f.eks. $\mathcal{K} = \mathbb{R}^2$, så vi skal betragte konvekse kegler uden linier.

Sætning 12.6. *Lad $\mathcal{K} \subseteq \mathbb{R}^n$ være en afsluttet konveks kegle som ikke er $\{0\}$. Hvis $\mathcal{K}$ ikke indeholder linier, er $\mathcal{K}$ den konvekse kegle frembragt af sine ekstremale retninger.*

Bevis. Beviset er et induktionsbevis efter dimensionen k af $\mathcal{K}$. Tilfældet $k = 0$ er ikke aktuelt da det er forudsat at $\mathcal{K} \neq \{0\}$. Induktionen starter derfor med $k = 1$ som behandles for sig, og induktionstrinnet tager så sit udgangspunkt i at nedenstående udsagn $\mathcal{U}_k$ er sandt for $k = 1$.

$\mathcal{U}_k$: Enhver afsluttet konveks kegle $\mathcal{K}$ som ikke er $\{0\}$-keglen, som er uden linier og af dimension mindre eller lig k indeholdt i et talrum $\mathbb{R}^\ell$, er den konvekse kegle frembragt af sine ekstremale retninger.

For $k = 1$ er $\mathcal{K} = \{\lambda b | \lambda \geq 0\}$ og da $\mathcal{K}$ ikke indeholder linier er $\mathcal{K}$ i sig selv lig den eneste ekstremale retning som denne kegle indeholder. Lad nu $k \geq 1$, $k \in \mathbb{N}$ og antag $\mathcal{U}_k$ er sand. Lad da $\mathcal{K} \subseteq \mathbb{R}^n$ være en afsluttet konveks kegle uden linier og af dimension

$k+1$. Lad $a \in \mathcal{K} \setminus \{0\}$. Da $\dim(\mathcal{K}) \geq k+1 \geq 2$ findes $b \in \mathcal{K} \setminus \{0\}$, så $a \notin \{tb | t \geq 0\}$. For en kegle $\mathcal{K}$ gælder åbenbart $\mathcal{K} \subseteq \text{recc}(\mathcal{K})$, da $\mathcal{K} + \mathcal{K} \subseteq \mathcal{K}$. På den anden side fås af recessionskeglens egenskaber at $\mathcal{K} = \mathcal{K} + \text{recc}(\mathcal{K}) \supset \text{recc}(\mathcal{K})$ da $0 \in \mathcal{K}$, så $\text{recc}(\mathcal{K}) = \mathcal{K}$. Specielt er $b \in \text{recc}(\mathcal{K}) \setminus \{0\}$. Vi kan ikke have $-b \in \text{recc}(\mathcal{K})$, da $b \in \text{recc}(\mathcal{K}) \cap (-\text{recc}(\mathcal{K}))$, og $b \neq 0$ ville betyde, at linien $\{tb | t \in \mathbb{R}\}$ er en delmængde af $\mathcal{K}$. Da $\mathcal{K}$ ingen linier har, er $-b \notin \text{recc}(\mathcal{K})$, så Sætning 10.6 giver

$$\{t \in \mathbb{R} | a + tb \in \mathcal{K}\} = [\alpha, \infty[\text{ for et } \alpha \leq 0 \,.$$

Bemærk specielt at $a + \alpha b \notin \text{ri}\,(\mathcal{K})$, idet $a + \alpha b$ må være et randpunkt for for $\mathcal{K}$ relativt til $\text{aff}(\mathcal{K})$ Defineres $d_0 := a + \alpha b$ så er $d_0 \in \mathcal{K} \setminus \text{ri}\,(\mathcal{K})$ og i henhold til Sætning 11.2 kan d_0 og $\text{ri}\,(\mathcal{K})$ separeres egentligt. Der findes derfor en en hyperplan $H(y, \alpha)$ så

$$\begin{array}{ll} \forall k \in \text{ri}\,(\mathcal{K}) : & y \cdot k \leq \alpha \text{ og } y \cdot d_0 \geq \alpha \\ \exists c_0 \in \text{ri}\,(\mathcal{K}) : & y \cdot c_0 < \alpha. \end{array}$$

Eftersom $\mathcal{K} = \overline{\text{ri}\,(\mathcal{K})}$ og $d_0 \in \mathcal{K}$ fås $y \cdot d_0 = \alpha$. Nu er $\mathcal{K}$ en kegle så for $n \in \mathbb{N}$ er $nd_0 \in \mathcal{K}$ og $n\alpha = n(d_0 \cdot y) = nd_0 \cdot y \leq \alpha$, hvoraf vi kan slutte at $\alpha \leq 0$. På den anden side er $0 \in \mathcal{K}$ så $0 \cdot y \leq \alpha$ og $\alpha = 0$. Idet $\alpha = 0$ fandt vi ovenfor et $c_0 \in \mathcal{K} \setminus H(y, 0)$ dvs. $y \cdot c_0 < 0$. Heraf kan vi slutte at $\mathcal{K} \cap H(y, 0)$ er en ikke–tom afsluttet konveks kegle uden linier af dimension mindst 1 da $d_0 \neq 0$, og af dimension højst k, da span $(\mathcal{K} \cap H(y, 0))$ ikke kan indeholde c_0, eftersom $c_0 \notin H(y, 0)$. Induktionsforudsætningen giver da, at $\mathcal{K} \cap H(y, 0)$ er den konvekse kegle frembragt af sine ekstremale retninger. Lad $h := \{\lambda d | \lambda \geq 0\}$ være en ikke–triviel ekstremal retning i $\mathcal{K} \cap H(y, 0)$, da bliver h også en ekstremal retning i $\mathcal{K}$. Antages nemlig $d = c_1 + c_2$ med $c_1, c_2 \in \mathcal{K}$, så fås $y \cdot c_1 \leq 0$, $y \cdot c_2 \leq 0$, og $0 = y \cdot d = y \cdot c_1 + y \cdot c_2$, så $y \cdot c_1 = 0$ og $y \cdot c_2 = 0$, dvs. $c_1, c_2 \in \mathcal{K} \cap H(y, 0)$, og dermed, da h er ekstremal i $\mathcal{K} \cap H(y, 0)$, at $c_1 \in h$ og $c_2 \in h$. Heraf ses at $\mathcal{K}$ har ekstremale retninger.

Defineres nu nu $\mathcal{H} := \text{cone}\,(\text{exray}(\mathcal{C}))$ så kan et indirekte argument vise os at $\mathcal{H} = \mathcal{K}$. Det antages derfor at der findes $a \in \mathcal{K} \setminus \mathcal{H}$, altså specielt $a \neq 0$. De allerede gennemførte argumenter viser at $\mathcal{H} \neq 0$ så ifølge ovenstående, findes $b \in \mathcal{H} \setminus \{0\}$. Processen fra før med a og b gentages nu, men med disse nye elementer a og b. Der findes følgelig et interval $[\alpha, \infty[$ med $\alpha \leq 0$, om hvilket det gælder:

$$a + \lambda b \in \mathcal{K} \;\text{ for }\; \lambda \geq \alpha \;\text{ og }\; a + \lambda b \notin \mathcal{K} \;\text{ for }\; \lambda < \alpha \,.$$

Lad igen $d_0 = a + \alpha b$. Vi ser, at $d_0 \notin \text{ri}\,(\mathcal{K})$ og d_0 kan skilles egentligt fra $\text{ri}\,(\mathcal{K})$, og som før fås at $d_0 \in \text{cone}\,(\text{exray}(\mathcal{C}))$. Nu er $a = d_0 - \alpha b$, og da $\alpha \leq 0$, er $-\alpha b \in \mathcal{H} = \text{cone}\,(\text{exray}(\mathcal{C}))$ og det ses at $a \in \text{cone}\,(\text{exray}(\mathcal{C}))$ i modstrid med antagelsen. Sætningen følger. $\square$

Vi er nu gennem de fleste anvendelser af separationssætningen, men vi mangler en enkelt, hvis bevis dog har mange lighedspunkter med beviset for Sætning 12.6.

Sætning 12.7. *Lad $\mathcal{C}$ være en afsluttet konveks mængde i $\mathbb{R}^n$. Hvis $\mathcal{C}$ ikke indeholder linier er $\mathcal{C} = \text{conv}\,(\text{ext}(\mathcal{C})) + \text{recc}(\mathcal{C})$.*

Bevis. Igen et induktionsbevis efter dimensionen af $\mathcal{C}$. Dimensionerne -1 og $\{0\}$ svarende til henholdsvis $\mathcal{C} = \emptyset$ og $\mathcal{C} = \{\text{et punkt}\}$ er vel problemfrie? Ligeledes giver Sætning 12.6 sammen med Sætning 7.1 resultatet, når $\mathcal{C}$ er begrænset. Vi kan derfor antage at $\mathcal{C}$ er ubegrænset, og dermed $\operatorname{recc}(\mathcal{C}) \neq 0$ i det følgende. Induktionsudsagnet $\mathcal{U}_k$ er som følger $(k \geq 1)$.

$\mathcal{U}_k$; Enhver ubegrænset, afsluttet, konveks delmængde $\mathcal{C}$ uden linier og af dimension mindre eller lig k i et talrum $\mathbb{R}^\ell$, tilfredsstiller identiteten $\mathcal{C} = \operatorname{conv}(\operatorname{ext}(\mathcal{C})) + \operatorname{recc}(\mathcal{C})$.

Først et bevis for $\mathcal{U}_1$: En konveks delmængde af $\mathbb{R}^n$ af dimension 1 er en konveks delmængde af en linie $\ell = \{a + tb | t \in \mathbb{R}\}$. Da $\mathcal{C}$ er afsluttet, konveks, og uden linier, må $\mathcal{C}$ være givet ved $t \in [\alpha, \infty[$ eller $t \in]-\infty, \beta]$. Lad os antage det første, så er $\operatorname{ext}(\mathcal{C}) = a + \alpha b$ og $\operatorname{recc}(\mathcal{C}) = \{\lambda b | \lambda \geq 0\}$, så $\mathcal{U}_1$ følger.

Lad nu $k \geq 1$, og lad $\mathcal{C}$ være en ubegrænset, afsluttet, konveks delmængde af $\mathbb{R}^n$ af dimension $k+1$ som ikke indeholder linier, og lad $c_0 \in \mathcal{C}$ være tilfældigt valgt. Det skal nu vises at $c_0 \in \operatorname{conv}(\operatorname{ext}(\mathcal{C})) + \operatorname{recc}(\mathcal{C})$. Da $\mathcal{C}$ er ubegrænset er $\operatorname{recc}(\mathcal{C}) \neq 0$, så der findes $b \in \operatorname{recc}(\mathcal{C}) \setminus \{0\}$. Da $\mathcal{C}$ ikke indeholder linier, er $-b \notin \operatorname{recc}(\mathcal{C})$, og ifølge Sætning 10.6 findes $\alpha \leq 0$, så $\{\lambda \in \mathbb{R} | c_0 + \lambda b \in \mathcal{C}\} = [\alpha, \infty[$. Lad $d_0 = c_0 + \alpha b$, så er $d_0 \notin \operatorname{ri}(\mathcal{C})$, og d_0 og $\operatorname{ri}(\mathcal{C})$ kan separeres egentligt ved hjælp af Sætning 11.2. Lad da $y \in \mathbb{R}^n \setminus \{0\}$ og $\alpha \in \mathbb{R}$ være fundne så $H(y, \alpha)$ separerer egentligt, dvs.

$$\forall c \in \operatorname{ri}(\mathcal{C}) : y \cdot c \leq \alpha \quad \text{og} \quad y \cdot d_0 \geq \alpha\,.$$

Da $\mathcal{C} = \overline{\operatorname{ri}(\mathcal{C})}$ er $y \cdot c \leq \alpha$ for alle $c \in \mathcal{C}$ så $y \cdot d_0 = \alpha$, og da separationen er egentlig, må der findes $c_0 \in \mathcal{C}$ så $y \cdot c_0 < \alpha$, og dermed for $\mathcal{D} = \{c \in \mathcal{C} | y \cdot c = \alpha\}$ gælder $\mathcal{D} \subsetneq \mathcal{C}$, $\mathcal{D} \neq \emptyset$. Da det for alle $d \in \mathcal{D}$ gælder at $y \cdot d = \alpha$, fås $\forall a \in \operatorname{aff}\mathcal{D} : y \cdot a = \alpha$, så $\operatorname{aff}\mathcal{D} \neq \operatorname{aff}(\mathcal{C})$, da der findes $c \in \mathcal{C}$ med $y \cdot c < \alpha$, dvs. $\dim(\mathcal{D}) \leq k$. Induktionsantagelsen giver nu at $\mathcal{D} = \operatorname{conv}(\operatorname{ext}(\mathcal{D})) + \operatorname{recc}(\mathcal{D})$. Som i beviset for Sætning 12.5, ses det, at $\operatorname{ext}(\mathcal{D}) \subseteq \operatorname{ext}(\mathcal{C})$. Det følger endvidere fra Sætning 10.6 (ii) at $\operatorname{recc}(\mathcal{D}) \subseteq \operatorname{recc}(\mathcal{C})$. Lad nemlig $e \in \operatorname{recc}(\mathcal{D})$, da gælder for d_0 at $\{d_0 + \lambda e | \lambda \geq 0\} \subseteq \mathcal{D} \subseteq \mathcal{C}$. Men $d_0 \in \mathcal{C}$. og derfor fås $e \in \operatorname{recc}(\mathcal{C})$ fra 10.6 (ii). Da $d_0 \in \mathcal{D}$ gælder altså

$$\begin{aligned} d_0 \in \mathcal{D} &= \operatorname{conv}(\operatorname{ext}(\mathcal{D})) + \operatorname{recc}(\mathcal{D}) \\ &\subseteq \operatorname{conv}(\operatorname{ext}(\mathcal{C})) + \operatorname{recc}(\mathcal{C})\,. \end{aligned}$$

Imidlertid er $c_0 = d_0 - \alpha b$ og $\alpha \leq 0$ og $b \in \operatorname{recc}(\mathcal{C})$, så $c_0 \in \operatorname{conv}(\operatorname{ext}(\mathcal{C})) + \operatorname{recc}(\mathcal{C}) + \operatorname{recc}(\mathcal{C})$, og da $\operatorname{recc}(\mathcal{C}) + \operatorname{recc}(\mathcal{C}) = \operatorname{recc}(\mathcal{C})$ eftersom $\operatorname{recc}(\mathcal{C})$ er en konveks kegle, fås $c_0 \in \operatorname{conv}(\operatorname{ext}(\mathcal{C})) + \operatorname{recc}(\mathcal{C})$ og sætningen følger, da c_0 var tilfældigt valgt i $\mathcal{C}$. ☐

Festens clou er nu sammenstykningen af disse resultater til følgende:

Sætning 12.8. *Lad $\mathcal{C}$ være en afsluttet konveks mængde indeholdt i $\mathbb{R}^n$. Hvis $\mathcal{C}$ ikke indeholder linier er*

$$\begin{aligned} \mathcal{C} &= \operatorname{conv}(\operatorname{ext}(\mathcal{C})) + \operatorname{cone}(\operatorname{exray}(\mathcal{C})) \\ &= \operatorname{conv}(\operatorname{ext}(\mathcal{C}) + \operatorname{exray}(\mathcal{C}))\,. \end{aligned}$$

Bevis. Vi bemærker først at for to delmængder $\mathcal{A}$ og $\mathcal{B}$ i $\mathbb{R}^n$ gælder $\operatorname{conv}(\mathcal{A}+\mathcal{B}) = \operatorname{conv}(\mathcal{A}) + \operatorname{conv}(\mathcal{B})$. Beviset for dette overlades til læseren ligsom det også overlades til samme person at vise at for en konveks mængde $\mathcal{C}$ er $\operatorname{exray}(\mathcal{C}) = \operatorname{exray}(\operatorname{recc}(\mathcal{C}))$ og $\operatorname{conv}(\operatorname{exray}(\operatorname{recc}(\mathcal{C}))) = \operatorname{cone}(\operatorname{exray}(\operatorname{recc}(\mathcal{C})))$. Det bemærkes at vi iøvrigt tidligere har overvejet at $\operatorname{recc}(\mathcal{C})$ er afsluttet når $\mathcal{C}$ er det. Vi vil herefter gennemføre beviset for sætningen ved at argumenterer for en række identiteter:

$$\begin{aligned}
\operatorname{conv}(\operatorname{ext}(\mathcal{C}) + \operatorname{exray}(\mathcal{C})) &= \operatorname{conv}(\operatorname{ext}(\mathcal{C})) + \operatorname{conv}(\operatorname{exray}(\mathcal{C})) \quad \text{iht. første påstand ovenfor}\\
&= \operatorname{conv}(\operatorname{ext}(\mathcal{C})) + \operatorname{conv}(\operatorname{exray}((\operatorname{recc}(\mathcal{C}))) \quad \text{iht. anden påstand}\\
&= \operatorname{conv}(\operatorname{ext}(\mathcal{C})) + \operatorname{cone}(\operatorname{exray}((\operatorname{recc}(\mathcal{C}))) \quad \text{iht. tredie påstand}\\
&= \operatorname{conv}(\operatorname{ext}(\mathcal{C})) + \operatorname{recc}(\mathcal{C}) \quad \text{iht. Sætning 12.6}\\
&= \mathcal{C} \quad \text{iht. Sætning 12.7}
\end{aligned}$$

□

Den sidste opgave vi står overfor bliver at beskrive konvekse polyedre.

Definition 12.9. En *konveks polytop* i $\mathbb{R}^n$ er det konvekse hylster af endeligt mange punkter i $\mathbb{R}^n$.

Et *konvekst polyeder* i $\mathbb{R}^n$ er fællesmængden af endeligt mange afsluttede halvrum i $\mathbb{R}^n$.

Eksempel 12.10. Lad $c \in \mathbb{R}^n$, $A \in \mathbf{M}(m,n)$ og $b \in \mathbb{R}^m$, $x \in \mathbb{R}^n$. Et lineært programmeringsproblem består i at bestemme $\sup\limits_{x\in\mathbb{R}^n} c \cdot x$ under bibetingelserne

$$\begin{aligned}
&{}_1A \cdot x \le b_1\,;\ {}_2A \cdot x \le b_2, \ldots, {}_mA \cdot x \le b_m\\
&x_1 \ge 0, \ldots, x_n \ge 0\,.
\end{aligned}$$

Man kan tænke på ulighederne ${}_iA \cdot x \le b_i$ som nogle begrænsninger af økonomisk eller teknisk art, som problemet lægger på valget af de mulige x–vektorer. Betingelserne $x_i \ge 0$ vil ofte være naturlige fra problemstillingen, hvis f.eks. x_i betegner forbruget af en vare eller et gode. Vender man sig til en mere formel matematisk opfattelse af problemet ses det, at mængden af "mulige x–er" $\mathcal{M}$ som kriteriefunktionen $x \to c \cdot x$ skal maksimeres over udgør et polyeder, idet $\mathcal{M}$ er fællesmængden af $(n+m)$ halvrum

$$\mathcal{M} = \left(\bigcap_{i=1}^{m} J({}_iA, b_i)\right) \bigcap \left(\bigcap_{j=1}^{n} J(-e_j, 0)\right)\,.$$

Vi anvender nu den lærte teori på $\mathcal{M}$. Først bemærkes det at $\mathcal{M}$ muligvis er tom, i hvilket tilfælde der gælder $\sup\{c \cdot x \mid x \in \mathcal{M}\} = -\infty$. Vi vil fremover antage at $\mathcal{M}$ ikke er tom og så starte med at bestemmes $\operatorname{recc}(\mathcal{M})$ samt det største underrum $\mathcal{L} \subseteq \operatorname{recc}(\mathcal{M})$. I dette tilfælde er $\mathcal{L} = 0$ eftersom alle koordinater er ikke negative. Vi har så at $\mathcal{M}$ er

et polyeder der er en fællesmængde af $n+m$ afsluttede halvrum, og det følger at $\mathcal{M}$ er en afsluttet konveks mængde uden linier.
Ifølge Sætning 12.8 er

$$\mathcal{M} = \operatorname{conv}(\operatorname{ext}(\mathcal{M})) + \operatorname{cone}(\operatorname{exray}(\mathcal{M})).$$

Da $\mathcal{M} \neq \emptyset$ er $\sup\{c \cdot m | m \in \mathcal{M}\}$ enten endelig eller uendelig. Det viser sig at det der afgør hvilket tilfælde der forekommer er om objektfunktionens koefficient-vektor c tilhørere normalkeglen $\operatorname{normc}(\operatorname{recc}(\mathcal{M}))$ eller ej.

1) Antag $c \notin \operatorname{normc}(\operatorname{recc}(\mathcal{M}))$. I dette tilfælde vil der findes en vektor $b \in \operatorname{recc}(\mathcal{M})$ med $c \cdot b > 0$, og for $m \in \mathcal{M}$, $n \in \mathbb{N}$ er $m_n := m + nb \in \mathcal{M}$ og det ses at $c \cdot m_n \to +\infty$ for $n \to \infty$. Altså

$$c \notin \operatorname{normc}(\operatorname{recc}(\mathcal{M}))) \implies \sup\{c \cdot m | m \in \mathcal{M}\} = +\infty\,.$$

2) Vi antager nu at $c \in \operatorname{normc}(\operatorname{recc}(\mathcal{M}))$. Da $\mathcal{M} = \operatorname{conv}(\operatorname{ext}(\mathcal{M})) + \operatorname{recc}(\mathcal{M})$ ses det, at for $m = x + b$ med $x \in \operatorname{conv}(\operatorname{ext}(\mathcal{M}))$ og $b \in \operatorname{recc}(\mathcal{M})$, er $c \cdot m = c \cdot x + c \cdot b \leq c \cdot x$, så da $0 \in \operatorname{recc}(\mathcal{M})$ fås $\sup\{c \cdot m | m \in \mathcal{M}\} = \sup\{c \cdot x | x \in \operatorname{conv}(\operatorname{ext}(\mathcal{M}))\} = \sup\{c \cdot x | x \in \operatorname{ext}(\mathcal{M})\}$, da funktionen $x \to c \cdot x$ er lineær.

Det vil sige, det er nok at gøre følgende for at løse problemet.

1) Afgør om $\mathcal{M} = \emptyset$.

2) Hvis $\mathcal{M} \neq \emptyset$ skal det afgøres om $c \in \operatorname{normc}(\operatorname{recc}(\mathcal{M}))$.

3) Hvis $\mathcal{M} \neq \emptyset$ og $c \in \operatorname{normc}(\operatorname{recc}(\mathcal{M}))$ bestemmes $\operatorname{ext}(\mathcal{M})$, og værdimængden $\{c \cdot x | x \in \operatorname{ext}(\mathcal{M})\}$.

Det fremgår af nedenstående mere detaljerede undersøgelse af polyedre, at hvis et sådant er fællesmængde af k halvrum i $\mathbb{R}^n$, så har det højst $\binom{k}{n}$ ekstremalpunkter. Maksimeringsopgaven i ovenstående lineære program er derfor endelig og pricipielt løsbar ved at bestemme objektfunktionens værdi i alle ekstremalpunkterne. Det næste afsnit viser hvorledes man konkret finder ekstremalpunkterne ud fra de givne hyperplaner. Disse overvejelser giver ikke en hensigtsmæssig algoritme til løsning af praktiske problemer. I praksis kan man ved hjælp af den såkaldte simplexalgoritme succesivt bestemme ekstremalpunkter således at objektfunktionens værdi vokser (svagt) ved hver itteration. Vi skal ikke gå nærmere ind på disse anvendelser af teorien her, men vil dog gøre opmærksom på et vigtigt teoretisk resultat som vi allerede nu kan indse: *Der er kun 3 muligheder for løsninger til et lineært programmeringsproblem:*

(1) $\mathcal{M} = \emptyset$ og $\sup\{c \cdot x | x \in \mathcal{M}\} = -\infty$

(2) $\mathcal{M} \neq \emptyset$ og $\sup\{c \cdot x | x \in \mathcal{M}\} = \infty$

(3) $\mathcal{M} \neq \emptyset$ og $\sup\{c \cdot x | x \in \mathcal{M}\} = \max\{c \cdot x | x \in \operatorname{ext}(\mathcal{M})\}$.

12.1 Detaljer om polyedre

Lad $y_1, \ldots, y_m \in \mathbb{R}^n \setminus \{0\}$ og $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ og

$$\mathcal{P} := \bigcap_{i=1}^{m} J(y_i, \alpha_i)\,.$$

I det følgende antages, at $\mathcal{P}$ er et ikke er tomt polyeder. På de næste sider gives en opskrift på hvorledes $\mathrm{recc}(\mathcal{P})$, $\mathcal{L} := \mathrm{recc}(\mathcal{P}) \cap (-\mathrm{recc}(\mathcal{P}))$, ekstremalpunkter og ekstremale retninger, bestemmes (såfremt disse mængder ikke er tomme eller udartede).

1) $\mathrm{recc}(\mathcal{P}) = \bigcap\limits_{i=1}^{m} J(y_i, 0)$.

Bevis. "$\mathrm{recc}(\mathcal{P}) \subseteq \bigcap\limits_{i=1}^{m} J(y_i, 0)$".

Lad $b \in \mathrm{recc}(\mathcal{P})$ og $p \in \mathcal{P}$, da er $p + \lambda b \in \mathcal{P}$ for $\lambda \geq 0$, så for $i \in \{1, \ldots, m\}$: $y_i \cdot (p + \lambda b) \leq \alpha_i$. Lad $\lambda \to +\infty$, da fås $y_i \cdot b \leq \frac{1}{\lambda}(\alpha_i - y_i \cdot p)$, så $y_i \cdot b \leq 0$, og $b \in \bigcap\limits_{i=1}^{m} J(y_i, 0)$.

Bevis. "$\bigcap\limits_{i=1}^{m} J(y_i, 0) \subseteq \mathrm{recc}(\mathcal{P})$".

Lad $p \in \mathcal{P}$ og $b \in J(y_i, 0)$ for $1 \leq i \leq m$, så er

$$(p + b) \cdot y_i \leq \alpha_i + 0 = \alpha_i \quad \text{dvs.} \quad p + b \in \mathcal{P} \quad \text{så} \quad b \in \mathrm{recc}(\mathcal{P})\,.$$

2) $\mathcal{L} = \mathrm{recc}(\mathcal{P}) \cap (-\mathrm{recc}(\mathcal{P})) = \bigcap\limits_{i=1}^{m} H(y_i, 0) = (\{y_1, \ldots, y_m\})^{\perp}$.

Bevis. Følger fra Sætning 7.3 og 1).

Analysen burde nu fortsætte med $\mathcal{P} \cap \mathcal{L}^{\perp}$ i $\mathcal{L}^{\perp}$, men for at forenkle argumenterne antages det i stedet at $\mathcal{P}$ er et polyeder for hvilket $\mathcal{L} = 0$, altså et polyeder uden linier. Problemet kunne også klares ved at vælge en ortonormalbasis $(z_1, \ldots, z_d)$ for $\mathcal{L}$, og så tilføje halvrummene $J(z_i, 0)$, $J(-z_i, 0)$ til listen over hvilke der dannes fællesmængde. Herved opnås at vi får beskrevet $\mathcal{P} \cap \mathcal{L}^{\perp}$ som et polyeder. Notationsmæssigt er det imidlertid nemmest at forudsætte, at $\mathcal{P}$ ikke har linier i de næste beregninger.

$\mathcal{P}$ antages ikke at indeholde linier. Fra 2) ses dette at være ensbetydende med, at $(\{y_1, \ldots, y_m\})^{\perp} = \{0\}$, men det giver så, at

$$\mathbb{R}^n = ((\{y_1, \ldots, y_m\})^{\perp})^{\perp} = \mathrm{span}\,(\{y_1, \ldots, y_m\})\,.$$

Altså sættet $\{y_1, \ldots, y_m\}$ udspænder $\mathbb{R}^n$, og må derfor indeholde mindst et sæt $(y_{i_1}, y_{i_2}, \ldots, y_{i_n})$ bestående af n lineært uafhængige vektorer. Et sådant sæt kaldes et *basissæt* fra $\{y_1, \ldots, y_m\}$. Mængden af samtlige basissæt betegnes $\mathcal{B}$. Nu kan $\mathrm{ext}(\mathcal{P})$ bestemmes:

3) $x \in \operatorname{ext}(\mathcal{P}) \iff (x \in \mathcal{P} \text{ og } \exists (y_{i_1}, \ldots, y_{i_n}) \in \mathcal{B} : x \cdot y_{i_1} = \alpha_{i_1}, \ldots, x \cdot y_{i_n} = \alpha_{i_n})$.

Bevis. Beviset føres ved at vise, at højresiden medfører venstresiden, og at vise, at negationen af højresiden medfører negationen af venstresiden. Bemærk, at Sætning 12.7 viser, at når $\mathcal{P} \neq \emptyset$ som antaget, er $\operatorname{ext}(\mathcal{P}) \neq \emptyset$. Antag nu, at $x \in \mathcal{P}$, og der findes $(y_{i_1}, \ldots, y_{i_n}) \in \mathcal{B}$ så $x \cdot y_{i_j} = \alpha_{i_j}$, $1 \leq j \leq n$. Lad $u, z \in \mathcal{P}$, $0 < \lambda < 1$ og antag, at $x = (1-\lambda)u + \lambda z$. Hvis dette medfører at $x = u = z$, er x ekstremal i $\mathcal{P}$. Da $u, z \in \mathcal{P}$, er $u \cdot y_{i_j} \leq \alpha_{i_j}$ og $z \cdot y_{i_j} \leq \alpha_{i_j}$, $1 \leq j \leq n$. Men $\alpha_{i_j} = x \cdot y_{i_j} = (1-\lambda)u \cdot y_{i_j} + \lambda z \cdot y_{i_j} \leq \alpha_{i_j}$, så $u \cdot y_{i_j} = \alpha_{i_j}$, og $z \cdot y_{i_j} = \alpha_{i_j}$, eftersom $0 < \lambda < 1$ for $1 \leq j \leq n$. Det er forudsat at sættet $(y_{i_1}, \ldots, y_{i_n})$ er et basissæt, altså lineært uafhængigt, derfor vil identiteterne $u \cdot y_{i_j} = z \cdot y_{i_j} = x \cdot y_{i_j}$ for $1 \leq j \leq n$ medføre at $u = z = x$. (Hvis du tvivler kan du se, at $(x-u)$ og $(x-z)$ begge tilhører $\{y_{i_1}, \ldots, y_{i_n}\}^{\perp} = (\operatorname{span}(y_{i_1}, \ldots, y_{i_n}))^{\perp} = (\mathbb{R}^n)^{\perp} = 0$). Heraf ses så, at x er et ekstremalpunkt i $\mathcal{P}$. Lad nu $x \in \mathcal{P}$ og $I = \{i \in \{1, \ldots, m\} | y_i \cdot x = \alpha_i\}$.

Det skal nu vises at hvis mængden $\{y_i | i \in I\}$ ikke indeholder et basissæt, så er x ikke et ekstremalpunkt. Vi antager derfor, at $\{y_i | i \in I\}$ ikke indeholder et basissæt, dvs. $\operatorname{span}(\{y_i | i \in I\}) \neq \mathbb{R}^n$. Lad da $z \in \mathbb{R}^n \setminus \{0\}$ være en vektor i $(\{y_i | i \in I\})^{\perp}$, dvs.

$$z \cdot y_i = 0 \text{ hvis } i \in I\ .$$

Da $\operatorname{span}\{y_1, \ldots, y_m\} = \mathbb{R}^n$, kan I ikke være hele mængden $\{1, \ldots, m\}$, så der findes $i \in \{1, \ldots, m\} \setminus I$. Lad nu $\lambda \in \mathbb{R}$, og definer $u_\lambda = x + \lambda z$, da ses

$$\begin{aligned} &\forall i \in I : y_i \cdot u_\lambda = y_i \cdot (x + \lambda z) = y_i \cdot x = \alpha_i \\ &\forall i \in \{1, \ldots, m\} \setminus I : \alpha_i - y_i \cdot u_\lambda = (\alpha_i - y_i \cdot x) - \lambda (y_i \cdot z)\,. \end{aligned}$$

Da det for alle $i \in \{1, \ldots, m\} \setminus I$ gælder at $(\alpha_i - y_i \cdot x) > 0$ ses det, at for alle sådanne i findes et interval $[-\delta_i, \delta_i]$ for et $\delta_i > 0$, så $\lambda \in [-\delta_i, \delta_i] \implies (\alpha_i - y_i \cdot x) - \lambda (y_i \cdot z) \geq 0$. Sæt $\delta = \min\{\delta_1, \ldots, \delta_m\} > 0$, da fås

$$\forall i \in \{1, \ldots, m\} \setminus I : y_i \cdot x + \delta y_i \cdot z \leq \alpha_i \text{ og } y_i \cdot x - \delta y_i \cdot z \leq \alpha_i\,,$$

dvs. $u_\delta = x + \delta z \in \mathcal{P}$ og $u_{-\delta} = x - \delta z \in \mathcal{P}$, og $x = \frac{1}{2} u_\delta + \frac{1}{2} u_{-\delta}$. Da $z \neq 0$ og $\delta > 0$, er $u_\delta \neq x$ og $u_{-\delta} \neq x$, og det følger, at x ikke er et ekstremalpunkt i $\mathcal{P}$. Hermed er beviset for 3) fuldført.

4) Lad $b \in \mathbb{R}^n \setminus \{0\}$ og $r = \{\lambda b | \lambda \geq 0\}$ være retningen i $\mathbb{R}^n$ med retningsvektor b.

Retningen r er ekstremal for $\mathcal{P} \iff \forall i \in \{1, \ldots, m\} : b \cdot y_i \leq 0$, og der findes $(n-1)$ lineært uafhængige vektorer $\{y_{i_1}, \ldots, y_{i_{n-1}}\} \subseteq \{y_i | 1 \leq i \leq m\}$, så $b \cdot y_{i_1} = b \cdot y_{i_2} = \cdots = b \cdot y_{i_{n-1}} = 0$.

Bevis. Beviset er analogt til beviset for 3), idet det først vises, at betingelsen til højre er tilstrækkelig til at sikre, at r er ekstremal. Dernæst vises det, at betingelsen er nødvendig ved at vise, at enhver retning r som ikke opfylder betingelsen, heller ikke er ekstremal. Antag nu, at betingelsen er opfyldt for $r = \{\lambda b | \lambda \geq 0\}$, dvs. $b \cdot y_i \leq 0$, og $b \cdot y_{i_1} = \cdots =$

$b \cdot y_{i_{n-1}} = 0$ for et lineært uafhængigt sæt $\{y_{i_1}, \dots, y_{i_{n-1}}\}$. Antag endvidere at $b = c + d$ med $c, d \in \mathrm{recc}(\mathcal{P})$, jeg skal så vise, at der findes ikke–negative skalarer α og β, så $c = \alpha b$ og $d = \beta b$. Da $c, d \in \mathrm{recc}(\mathcal{P})$ fås fra 1), at for $1 \le i \le m$ er $c \cdot y_i \le 0$ og $d \cdot y_i \le 0$, men for $j \in \{1, 2, \dots, n-1\}$ er $y_{i_j} \cdot (c + d) = y_{i_j} \cdot b = 0$, så $y_{i_j} \cdot c = y_{i_j} \cdot d = 0$ for $1 \le j \le n-1$. Da sættet $\{y_{i_1}, \dots, y_{i_{n-1}}\}$ består af $(n-1)$ lineært uafhængige vektorer, og $b \neq 0$ er ortogonal på alle $(n-1)$, må $c, d \in \mathrm{span}\,(b) = (\{y_{i_1}, \dots, y_{i_{n-1}}\})^{\perp}$, altså $c = \alpha b$ og $d = \beta b$ for et par $\alpha, \beta \in \mathbb{R}$. Da $\mathcal{P}$ ikke indeholder linier, kan $\mathrm{recc}(\mathcal{P})$ ikke indeholde linier, så $\alpha \ge 0$ og $\beta \ge 0$ da f.eks. $\alpha < 0$ ville medføre, at $\{\lambda b | \lambda \in \mathbb{R}\} \subseteq \mathrm{recc}(\mathcal{P})$, idet $\mathrm{recc}(\mathcal{P})$ er en konveks kegle. Det er altså vist, at retningen $r = \{\lambda b | \lambda \ge 0\}$ er en ekstremal retning. Lad nu $b \in \mathrm{recc}(\mathcal{P}) \setminus \{0\}$ og antag, at

$$\dim(\mathrm{span}\,(\{y_i | 1 \le i \le m\,,\ y_i \cdot b = 0\})) < (n-1)\,.$$

Vi deler som før indeksmængden $\{1, \dots, m\}$ op m.h.t. om $y_i \cdot b = 0$ eller $y_i \cdot b < 0$, så lad $I := \{i \in \{1, \dots, m\} | y_i \cdot b = 0\}$. Da $\dim(\mathrm{span}(\,\{y_i | i \in I\})) \le n-2$, findes $z \neq 0$ så $z \cdot b = 0$ og $z \cdot y_i = 0$ for $i \in I$. Vi kan så genbruge argumenterne fra beviset under 3) ved at definere $u_\lambda = b + \lambda z$, $\lambda \in \mathbb{R}$. Vi har

$$\begin{aligned} &\forall i \in I : y_i \cdot u_\lambda = y_i \cdot b + \lambda y_i \cdot z = 0 \\ &\forall i \in \{1, \dots, m\} \setminus I : y_i \cdot u_\lambda = y_i \cdot b + \lambda y_i \cdot z < 0 \text{ når } \lambda y_i \cdot z < -y_i \cdot b\,. \end{aligned}$$

Idet det for $i \in \{1, \dots, m\} \setminus I : -y_i \cdot b > 0$, findes for hvert sådant i et $\delta_i > 0$, så $y_i \cdot b + \lambda y_i \cdot z \le 0$ for $\lambda \in [-\delta_i, \delta_i]$. Lad $\delta = \min\{\delta_i | i \in \{1, \dots, m\} \setminus I\}$, så gælder $\forall i \in \{1, \dots, m\}$, $y_i \cdot u_\delta \le 0$, og $y_i \cdot u_{-\delta} \le 0$, dvs. $u_\delta \in \mathrm{recc}(\mathcal{C})$ og $u_{-\delta} \in \mathrm{recc}(\mathcal{C})$, og $b = \frac{1}{2} u_\delta + \frac{1}{2} u_{-\delta}$. Da u_δ og b ikke er proportionale, er b ikke en ekstremal retning.

Eksempel 12.11. Lad $\mathcal{P} = \{(x_1, x_2, x_3) \in \mathbb{R}^3 | x_1 \ge 0\,,\ x_2 \ge 0\,,\ x_3 \ge 0\,,\ x_3 - x_2 \le 1\,,\ 2x_2 - x_1 \ge -1\}$. $\mathcal{P}$ er altså et polyeder $\mathcal{P} = \bigcap_{i=1}^{5} J(y_i, \alpha_i)$, hvor

$$\begin{aligned} &y_1 = (-1, 0, 0)\,,\ \alpha_1 = 0\,;\ y_2 = (0, -1, 0)\,,\ \alpha_2 = 0\,;\ y_3 = (0, 0, -1)\,,\ \alpha_3 = 0\,; \\ &y_4 = (0, -1, 1)\,,\ \alpha_4 = 1\,;\ y_5 = (1, -2, 0)\,,\ \alpha_5 = 1\,. \end{aligned}$$

1)

$$\begin{aligned} \mathrm{recc}(\mathcal{P}) = \{b \in \mathbb{R}^3\ |y_1 \cdot b \le 0\,,\ y_2 \cdot b \le 0\,,\ y_3 \cdot b \le 0 \\ y_4 \cdot b \le 0\,,\ y_5 \cdot b \le 0\}\,. \end{aligned}$$

Da $y_1 \cdot b = -b_1$, $y_2 \cdot b = -b_2$, $y_3 \cdot b = -b_3$ ses, at for $b \in \mathrm{recc}(\mathcal{P})$ er $b_1 \ge 0$, $b_2 \ge 0$, $b_3 \ge 0$, så $\mathrm{recc}(\mathcal{P})$ har ingen linier, og dermed har $\mathcal{P}$ heller ingen linier.

Vi kan derfor gå til 3) og 4) for at bestemme ekstremalpunkter og ekstremale retninger.

3) Proceduren her er arbejdskrævende, men på den anden side let at løse v.h.a. computerpower. I princippet skal vi søge alle lineært uafhængige sæt af 3 vektorer blandt mængden $\{y_1, \dots, y_5\}$. Det er nemmere at udtage alle sæt på 3 vektorer $\{y_{i_1}, y_{i_2}, y_{i_3}\}$ og så prøve at løse ligningssystemet $y_{i_j} \cdot x = \alpha_j$, $1 \le j \le 3$. Såfremt

sættet ikke er lineært uafhængigt er der enten ingen løsning eller uendeligt mange løsninger. Der er $\binom{5}{3} = 10$ sæt bestående af 3 y_i'er, så der er højst 10 ekstremalpunkter. Jeg tager de forskellige kombinationer ved at betegne dem med de indices der indgår, altså $\{1,2,3\}$ betyder, at jeg prøver at løse $y_1 \cdot x = \alpha_1$, $y_2 \cdot x = \alpha_2$, $y_3 \cdot x = \alpha_3$, $\{1,2,3\}$: løsning $(0,0,0)$ og $y_4 \cdot 0 = 0 \leq \alpha_4$ og $y_5 \cdot 0 = 0 \leq 1 = \alpha_5$.

heraf følger at $0 = (0,0,0)$ **er et ekstremalpunkt**. Processen gentages med de andre kombinationer af indices og bliver samlet i nedenstående skema. Når vi finder et ekstremalpunkt bliver det markeret med $\clubsuit$.

$$\clubsuit\{1,2,3\}: \text{ Løsning } (0,0,0)\textbf{er et ekstremalpunkt}.$$

$$\clubsuit\begin{cases}\{1,2,4\}: \text{Løsning } \ (0,0,1)\,,\ y_3 \cdot (0,0,1) = -1 < 0 = \alpha_3 \\ y_5 \cdot (0,0,1) = 0 < 1 = \alpha_5 \text{ dvs. } (0,0,1) \textbf{ er et ekstremalpunkt}.\end{cases}$$

$$\div\{1,2,5\}: \text{ Løsning findes ikke. Sættet } \{y_1, y_2, y_5\} \text{ er lineært afhængigt.}$$

$$\div\{1,3,4\}: \text{ Løsning } (0,-1,0), \text{ men } (0,-1,0) \notin \mathcal{P} \text{ da } y_2 \cdot (0,-1,0) = 1.$$

$$\div\{1,3,5\}: \text{ Løsning } (0,-\frac{1}{2},0), \text{ men } (0,-\frac{1}{2},0) \notin \mathcal{P} \text{ da } y_2 \cdot (0,-\frac{1}{2},0) = \frac{1}{2} > 0\,.$$

$$\div\{1,4,5\}: \text{ Løsning } (0,-\frac{1}{2},\frac{1}{2}) \text{ men } (0,-\frac{1}{2},\frac{1}{2}) \notin \mathcal{P} \text{ da } y_2 \cdot (0,-\frac{1}{2},0) = \frac{1}{2} > 0\,.$$

$$\div\{2,3,4\}: \text{ Løsning findes ikke. Sættet } \{y_2, y_3, y_4\} \text{ er lineært afhængigt.}$$

$$\clubsuit\begin{cases}\{2,3,5\}: \text{Løsning } \ (1,0,0)\,,\ y_1 \cdot (1,0,0) = -1 < 0 = \alpha_1 \\ y_4 \cdot (1,0,0) = 0 < 1 = \alpha_4. \text{ vs. } (1,0,0) \textbf{ er et ekstremalpunkt}.\end{cases}$$

$$\clubsuit\begin{cases}\{2,4,5\}: \text{Løsning } \ (1,0,1)\,;\ y_1 \cdot (1,0,1) = -1 < 0 = \alpha_1 \\ y_3 \cdot (1,0,1) = -1 < 0 = \alpha_3. \text{ dvs. } (1,0,1) \textbf{ er et ekstremalpunkt}.\end{cases}$$

$$\div\{3,4,5\}: \text{ Løsning } (-1,-1,0) \text{ men } y_1 \cdot (-1,-1,0) = 1 > 0 = \alpha_1$$
$$\text{så } (-1,-1,0) \notin \mathcal{P}\,.$$

Der er altså 4 ekstremalpunkter

$$\operatorname{ext}(\mathcal{P}) = \{(0,0,0),(0,0,1),(1,0,0),(1,0,1)\}\,.$$

4) Fastlæggelse af de ekstremale retninger er lige så omfattende et projekt, idet vi skal undersøge alle par $\{y_i, y_j\}$ med $y_i \neq y_j$, finde en normalretning n til disse 2, og så undersøge, om denne ligger i $\operatorname{recc}(\mathcal{P})$. Jeg går frem som før. Et par af tal $\{i,j\}$ betyder, at jeg finder en fælles normalvektor til y_i og y_j, og dernæst undersøger, om denne eller dens modsatte ligger i $\operatorname{recc}(\mathcal{P})$.

$\div \{1,2\} : n = (0,0,1) \quad n \cdot y_3 = -1$ ok. $n \cdot y_4 = 1 > 0$ så
hverken n eller $-n$ tilhører $\text{recc}(\mathcal{P})$.

$$\clubsuit \begin{cases} \{1,3\} : n = (0,1,0) \quad n \cdot y_2 = -2\,,\, n \cdot y_4 = -1\,,\, n \cdot y_5 = -2 \\ \qquad \text{dvs. } (0,1,0) \text{ er retningsvektor for en ekstremalretning} \end{cases}$$

$$\clubsuit \begin{cases} \{1,4\} : n = (0,1,1) \quad n \cdot y_2 = -1\,;\, n \cdot y_3 = -1\,,\, n \cdot y_5 = -2 \\ \qquad \text{dvs. } (0,1,1) \text{ er retningsvektor for en ekstremalretning} \end{cases}$$

$\div \{1,5\} : n = (0,0,1)$. Denne er fundet under $\{1,2\}$

$\div \{2,3\} : n = (1,0,0)$. Da $n \cdot y_1 = -1$ og $n \cdot y_5 = 1$ kan
hverken $n \in \text{recc}(\mathcal{P})$ eller $-n \in \text{recc}(\mathcal{P})$ være muligt

$\div \{2,4\} : n = (1,0,0)$ Denne er fundet under $\{2,3\}$

$\div \{2,5\} : n = (0,0,1)$. Denne er fundet under $\{1,2\}$

$\div \{3,4\} : n = (1,0,0)$ .Denne er fundet under $\{2,3\}$

$$\clubsuit \begin{cases} \{3,5\} : n = (2,1,0) \quad n \cdot y_1 = -2 < 0 \quad n \cdot y_2 = -1 < 0 \\ \qquad n \cdot y_4 = -1 < 0 \text{ dvs. } n \in \text{recc}(\mathcal{P}) \text{ og} \\ \qquad (2,1,0) \text{ er retningsvektor for en ekstremal retning}\,. \end{cases}$$

$$\clubsuit \begin{cases} \{4,5\} : n = (2,1,1) \quad n \cdot y_1 = -2 < 0\,,\, n \cdot y_2 = -1 < 0\,,\, n \cdot y_3 = -1 < 0 \\ \qquad \text{dvs. } (2,1,1) \text{ er retningsvektor for en ekstremal retning}\,. \end{cases}$$

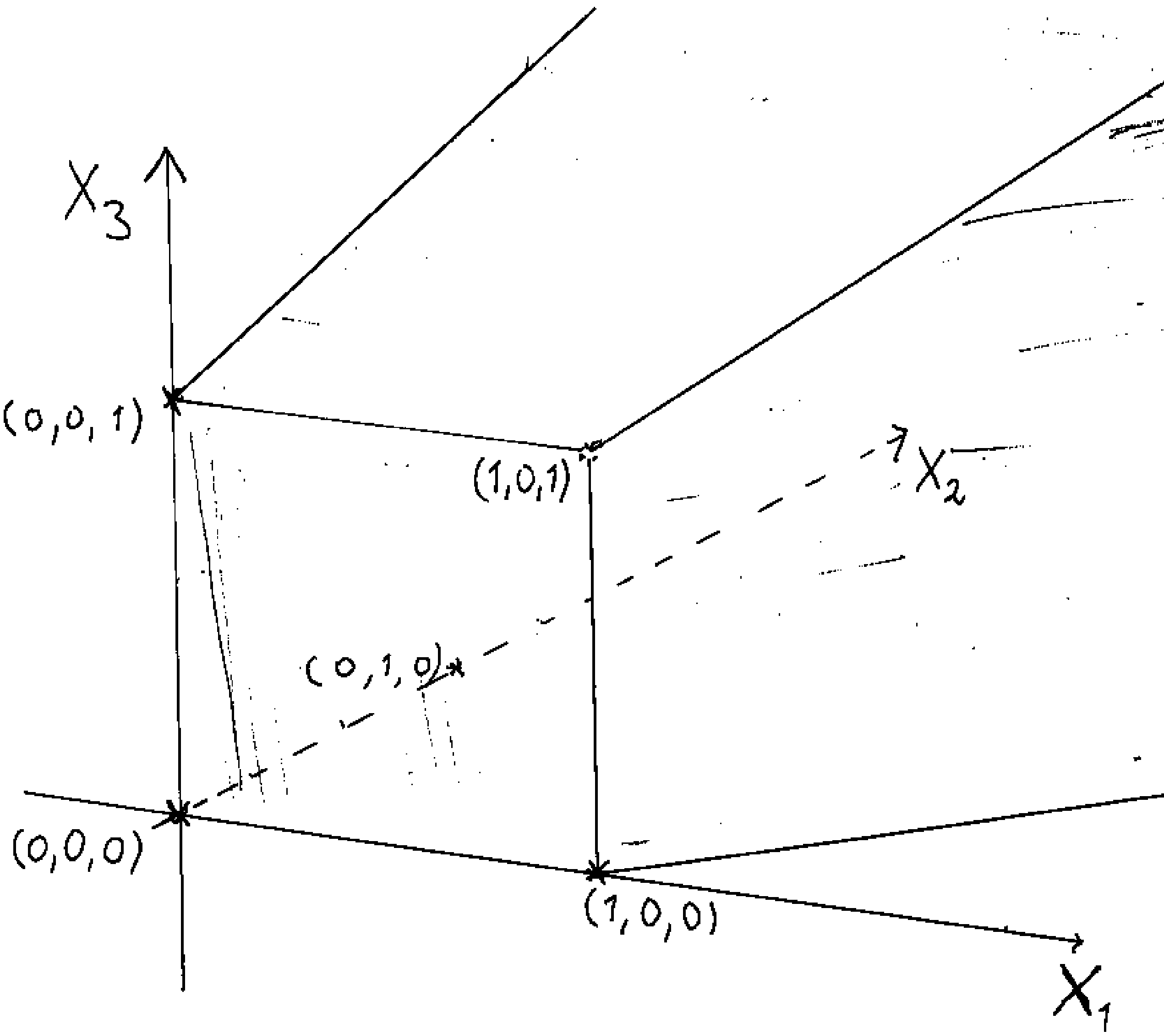
Der er altså 4 ekstremale retninger

$$\{(0,1,0),(0,1,1),(2,1,0),(2,1,1)\}$$

og

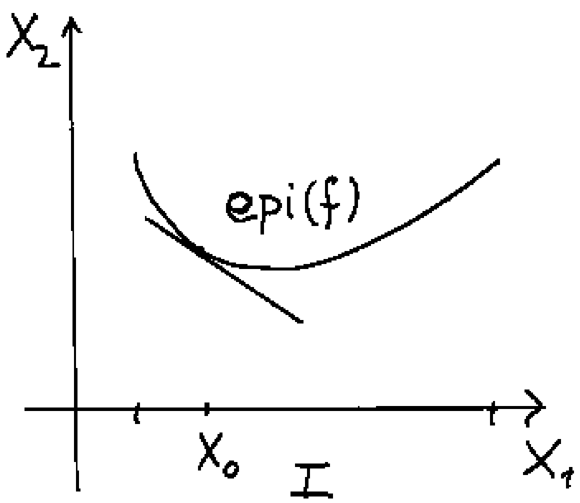
$$\begin{aligned} \mathcal{P} &= \text{conv}\,((0,0,0),(1,0,0),(0,0,1),(1,0,1)) \\ &\quad + \text{cone}\,((0,1,0),(0,1,1),(2,1,0),(2,1,1))\,. \end{aligned}$$

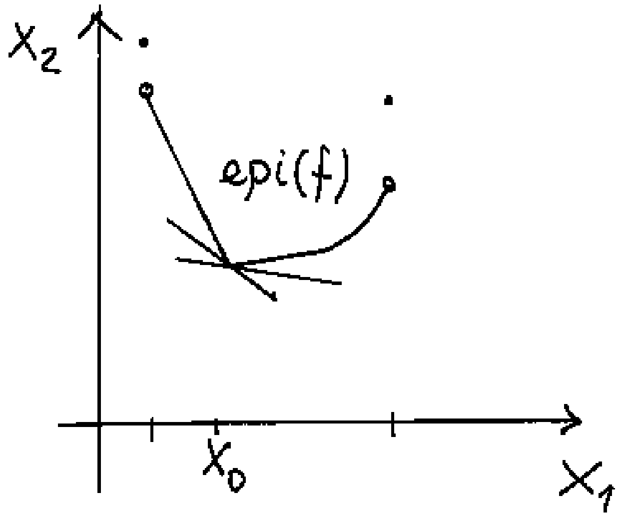
Man kan også prøve at tegne sig ud af problemerne, men det er dog lettere at tegne, når regningerne er gjorte.


X3
(0,0,1)
(1,0,1)
X2
(0,1,0)
(0,0,0)
(1,0,0)
X1

13 Konvekse funktioner af én variabel

Vi betragter i det følgende reelle funktioner defineret på et interval $I \subseteq \mathbb{R}$, og vil opstille kriterier for konveksitet af sådanne. Det vigtigste resultat er Sætning 13.4, som fortæller, at en 2 gange kontinuert differentiabel reel funktion på et interval er konveks hvis og kun hvis den dobbelt afledede funktion er ikke–negativ. Kapitlet indeholder også et vigtigt resultat om konvekse funktioners egenskaber, nemlig Sætning 13.6 som viser, at for en konveks funktion $f : I \to \mathbb{R}$ og et indre punkt $x_0 \in \mathring{I}$, findes $\alpha \in \mathbb{R}$, så for alle $x \in I$ er $f(x) \geq f(x_0) + \alpha(x - x_0)$. Resultatet er illustreret i nedenstående tegninger.





Tegningerne viser, at hvis f er differentiabel i x_0 må $\alpha = f'(x_0)$, men hvis f ikke er differentiabel i x_0, er der mange muligheder for valget af α. Bemærk, at hvis vi lader $\mathrm{affmin}(f)$ betyde mængden af affine minoranter til f på intervallet I;

$$\mathrm{affmin}(f) = \{\varphi : I \to \mathbb{R} \,|\varphi \text{ er affin og } \forall x \in I : \varphi(x) \leq f(x)\}$$

, så defineres der en konveks funktion $g : I \to \mathbb{R}$ ved

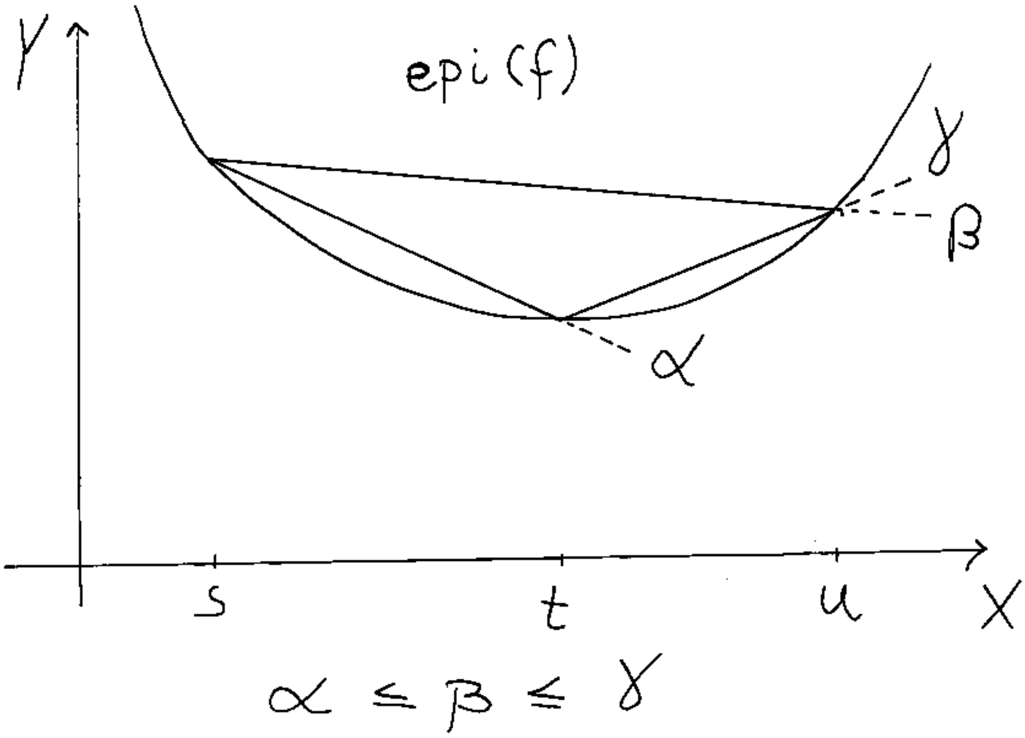
$$\forall x \in I \quad g(x) := \sup\{\varphi(x) | \varphi \in \mathrm{affmin}(f)\} \text{ og}$$

$$\forall x \in I : g(x) \leq f(x) \quad \text{og} \quad \forall x \in \mathring{I} : g(x) = f(x)\,.$$

Lemma 13.1. *Lad $I \subseteq \mathbb{R}$ være et interval, og $f : I \to \mathbb{R}$ en konveks funktion. For vilkårlige 3 punkter $s < t < u$ fra I gælder*

$$\frac{f(t) - f(s)}{t - s} \leq \frac{f(u) - f(s)}{u - s} \leq \frac{f(u) - f(t)}{u - t}\,.$$

Bevis. Resultatet følger af konveksiteten af epi(f) som illustreret herunder



De 3 sekanthældninger benævnes α, β, γ, og det ses, at da $(t, f(t))$ ligger under sekanten $[(s, f(s)), (u, f(u))]$, må $\alpha \leq \beta \leq \gamma$. □

Ud fra lemmaet ses det, at enhver sekanthældning til højre ud fra t er større eller lig α, dvs. det giver mening at definere $f'_+(t)$, *eller den højreafledede af f i t* som infimum over sekanthældningerne til højre. Analogt er enhver sekanthældning for en sekant ud fra $(t, f(t))$ som peger til venstre opadtil begrænset af γ, så vi kan definere $f'_-(t)$, *den venstreafledede af f i t* som supremum af de venstre sekanthældninger ud fra t.

Definition 13.2. Lad $I \subseteq \mathbb{R}$ være et interval og $f : I \to \mathbb{R}$ en konveks funktion. For ethvert indre punkt $t \in I$ defineres

$$f'_+(t) = \inf \left\{ \frac{f(u) - f(t)}{u - t} \;\middle|\; u \in I \,,\, u > t \right\}$$
$$f'_-(t) = \sup \left\{ \frac{f(t) - f(s)}{t - s} \;\middle|\; s \in I \,,\, s < t \right\} .$$

Sætning 13.3. *Lad $I \subseteq \mathbb{R}$ være et interval og $f : I \to \mathbb{R}$ en konveks funktion, da gælder*

i) $\forall t \in \overset{\circ}{I} : f'_-(t) \leq f'_+(t)$.

ii) $\overset{\circ}{I} \ni t \to f'_-(t)$ *er en voksende funktion.*

iii) $\overset{\circ}{I} \ni t \to f'_+(t)$ *er en voksende funktion.*

iv) $\forall t \in \mathring{I} : \alpha \in [f'_-(t), f'_+(t)] \iff \forall x \in I : f(x) \geq f(t) + \alpha(x-t)$.

Bevis.

Ad i) Kig på t, heraf ses, at enhver venstre hældning (α) er mindre eller lig enhver højrehældning (γ), så $f'_-(t) \leq f'_+(t)$.

Ad ii) Kig f.eks. på t, u på tegningen fra før. Heraf aflæses, at $f'_-(t) \leq f'_+(t) \leq \gamma \leq f'_-(u)$, og det ses, at $f'_-(t)$ er en voksende funktion.

Ad iii) Kig her på s og t, og aflæs $f'_+(s) \leq \alpha \leq f'_-(t) \leq f'_+(t)$.

Ad iv) Antag $\alpha \in [f'_-(t), f'_+(t)]$, da gælder for $x > t$, at $\alpha \leq f'_+(t) \leq \frac{f(x)-f(t)}{x-t}$, så for $x > t$ fås (da $x - t > 0$)

$$f(x) - f(t) \geq \alpha(x-t) \quad \text{og} \quad f(x) \geq f(t) + \alpha(x-t)\,.$$

Lad nu $x < t$, da er

$$\frac{f(t)-f(x)}{t-x} \leq f'_-(t) \leq \alpha\,.$$

Da $x < t$ er $t - x > 0$, så

$$f(t) - f(x) \leq \alpha(t-x) \quad \text{og} \quad f(x) \geq f(t) + \alpha(x-t)\,.$$

Lad endelig $x = t$, så er $f(x) = f(t) + \alpha(x-t)$, og ved sammenstykning fås, at for alle $x \in I$ er $f(x) \geq f(t) + \alpha(x-t)$.

For at vise den anden implikation antages nu, at $\alpha \in \mathbb{R}$ opfylder

$$\forall x \in I : f(x) \geq f(t) + \alpha(x-t)\,.$$

Beviset deles igen mht. om $x > t$ eller $x < t$. For $x > t$ fås, idet $(x-t) > 0$, at

$$\forall x > t : \frac{f(x)-f(t)}{x-t} \geq \alpha\,.$$

Da $f'_+(t)$ er infimum over alle højre sekanthældninger, er $\alpha \leq f'_+(t)$. Helt analogt fås for $x < t$

$$\forall x < t : \frac{f(t)-f(x)}{t-x} \leq \alpha \quad \text{så} \quad \alpha \geq f'_-(t)\,.$$

□

Med udgangspunkt i Sætning 13.3 har vi en god del af karakteriseringen af de differentiable konvekse funktioner.

Sætning 13.4. *Lad $I \subseteq \mathbb{R}$ være et åbent interval, og $f : I \to \mathbb{R}$ en differentiabel reel funktion, da gælder*

i) f *er konveks* $\iff$ f' *er voksende.*

ii) Hvis $f \in C^2(I, \mathbb{R})$ *gælder*

$$f \quad \text{er konveks} \quad \iff f'' \geq 0\,.$$

Bevis.

Ad i) At være differentiabel betyder, at $f'_+(t) = f'_-(t) = f'(t)$. Så hvis f er konveks, er $f'(t)$ voksende ifølge Sætning 13.3.

Antag nu, at $f'(t)$ er voksende og at f ikke er konveks, da findes punkter $s < t < u$ i I, så punktet $(t, f(t))$ ligger over sekanten mellem $(s, f(s))$ og $(u, f(u))$

Idet $\alpha := \frac{f(t)-f(s)}{t-s}$, $\gamma := \frac{f(u)-f(t)}{u-t}$, ses det, at antagelsen om at $(t, f(t))$ ligger over sekanten mellem $(s, f(s))$ og $(u, f(u))$ betyder, at $\alpha > \gamma$. Fra middelværdisætningen ved vi imidlertid, at der findes $x \in]s, t[$ så $f'(x) = \alpha$ og $y \in]t, u[$ så $f'(y) = \gamma$, og da f' er voksende må $\gamma \geq \alpha$, hvilket er i modstrid med $\alpha > \gamma$. Antagelsen om at f ikke er konveks må da være forkert, og f er konveks når f' er voksende.

Ad ii) Når $f \in C^2(I, \mathbb{R})$ er f' voksende ensbetydende med, at for alle $x \in I$ er $f''(x) \geq 0$, eller kort, $f'' \geq 0$. □

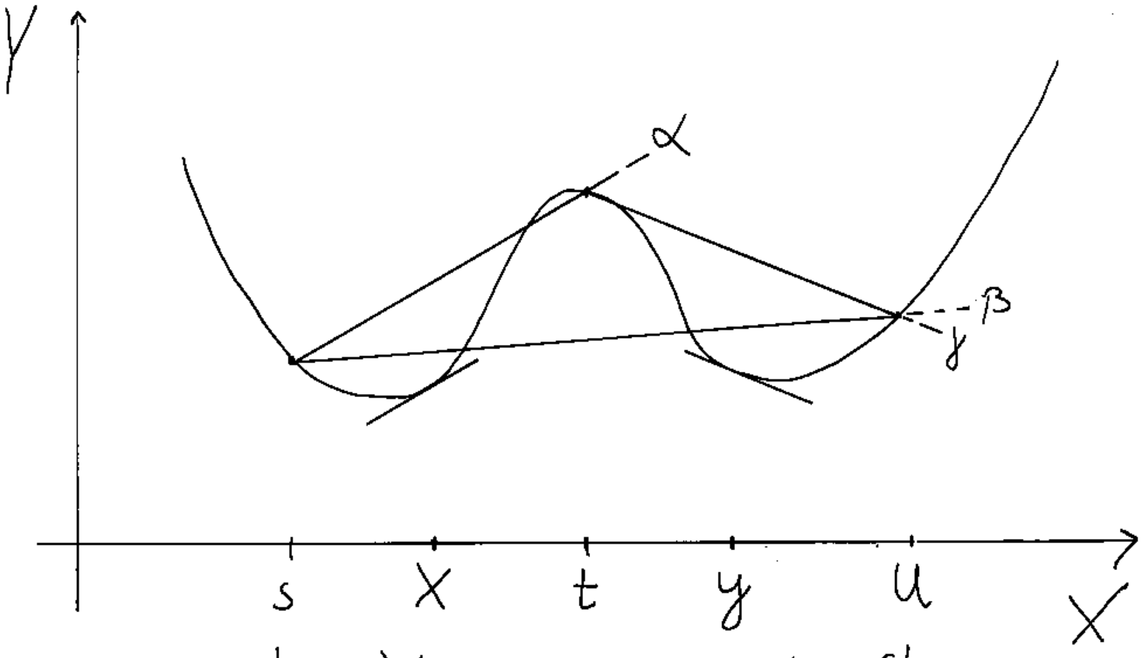
Sætning 13.3 iv) er meget væsentlig. Dette resultat gør det rimeligt allerede nu at introducere begrebet *subdifferentialet*.

Definition 13.5. Lad $\mathcal{D} \subseteq \mathbb{R}^n$ være konveks, og $f : \mathcal{D} \to \mathbb{R}$ en konveks funktion. For $x_0 \in \mathcal{D}$ defineres subdifferentialet — $\partial f(x_0)$ — ved

$$\partial f(x_0) = \{y \in \mathbb{R}^n | \forall x \in \mathcal{D} : f(x) \geq f(x_0) + y \cdot (x - x_0)\}\,.$$

Med dette begreb til rådighed kan Sætning 13.3 iv) omformuleres til

Sætning 13.6. *Lad* $I \subseteq \mathbb{R}$ *være et interval og* $f : I \to \mathbb{R}$ *en konveks funktion. For ethvert indre punkt* $x_0 \in I^\circ$ *er* $\partial f(x_0) = [f'_-(x_0), f'_+(x_0)]$.



$$f'(y) = \gamma \leq \beta \leq \alpha = f'(x)$$

14 Konvekse funktioner af flere variable

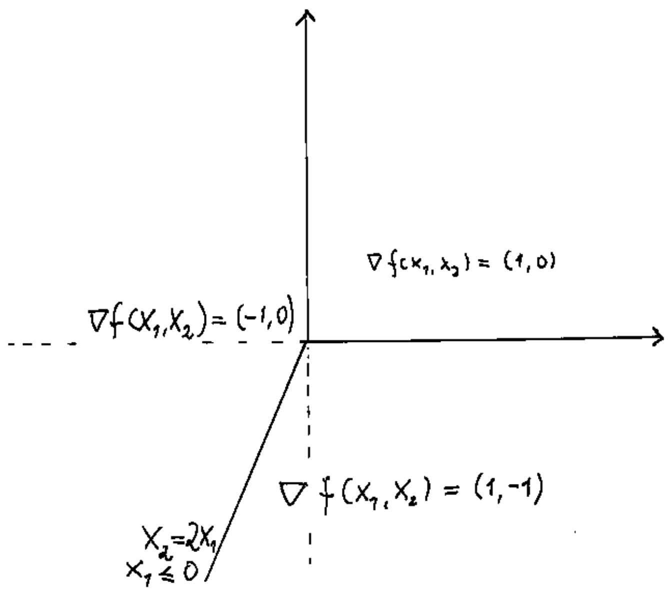
I analogi med kapitel 13 er der her 2 resultater, som er meget vigtige. Det første består i at karakterisere konveksitet af en C^2–funktion ved hjælp af funktionens partielle afledede af 2. orden. Dette aspekt behandles i Matematisk Analyse II af Knut Sydsæter m.fl. i Kapitel 4. Derimod behandler denne bog ikke konvekse funktioner som ikke er differentiable (mindst C^2), og det er en skam, da mange optimeringsproblemer vedrører funktioner som ikke er differentiable, men dog konvekse eller konkave.

Som eksempel kan man i 2 variable kigge på funktionen $f(x_1, x_2) = \max\{|x_1|, x_1 - x_2\}$

$$f(x_1,x_2) = \begin{cases} x_1 & x_1 \geq 0 \text{ og } x_2 \geq 0 \\ x_1 - x_2 & x_1 \geq 0 \text{ og } x_2 \leq 0 \\ -x_1 & x_1 \leq 0 \text{ og } x_2 \geq 2x_1 \\ x_1 - x_2 & x_1 \leq 0 \text{ og } x_2 \leq 2x_1 \end{cases} \quad \begin{array}{ll} \bigtriangledown f(x_1,x_2) = (1,0) & x_1 > 0, x_2 > 0 \\ \bigtriangledown f(x_1,x_2) = (1,-1) & x_1 > 0, x_2 < 0 \\ \bigtriangledown f(x_1,x_2) = (-1,0) & x_1 < 0, x_2 > 2x_1 \\ \bigtriangledown f(x_1,x_2) = (1,-1) & x_1 < 0, x_2 < 2x_1\,. \end{array}$$

Sådanne funktioner optræder i forbindelse med praktiske problemer, så det er langt fra patologisk at interessere sig for ikke–differentiable konvekse funktioner.

Nedenstående skitse viser de 3 områder, hvor $\bigtriangledown f(x_1, x_2)$ er defineret og $\bigtriangledown f(x_1, x_2)$ er anført.



De 3 halvlinier udgør de punkter hvor f ikke er differentiabel.

Standard optimeringsopgaven for en konveks funktion, er at minimere denne. At løse denne opgave vil i praksis sige at beregne subdifferentialer (13.5), idet der gælder følgende VIGTIGE sætning.

Sætning 14.1. *Lad $\mathcal{D} \subseteq \mathbb{R}^n$ være konveks, $f : \mathcal{D} \to \mathbb{R}$ en konveks funktion, og $x_0 \in \mathcal{D}$. Da er $f(x_0)$ en mindsteværdi for f hvis og kun hvis $0 \in \partial f(x_0)$.*

Bevis. Selv om sætningen er vigtig er beviset nemt, idet sætningens indhold blot er en let modifikation af et aspekt af definitionen på begrebet subdifferential. Antag $f(x_0)$ er en mindsteværdi, da er $f(x) \geq f(x_0)$ for alle $x \in \mathcal{D}$, og vi har med $y = 0 = (0, \dots, 0) \in \mathbb{R}^n$ $\forall x \in \mathcal{D} : f(x) \geq f(x_0) + y \cdot (x - x_0)$, så $y = (0, \dots, 0) \in \partial f(x_0)$.

Antag nu, at $y = (0, \dots, 0) \in \partial f(x_0)$, da fås $\forall x \in \mathcal{D} : f(x) \geq f(x_0) + y \cdot (x - x_0) = f(x_0)$, og $f(x_0)$ er en mindsteværdi. □

Det væsentlige indhold i sætningen er at der her gives en tilstrækkelig betingelse for at et punkt er et mindsteværdipunkt. Dette er naturligvis især vigtigt når man behandler teoretiske modeller hvor de indgående størrelser ikke er tal men parametre hvis indbyrdes afhængighed man prøver at udtale sig om.

Det reelle problem i forbindelse med opgaven at minimere en konveks funktion af flere variable, består derfor i at bestemme $\partial f(x)$ for x i f's definitionsmængde. Det vises først, at hvis x er et indre punkt i definitionsmængden, så er $\partial f(x)$ ikke tomt.

Sætning 14.2. *Lad $\mathcal{D} \subseteq \mathbb{R}^n$ være en konveks mængde, $f : \mathcal{D} \to \mathbb{R}$ en konveks funktion, og $x \in \mathcal{D}$.*

Hvis x er et indre punkt i $\mathcal{D}$, er subdifferentialet $\partial f(x)$ ikke tomt.

Bevis. Antag $x_0 \in \overset{\circ}{\mathcal{D}}$, og betragt $(x_0, f(x_0)) \in \operatorname{epi}(f)$. Eftersom det for alle $\alpha < f(x_0)$ gælder at $(x_0, \alpha) \notin \operatorname{epi}(f)$ ses det, at $(x_0, f(x_0))$ ikke kan være et indre punkt i $\operatorname{epi}(f)$. Dernæst vises, at $\operatorname{epi}(f)^\circ \neq \emptyset$ for at kunne anvende Sætning 11.1 til at separere $(x_0, f(x_0))$ fra $\operatorname{epi}(f)^\circ$. Antag nu $\operatorname{epi}(f)^\circ = \emptyset$, så er $\dim(\operatorname{epi}(f)) \leq n < n + 1$ ifølge Sætning 10.3 . Heraf følger, at $\operatorname{aff}(\operatorname{epi}(f)) \neq \mathbb{R}^{n+1}$. Lad $\operatorname{aff}(\operatorname{epi}(f)) = z_0 + \mathcal{L}$, hvor $\mathcal{L}$ er et lineært underrum af dimension lavere end $(n+1)$, dvs. der findes en vektor $y \in \mathcal{L}^\perp \setminus \{0\}$, $y = (y_1, \dots, y_{n+1})$. Efter sin konstruktion gælder for $a \in \operatorname{aff}(\operatorname{epi}(f))$, at $y \cdot a = y \cdot z_0$, dvs.

$$\forall x \in \mathcal{D}\ \forall t \geq f(x) :\ y \cdot (x_1, \dots, x_n, t) = y \cdot z_0\,.$$

Da t kan variere frit i $[f(x), \infty[$ for fastholdt x, må $y_{n+1} = 0$. Da $x_0 \in \overset{\circ}{\mathcal{D}}$ findes $r > 0$, så $x_0 + K(0, r) \subseteq \mathcal{D}$, altså fås, $\forall u \in \mathbb{R}^n$ med $\|u\| < r$ at der gælder $x_0 + u \in \mathcal{D}$, så $y \cdot z_0 = (y_1, \dots, y_n) \cdot (x_0 + u) = (y_1, \dots, y_n) \cdot x_0$, og $(y_1, \dots, y_n) \cdot u = 0$ for $\|u\| < r$, så $(y_1, \dots, y_n) = 0$ og $y = 0$, hvilket er i modstrid med antagelsen $y \neq 0$. Dermed må $\dim(\operatorname{epi}(f)) = n + 1$ og $\operatorname{epi}(f)^\circ \neq \emptyset$.

Lad da $y = (y_1, \dots, y_{n+1}) \in \mathbb{R}^{n+1} \setminus \{0\}$ og $\alpha \in \mathbb{R}$ være valgt så $H(y, \alpha)$ separerer $(x_0, f(x_0))$ og $\overline{\operatorname{epi}(f)}^\circ$, således at $\overset{\circ}{\operatorname{epi}}(f) \subseteq J(-y, -\alpha)$ og $(x_0, f(x_0)) \in J(y, \alpha)$. Af Sætning 10.5 fås $\overline{\operatorname{epi}(f)} \subseteq J(-y, -\alpha)$, og dermed specielt, at $(x_0, f(x_0)) \in J(y, \alpha) \cap J(-y, -\alpha) = H(y, \alpha)$. Dette betyder, at der gælder

$$\forall x \in \mathcal{D}\ \forall t \geq f(x) : y \cdot (x_0, f(x_0)) \leq y \cdot (x, t)\,. \tag{1}$$

Som før ses det, at hvis $y_{n+1} = 0$, så gælder der, da $x_0 \in \mathcal{D}^\circ$, at der findes $r > 0$ så $\forall u \in K(0,r) : x_0 + u \in \mathcal{D}$, og dermed for $t \geq f(x_0 + u)$

$$(y_1, \dots, y_n, 0) \cdot (x_0, f(x_0)) \leq (y_1, \dots, y_n, 0) \cdot (x_0 + u, t)$$
$$(y_1, \dots, y_n) \cdot x_0 \leq (y_1, \dots, y_n) \cdot x_0 + (y_1, \dots, y_n) \cdot u\,,$$

men dette er umuligt når $y \neq 0$, da $u = -\frac{r}{2\|y\|} y \in K(0,r)$ og $y \cdot u < 0$. Vi ser altså at $y_{n+1} \neq 0$, og da (1) gælder for alle $t \geq f(x)$, må $y_{n+1} > 0$. Lad nu $a \in \mathbb{R}^n$ være defineret ved $a_i = \frac{y_i}{y_{n+1}}$, herved kan (1) omformes (ved division med $y_{n+1} > 0$) til (2)

$$\forall x \in \mathcal{D}\ \forall t \geq f(x) :\ (a_1, \dots, a_n, 1) \cdot (x_0, f(x_0)) \leq (a_1, \dots, a_n, 1) \cdot (x, t)\,. \tag{2}$$

For $t = f(x)$ fås så

$$\forall x \in \mathcal{D} : a \cdot x_0 + f(x_0) \leq a \cdot x + f(x) \tag{3}$$
$$\forall x \in \mathcal{D} : f(x) \geq f(x_0) + (-a) \cdot (x - x_0)\,. \tag{4}$$

Heraf ses $-a \in \partial f(x_0)$ og $\partial f(x_0) \neq \emptyset$, når x_0 er et indre punkt. □

Hvis f er differentiabel og konveks gælder følgende:

Sætning 14.3. *Lad $\mathcal{D} \subseteq \mathbb{R}^n$ være åben og konveks, og $f : \mathcal{D} \to \mathbb{R}$ en konveks funktion. Hvis $f \in C^1(\mathcal{D}, \mathbb{R})$, så gælder $\forall x \in \mathcal{D} : \partial f(x) = \bigtriangledown f(x)$.*

Bevis. Lad $x_0 \in \mathcal{D}$ og definer $g(x) := f(x) - (f(x_0) + \bigtriangledown f(x_0) \cdot (x - x_0))$, da er $g(x)$ konveks som sum af en konveks funktion og en affin funktion. (Jensens ulighed er klart opfyldt for g hvis den er det for f). Gradienten $\bigtriangledown g(x_0) = 0$, så ifølge Sætning 0.4 er $g(x) \geq g(x_0) = 0$ for alle $x \in \mathcal{D}$, og dermed

$$\forall x \in \mathcal{D} : f(x) \geq f(x_0) + \bigtriangledown f(x_0) \cdot (x - x_0)\,.$$

Heraf ses $\bigtriangledown f(x_0) \in \partial f(x_0)$.

Antag nu $a \in \partial f(x_0)$, så gælder

$$\forall x \in \mathcal{D} : f(x) \geq f(x_0) + a \cdot (x - x_0)\,,$$

og for $g(x) := f(x) - (f(x_0) + a \cdot (x - x_0))$, er $g(x)$ en differentiabel funktion som opfylder $g(x) \geq 0 = g(x_0)$. Der er mindsteværdi i x_0, og dermed $0 = \bigtriangledown g(x_0) = \bigtriangledown f(x_0) - a$. □

Uden bevis nævnes en nyttig sætning.

Sætning 14.4. *Lad $\mathcal{D} \subseteq \mathbb{R}^n$ være konveks, $f, g : \mathcal{D} \to \mathbb{R}$ konvekse funktioner, $\alpha, \beta \in \mathbb{R}^+$ og $x_0 \in \mathcal{D}^\circ$. Subdifferentialet $\partial(\alpha f + \beta g)(x_0)$ er $\alpha \partial f(x_0) + \beta \partial g(x_0)$.*

Et andet nyttigt resultat er

Sætning 14.5. *Lad $\mathcal{D} \subseteq \mathbb{R}^n$ være konveks, og $f : \mathcal{D} \to \mathbb{R}$ konveks. For $x \in \mathcal{D}$ er $\partial f(x)$ afsluttet og konvekst.*

Bevis. Hvis $\partial f(x) = \emptyset$ eller et punkt, er sagen klar.

Antag nu, at $a, b \in \partial f(x_0)$ og $0 \le \lambda \le 1$, så for $x \in \mathcal{D}$ er

$$\begin{aligned}&f(x) - (f(x_0) + ((1-\lambda)a + \lambda b)\cdot(x - x_0))\\ &= (1-\lambda)(f(x) - (f(x_0) + a\cdot(x-x_0)) + \lambda(f(x) - (f(x_0) + b\cdot(x-x_0)) \ge 0\,.\end{aligned}$$

og $(1-\lambda)a + \lambda b \in \partial f(x_0)$, så $\partial f(x_0)$ er konveks. Hvis $(a_n)_{n\in\mathbb{N}}$ opfylder $a_n \to a$ for $n \to \infty$ og $a_n \in \partial f(x_0)$, fås for $x \in \mathcal{D}$

$$f(x) - (f(x_0) + a\cdot(x-x_0)) = \lim_{n\to\infty}(f(x) - (f(x_0) + a_n\cdot(x-x_0))) \ge 0\,,$$

da $a_n \in \partial f(x_0)$. Heraf ses $a \in \partial f(x_0)$, og $\partial f(x_0)$ er afsluttet. □

Kapitlet slutter med en opskrift på beregning af subdifferentialer for funktioner der er supremum af affine funktioner.

Sætning 14.6. *Lad $y_1, \ldots, y_m \in \mathbb{R}^n$, $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ og definer $f : \mathbb{R}^n \to \mathbb{R}$ ved*

$$\forall x \in \mathbb{R}^n \quad f(x) := \max\{y_i \cdot x + \alpha_i | 1 \le i \le m\}.$$

Funktionen $f(x)$ er konveks, og

$$\partial f(x_0) = \operatorname{conv}\{y_j | 1 \le j \le m \ \text{ og } \ f(x_0) = y_j \cdot x_0 + \alpha_j\}\,.$$

Bevis. Givet $x_0 \in \mathbb{R}^n$ er der altid mindst et $j \in \{1, \ldots, m\}$ så $f(x_0) = y_j \cdot x_0 + \alpha_j$, så vi kan definere en delmængde $J(x_0)$ (ikke–tom) af indeksmængden ved

$$J(x_0) = \{j \in \{1, \ldots, m\} \mid f(x_0) = y_j \cdot x_0 + \alpha_j\}\,.$$

Det skal først vises at for $j \in J(x_0)$ er $y_j \in \partial f(x_0)$, idet dette resultat sammen med Sætning 14.5 viser at $\operatorname{conv}(\{y_j | j \in J(x_0)\}) \subseteq \partial f(x_0)$. Hold da $j \in J(x_0)$ fast. Efter definitionen af f som det punktvise maksimum af de affine funktioner, fås

$$\begin{aligned}&\forall x \in \mathbb{R}^n : f(x) \ge y_j \cdot x + \alpha_j = y_j \cdot x_0 + \alpha_j + y_j \cdot (x - x_0)\\ &\text{da } \ j \in J(x_0) \qquad\qquad\qquad = f(x_0) + y_j \cdot (x - x_0)\,,\end{aligned}$$

altså $j \in J(x_0) \implies y_j \in \partial f(x_0)$, som ønsket og $\operatorname{conv}(\{y_j | j \in J(x_0)\}) \subseteq \partial f(x_0)$.

For at vise den anden inklusion $\operatorname{conv}(\{y_j | j \in J(x_0)\}) \supseteq \partial f(x_0)$ betragtes et $z \notin \operatorname{conv}(\{y_j | j \in J(x_0)\})$, $(z \in \mathbb{R}^n)$ og det vises at $z \notin \partial f(x_0)$. Med henblik på at vise dette bemærkes det at da det konvekse hylster af endeligt mange punkter er afsluttet, kan z og $\operatorname{conv}(\{y_j | j \in J(x_0)\})$ separeres stærkt. Der findes derfor $\gamma < \beta$ og $u \in \mathbb{R}^n \setminus \{0\}$, så

$$\operatorname{conv}(\{y_j | j \in J(x_0)\}) \subseteq J(u, \gamma) \ \text{ og } \ u \cdot z \ge \beta\,.$$

Specielt fås for

$$\forall j \in J(x_0) : u \cdot y_j \le \gamma \ \text{ og } \ u \cdot z \ge \beta > \gamma\,. \tag{1}$$

Dette kan udnyttes til at vise $z \notin \partial f(x_0)$, men vejen hertil er lidt snørklet.

Først bemærkes, at $f(x_0) = y_j \cdot x_0 + \alpha_j$ for $j \in J(x_0)$, og $f(x_0) > y_j \cdot x_0 + \alpha_j$ for alle $j \in \{1, \dots, m\} \setminus J(x_0)$. Da de affine funktioner $\varphi_j(x) := y_j \cdot x + \alpha_j$ alle er kontinuerte, er funktionen $f(x) = \sup\{\varphi_j(x) | 1 \le j \le m\}$ også kontinuert, og fællesmængden

$$\mathcal{V} = \bigcap_{\substack{j \notin J(x_0) \\ 1 \le j \le m}} \{x \in \mathbb{R}^n | (f - \varphi_j)^{-1}(]0, \infty[)\}$$

er en åben mængde som indeholder x_0. Der findes da $r > 0$ så $\overline{K(x_0, r)} \subseteq \mathcal{V}$, eller med andre ord

$$\forall x \in \mathbb{R}^n : \|x - x_0\| \le r \implies \forall_j \notin J(x_0) \text{ er } f(x) > \varphi_j(x) .$$

Da $f(x) = \sup\limits_{1 \le j \le m} \varphi_j(x)$ får man derfor af ovenstående udsagn, at der må gælde

$$\forall x \in \mathbb{R}^n : \|x - x_0\| \le r \implies f(x) = \max_{j \in J(x_0)} \varphi_j(x_0) . \tag{2}$$

(Dette kan gives en geometrisk fortolkning, men den udelades). Vælges nu $x_1 = x_0 + \frac{r}{\|u\|} u$, så gælder ifølge ovenstående (1) og (2)

$$f(x_1) = \max_{j \in J(x_0)} (y_j \cdot x_1 + \alpha_j) = (y_j \cdot x_0 + \alpha_j) + \max_{j \in J(x_0)} \left(\frac{r}{\|u\|} u \cdot y_j \right) \le f(x_0) + \frac{r}{\|u\|} \gamma .$$

Det er som bekendt målet med disse argumenter at vise at $z \notin \partial f(x_0)$. For at vise dette er det blot nødvendigt at vise at den affine funktion ψ, som er defineret ved $\psi(x) := f(x_0) + z \cdot (x - x_0)$, ikke er en affin minorant for f. Men de netop gennemførte regninger viser

$$\psi(x_1) = f(x_0) + \frac{r}{\|u\|} z \cdot u \ge f(x_0) + \frac{\beta \cdot r}{\|u\|} > f(x_0) + \frac{\gamma r}{\|u\|} \ge f(x_1) .$$

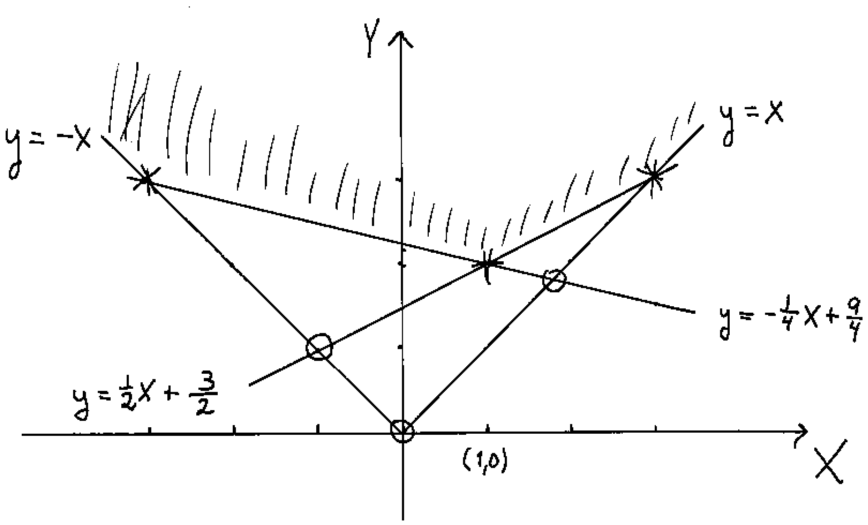
så ψ er ikke en affin minorant, $z \notin \partial(f x_0)$ og $\operatorname{conv}(\{y_j | j \in J(x_0)\}) = \partial f(x_0)$. Sætnigen følger. $\square$

Eksempler 14.7. Vi regner et eksempel med $n = 1$ først for at få lidt intuition for anvendelsen af Sætning *14.6.*

Lad

$$\begin{aligned} f(x) &= \max\{|x|, -\frac{1}{4}x + \frac{9}{4}, \frac{1}{2}x + \frac{3}{2}\} \\ &= \max\{x, -x, -\frac{1}{4}x + \frac{9}{4}, \frac{1}{2}x + \frac{3}{2}\} \\ &= \max\{\varphi_1(x), \varphi_2(x), \varphi_3(x), \varphi_4(x)\} . \end{aligned}$$

Her kan opgaven løses grafisk, så lad os gøre det først



$f(x)$	$\varphi_i(x)$	Interval	$J(x)$	$\partial f(x)$
$-x$	$-x$	$-\infty < x < -3$	2	-1
3	$-x, -\frac{1}{4}x + \frac{9}{4}$	-3	$2, 3$	$[-1, -\frac{1}{4}]$
$-\frac{1}{4}x + 9$	$-\frac{1}{4}x + \frac{9}{4}$	$-3 < x < 1$	3	$-\frac{1}{4}$
2	$-\frac{1}{4}x + \frac{9}{4}, \frac{1}{2}x + \frac{3}{2}$	1	$3, 4$	$[-\frac{1}{4}, \frac{1}{2}]$
$\frac{1}{2}x + \frac{3}{2}$	$\frac{1}{2}x + \frac{3}{2}$	$1 < x < 3$	4	$\frac{1}{2}$
3	$\frac{1}{2}x + \frac{3}{2}, x$	3	$1, 4$	$[\frac{1}{2}, 1]$
x	x	$3 < x < \infty$	1	1

Det ses, at $J(x)$ består af et enkelt tal for alle x bortset fra nogle få — endeligt mange – punkter. Idet $\varphi_i'(x) \neq 0$ for alle i og alle x, ses det at muligheden for at $0 \in \partial f(x)$ kun kan være til stede i knækpunkterne. Da $0 \in [-\frac{1}{4}, \frac{1}{2}] = \partial f(1)$ er mindsteværdien $f(1) = 2$.

Man kan også tænke på epi(f) som et polyeder, der er bestemt ved:

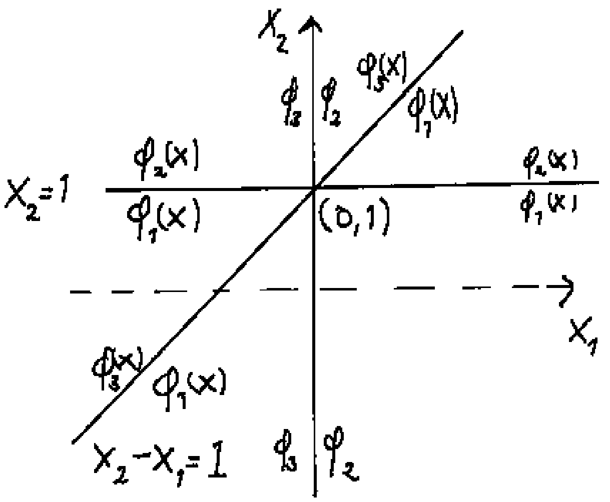
$$\begin{aligned}\text{epi}(f) &= \{(x, y) \in \mathbb{R}^2 | y \geq x\,,\ y \geq -x\,,\ y \geq -\frac{1}{4}x + \frac{9}{4}\,,\ y \geq \frac{1}{2}x + \frac{3}{2}\} \\ &= J(1, -1), 0) \cap J((-1, -1), 0) \cap J((-\frac{1}{4}, -1), -\frac{9}{4}) \cap J((\frac{1}{2}, -1), -\frac{3}{2})\,.\end{aligned}$$

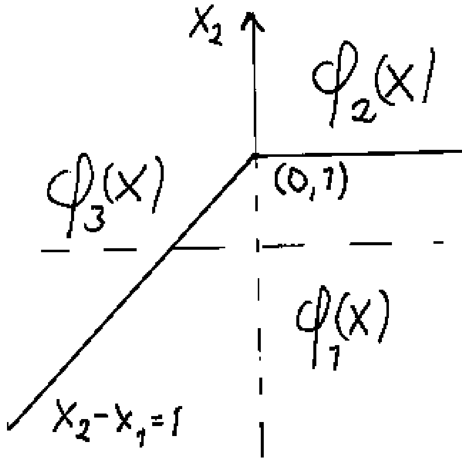
For så at bruge teorien til at bestemme hjørnerne med, og blandt disse vælge hjørnet med minimal y værdi. Pas dog på, man skal også igennem overvejelserne om epi(f) indeholder linier eller ej.

Lad os prøve et eksempel med funktioner af 2 variable

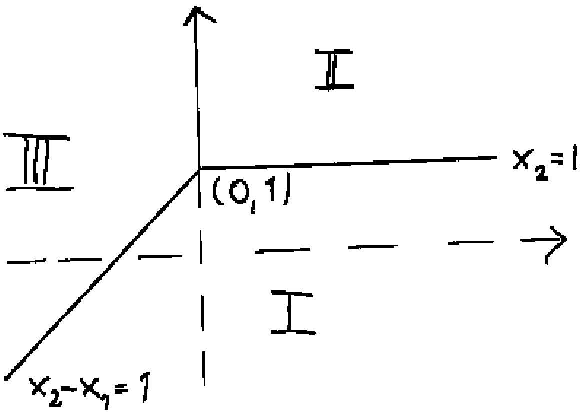
$$f(x_1, x_2) = \max\{x_1 - 3x_2 + 2\,,\ x_1 + x_2 - 2\,,\ -3x_1 + x_2 - 2\} = \max\{\varphi_1(x), \varphi_2(x), \varphi_3(x)\}\,.$$

$$\begin{aligned}
\ell_1 &:= \{x | \varphi_2(x) = \varphi_3(x)\} = \{(x_1, x_2) | x_1 = 0\}\ \varphi_2(x) > \varphi_3(x) \iff x_1 > 0 \\
\ell_2 &:= \{x | \varphi_1(x) = \varphi_3(x)\} = \{(x_1, x_2) | x_1 - x_2 = -1\} \varphi_1(x) > \varphi_3(x) \iff x_1 - x_2 > -1 \\
\ell_3 &:= \{x | \varphi_1(x) = \varphi_2(x)\} = \{(x_1, x_2) | x_2 = 1\}\ \varphi_1(x) > \varphi_2(x) \iff x_2 < 1
\end{aligned}$$





Skitsen nedenfor viser områderne hvor $f(x) = \varphi_1(x)$, $\varphi_2(x)$ eller $\varphi_3(x)$. Grænserne mellem disse områder er de 3 halvlinier p, q, r ud fra punktet $(0, 1)$



Sætning 14.6 giver nu, at

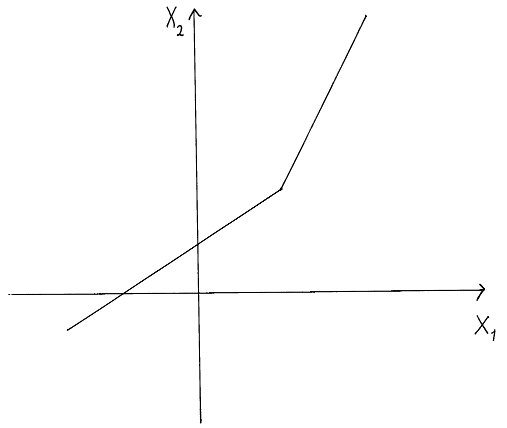
$$\partial f(x) = \begin{cases} (1,-3) & x \in \mathrm{I}^\circ \\ (1,1) & x \in \mathrm{II}^\circ \\ (-3,1) & x \in \mathrm{III}^\circ \\ [(1,-3),(-3,1)] & x \in p \setminus \{(0,1)\} \\ [(1,-3),(1,1)] & x \in q \setminus \{(0,1)\} \\ [(1,1),(-3,1)] & x \in r \setminus \{(0,1)\} \\ \operatorname{conv}(\{(1,1),(-3,1),(1,-3)\}) & x = (0,1)\,. \end{cases}$$

Det ses, at

$$\begin{aligned}(0,0) &= \frac{1}{2}(1,1) + \frac{1}{4}(-3,1) + \frac{1}{4}(1,-3)\\ &\in \operatorname{conv}\left(\{(1,1),(-3,1),(1,-3)\}\right),\end{aligned}$$

så f har mindsteværdi i $(0,1)$, og værdien er -1. Såfremt man kun er interesseret i mindsteværdien, søger man i analogi med metoden til at finde hjørnerne i et polyeder, efter mulige punkter som fastlægges af at $J(x)$ indeholder $n+1$ indices. I dette tilfælde er der kun en mulighed; $\varphi_1(x) = \varphi_2(x) = \varphi_3(x)$. I det generelle tilfælde undersøges herefter om $f(x) = \varphi_{i_1}(x) = \varphi_{i_{n+1}}(x)$. Dette behøver ikke altid at være tilfældet, som tegningen til det første eksempel viser i punkterne $x = -1$, $x = 0$, $x = \frac{9}{5}$. Blandt de punkter hvor $\varphi_{i_j}(x) = f(x)$ for $1 \leq j \leq n+1$ findes så dem med mindst værdi, og det undersøges om $0 \in \partial f(x)$ for et af disse punkter.

Dette check er nødvendigt som nedenstående skitse viser. Lad $f(x) = \max\{x+1, 2x\}$



Her er $\inf f(x) = -\infty$, men der er et knækpunkt i punktet $(1,2)$ og subdifferentialet i punktet 1 er $\partial f(1) = [1,2]$.

Ovenstående er ikke en gennemgang i alle detaljer, idet f. eks. tilfældene hvor polyederet $\operatorname{epi}(f)$ har linier ikke er behandlet.

Indeks

Opgavesamling

Indledning.

Opgavesamlinger vokser ved knopskydninger og ved import fra kolleger i ind- og udland. Denne samling er ingen undtagelse i så henseende idet den er baseret på et udvalg fra nedennævnte 4 kilder:

Tage Bai Andersen:

Opgaver til optimering (konveks analyse), I, II. Aarhus Universitet 1984.

Arne Brøndsted:

Noter til Matematik 2 OK. Københavns Universitet 1990.

Arne Strøm & Knut Sydsæter:

Eksamensopgaver, Matematikk for økonomer. Universitetet i Oslo 1995.

Staben, Matematisk Afdeling:

Eksamensopgaver i Matematik 2 OK. Københavns Universitet.

Indhold.

Opgaverne er inddelt i 5 afsnit svarende til hovedemnerne i Matematik 2 OK. En samling af tidligere eksamensopgaver uddeles til kursets deltagere.

Indholdsfortegnelse

1. Konvekse Mængder

OPGAVE 1. Vis, at der for vilkårlige lineære underrum L_1 og L_2 i $\mathbb{R}^n$ og vilkårlige vektorer z_1 og z_2 i $\mathbb{R}^n$ gælder

$$z_1 + L_1 = z_2 + L_2 \Rightarrow L_1 = L_2.$$

OPGAVE 2. Lad L være et vilkårligt lineært underrum i $\mathbb{R}^n$, og lad z_1 og z_2 være vilkårlige vektorer i $\mathbb{R}^n$. Vis, at følgende betingelser er ensbetydende:

(a) $z_1 + L = z_2 + L$.

(b) $z_2 - z_1 \in L$.

(c) $z_2 \in z_1 + L$.

(d) $(z_1 + L) \cap (z_2 + L) \neq \emptyset$.

(Ækvivalensen af (a) og (c) er postuleret i noterne side 2.2^{3-4}.)

OPGAVE 3. Lad $A_1 = z_1 + L_1, \ldots, A_k = z_k + L_k$ være affine underrum i $\mathbb{R}^n$. Vis, at det affine underrum

$$\bigcap_{i=1}^{k} A_i$$

enten er $\emptyset$ eller har formen

$$z + \bigcap_{i=1}^{k} L_i.$$

OPGAVE 4. I $\mathbb{R}^4$ betragtes mængden

$$A := \{(x_1, x_2, x_3, x_4) \mid x_1 - x_2 + 3x_3 + x_4 = 3\}.$$

1° Begrund, at A er et affint underrum og bestem $\dim A$.

2° Angiv en affin basis for A, og fremstil punktet $(1, -1, 0, 1)$ som affin kombination af punkterne i den valgte basis.

OPGAVE 5. Vis, at en afbildning $f : \mathbf{R}^n \to \mathbb{R}^m$ er affin, hvis og kun hvis der findes en $m \times n$ matrix A og en $m \times 1$ matrix B, således at

$$f(X) = B + AX.$$

(Punkterne i $\mathbb{R}^n$ opfattes her som $n \times 1$ søjlematricen X. Billedpunkterne $f(X)$ i $\mathbb{R}^m$ opfattes som $m \times 1$ søjlematricer.)

OPGAVE 6. I $\mathbb{R}^3$ betragtes punkterne

$$\begin{aligned} &x_1 = (0,0,0), \quad x_2 = (1,0,0), \quad x_3 = (0,1,0), \quad x_4 = (0,0,1), \\ &y_1 = (1,2,0), \quad y_2 = (3,-1,-1), \quad y_3 = (2,4,1), \quad y_4 = (0,1,0). \end{aligned}$$

Gør rede for, at der findes en og kun én affin afbildning $f : \mathbb{R}^3 \to \mathbb{R}^3$ med $f(x_i) = y_i$, $i = 1, 2, 3, 4$. Angiv f på matrixform som i Sætning 2.10. Undersøg, om f er bijektiv, altså en affin isomorfi.

OPGAVE 7. Når C er en delmængde af $\mathbb{R}^n$, og λ er et reelt tal, sætter man

$$\lambda C := \{\lambda x \mid x \in C\}.$$

I stedet for $(-1)C$ skriver man $-C$. Man siger, at C er symmetrisk, når $C = -C$. Vis, at når C er konveks, så er også λC konveks for alle $\lambda \in \mathbb{R}$.

OPGAVE 8. Når C_1 og C_2 er delmængder af $\mathbb{R}^n$, sætter man

$$C_1 + C_2 := \{x_1 + x_2 \mid x_1 \in C, x_2 \in C\}.$$

Når C_1 er en étpunkts mængde $\{x\}$, skriver man også $x + C_2$ i stedet for $\{x\} + C_2$.

1° Vis, at når C_1 og C_2 er konvekse, så er også $C_1 + C_2$ konveks.

2° Generalisér til et vilkårligt (endeligt) antal k af konvekse mængder $C_1, \dots, C_k$.

OPGAVE 9. Bevis, at afsluttede liniestykker $[x_0, x_1]$, halvåbne liniestykker $[x_0, x_1[$ og åbne liniestykker $]x_0, x_1[$ i $\mathbb{R}^n$ er konvekse mængder.

OPGAVE 10. Lad C være trekanten i $\mathbb{R}^2$ med hjørner $(2,2)$, $(4,2)$, og $(2,4)$, og lad D være den afsluttede cirkelskive med centrum i $(0,0)$ og radius 1.

1° Tegn $C, D, \frac{1}{2}C, -2C, -D, (1,1)+C, (1,1)+D, 3((1,1)+C), 3((1,1)+D), C+D$ og $C - D$.

2° Er C symmetrisk? Er D? Er $(1,1) + D$?

OPGAVE 11. For $x \in \mathbb{R}^n$ og $r \in \mathbb{R}_+$ betegner $K(x,r)$ som sædvanlig den åbne kugle med centrum i x og radius r.

1° Vis, at $K(x,r)$ er en konveks mængde.

2° Vis, at $K(0,r) = rK(0,1)$.

3° Vis, at $K(x,r) = x + K(0,r)$.

4° Vis, at $K(0,r) + K(0,s) = (r+s)K(0,1)$.

OPGAVE 12. Sæt $M := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 x_2 = 1, x_1 > 0, x_2 > 0\}$.

1° Bestem $\operatorname{conv} M$.

2° Bestem $\operatorname{conv}(M_1 \cup M)$, hvor $M_1 = \{(0,0)\}$.

3° Bestem $\operatorname{conv}(M_2 \cup M)$, hvor $M_2 = \{(0,0), (0,1)\}$.

4° Bestem $\operatorname{conv}(M_3 \cup M)$, hvor M_3 er x_2-aksen.

OPGAVE 13. Lad C_1 og C_2 være konvekse mængder i $\mathbb{R}^n$. Vis, at $\operatorname{conv}(C_1 \cup C_2)$ "sædvanligvis" består af alle liniestykker $[x_1, x_2]$, hvor $x_1 \in C_1$ og $x_2 \in C_2$. Forklar, hvad der menes med "sædvanligvis".

OPGAVE 14. Lad M være en mængde bestående af en linie og et punkt uden for linien. Bestem $\operatorname{conv} M$. Er M afsluttet? Er $\operatorname{conv} M$ afsluttet?

OPGAVE 15. Lad C være en afsluttet konveks mængde i $\mathbb{R}^n$. Et punkt $x \in C$ siges at være et *ekstremt punkt*, dersom $C \setminus \{x\}$ er konveks. Mængden af C's ekstreme punkter betegnes $\operatorname{ext} C$.

1° Gør rede for, at $x \in C$ er et ekstremt punkt, hvis og kun hvis der ikke findes to forskellige punkter y og z i C, således at x ligger på det åbne liniestykke

$$]y, z[:= \{(1-\lambda)y + \lambda z \mid \lambda \in]0,1[\}.$$

2° Bestem de ekstreme punkter i følgende mængder:

$$\begin{aligned} C_1 &:= \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}. \\ C_2 &:= \{(x_1, x_2) \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\}. \\ C_3 &:= \{(x_1, x_2) \in \mathbb{R}^2 \mid 0 \leq x_1, 0 \leq x_2\}. \\ C_4 &:= \{(x_1, x_2) \in \mathbb{R}^2 \mid 0 \leq x_1\}. \end{aligned}$$

OPGAVE 16. I $\mathbb{R}^3$ betragtes mængderne

$$\begin{aligned} C_1 &:= \{(x_1, x_2, x_3) \mid x_1^2 + (x_2 - 1)^2 \leq 1,\ x_3 = 0\}, \\ C_2 &:= \{(x_1, x_2, x_3) \mid x_1 = x_2 = 0,\ -1 \leq x_3 \leq 1\}. \end{aligned}$$

1° Tegn og beskriv $C := \operatorname{conv}(C_1 \cup C_2)$.

2° Angiv $\operatorname{ext} C$.

3° Undersøg, om $\operatorname{ext} C$ er afsluttet.

OPGAVE 17. Lad C være en afsluttet konveks mængde, og lad M være en delmængde af C med $C = \operatorname{conv} M$. Vis, at $\operatorname{ext} C \subseteq M$.

OPGAVE 18. I $\mathbb{R}^2$ betragtes punkterne $x_1 = (1, 5)$, $x_2 = (-4, -3)$, $x_3 = (4, -5)$, $x_4 = (0, 0)$ og $x_5 = (-\frac{3}{2}, 1)$.

1° Vis, at x_4 og x_5 er konvekse kombinationer af x_1, x_2 og x_3. Angiv sådanne konvekse kombinationer.

2° Beskriv det konvekse hylster af $\{x_1, x_2, x_3, x_4, x_5\}$.

OPGAVE 19. I $\mathbb{R}^3$ betragtes punkterne $x_1 = (2, 1, -1)$, $x_2 = (7, 3, -9)$, $x_3 = (-8, 7, -7)$ og $x_4 = (0, 4, 6)$. Beskriv det konvekse hylster af $\{x_1, x_2, x_3, x_4\}$.

OPGAVE 20. I $\mathbb{R}^3$ betragtes punkterne $x_1 = (2, 1, -1)$, $x_2 = (7, 3, -9)$, $x_3 = (8, 7, -7)$ og $x_4 = (0, 4, 6)$. Beskriv det konvekse hylster af $\{x_1, x_2, x_3, x_4\}$.

OPGAVE 21. Lad $y = (y_1, \ldots, y_n)$ være en vektor i $\mathbb{R}^n$ med $y_1 > 0, \ldots, y_n > 0$. Bevis, at for alle $\alpha \in \mathbb{R}$ er $\overline{\mathbb{R}}_+^n \cap J(y, \alpha)$ et begrænset konvekst polyeder. Angiv hjørnerne i $\overline{\mathbb{R}}_+^n \cap J(y, 1)$.

OPGAVE 22. Lad K være en konveks kegle. Vis, at $\operatorname{span} K = K - K$.

OPGAVE 23. Lad K være en konveks kegle. Vis, at $(-K) \cap K$ er det "største" lineære underrum indeholdt i K.

OPGAVE 24. Lad K være en konveks kegle. Ved K's *normalkegle* forstås mængden

$$K^\circ := \{y \in \mathbb{R}^n \mid \forall x \in K : x \cdot y \leq 0\}.$$

Vis, at K° er en afsluttet konveks kegle.

OPGAVE 25. Angiv en ikke-afsluttet konveks kegle i $\mathbb{R}^n$. Findes der åbne konvekse kegler i $\mathbb{R}^n$?

OPGAVE 26. Ved det *afsluttede konvekse hylster* $\operatorname{cl}\operatorname{conv} M$ af en mægnde M i $\mathbb{R}^n$ forstås fællesmængden af alle afsluttede konvekse mængder, som indeholder M. Gør rede for, at $\operatorname{cl}\operatorname{conv} M$ er den "mindste" afsluttede konvekse mængde, som indeholder M. Bevis, at $\operatorname{cl}\operatorname{conv} M = \operatorname{cl}(\operatorname{conv} M)$.

OPGAVE 27. Gør rede for, at der for vilkårlige mængder M_1 og M_2 i $\mathbb{R}^n$ gælder

$$M_1 \subseteq M_2 \Rightarrow \mathring{M}_1 \subseteq \mathring{M}_2.$$

Undersøg, om der for konvekse mængder C_1 og C_2 gælder

$$C_1 \subseteq C_2 \overset{?}{\Rightarrow} \operatorname{ri} C_1 \subseteq \operatorname{ri} C_2.$$

OPGAVE 28. Lad $P = \operatorname{conv}\{x_1, \ldots, x_k\}$ være en konveks polytop i $\mathbb{R}^n$. Bevis, at et punkt $x \in \mathbb{R}^n$ ligger i $\operatorname{ri} P$, hvis og kun hvis x har en fremstilling

$$x = \sum_{i=1}^{k}{}^{c}\lambda_i x_i,$$

hvori alle λ_i er > 0.

OPGAVE 29. Lad $f : \mathbb{R}^n \to \mathbb{R}^m$ være en affin afbildning. Vis, at der for enhver konveks mængde C i $\mathbb{R}^n$ gælder $\operatorname{ri} f(C) = f(\operatorname{ri} C)$.

OPGAVE 29. Lad $f : \mathbb{R}^n \to \mathbb{R}^m$ være en affin afbildning. Vis, at der for enhver konveks mængde C i $\mathbb{R}^n$ gælder $\operatorname{ri} f(C) = f(\operatorname{ri} C)$.

OPGAVE 30. Vis, at hvis $\mathcal{C}$ er en afsluttet konveks mængde så er $\operatorname{recc}(\mathcal{C})$ afsluttet.

OPGAVE 31. Vis, at for enhver konveks mængde $\mathcal{C}$ gælder $\operatorname{recc}(\mathcal{C}) \subseteq \operatorname{recc}(\overline{\mathcal{C}})$.

OPGAVE 32. Lad C være en ikke-tom afsluttet konveks mængde i $\mathbb{R}^n$. Ved en *side* i C forstås en konveks delmængde F af C med egenskaben, at hvis y og z er to forskellige punkter i C, således at $]y, z[\cap F \neq \emptyset$, så ligger både y og z i F. En side F siges at være *egentlig*, dersom $F \neq \emptyset$ og $F \neq C$.

1° Vis, at et punkt $x \in C$ er et ekstremt punkt, jf. Opgave 15, hvis og kun hvis $\{x\}$ er en side.

2° Angiv alle sider i en terning, i en afsluttet kugle i $\mathbb{R}^3$, i et afsluttet halvrum i $\mathbb{R}^n$ og i et n-simpleks.

3° Gør rede for, at hvis H er en støttehyperplan for C med $C \nsubseteq H$, så er $H \cap C$ en egentlig side i C. -En side af denne art kaldes en *eksponeret side.* Et punkt x i C kaldes et *eksponeret punkt*, dersom $\{x\}$ er en eksponeret side.

4° Giv et eksempel på en afsluttet konveks mængde C i $\mathbb{R}^2$, som har ikke-eksponerede sider.

5° Bevis, at enhver side i C er afsluttet.

6° Lad F være en side i C, og lad G være en delmængde af F. Vis, at G er en side i F, hvis (og kun hvis) G er en side i C.

7° Gør rede for, at der for enhver delmængde M af C findes en "mindste" side F i C, som indeholder M.

8° Lad F være en side i C og lad x være et punkt i F. Vis, at F er den mindste side i C, som indeholder x, hvis og kun hvis $x \in \operatorname{ri} F$.

9° En *facet* i C er en side F med $\dim F = \dim C - 1$. Vis, at enhver ikke tom facet er en eksponeret side.

OPGAVE 33. Lad B og C være de konvekse mængder i $\mathbb{R}^2$:

$$B = \{(x, y) \mid x^2 + y^2 \leq 1\}\,,\ C = \{(x, y) \mid |x| \leq 1\,,\ |y| \leq 1\}.$$

Hvilke delmængder S af randen for B kan man fjerne så resten $B \setminus S$ stadig er konveks?

Det tilsvarende spørgsmål stilles for C.

OPGAVE 34. Lad C være konveks og $\alpha, \beta \in \mathbb{R}$. Vis, at for $\alpha > 0$, $\beta > 0$, så er

$$(\alpha + \beta)C = \alpha C + \beta C\,.$$

Gælder en tilsvarende formel for vilkårlige $\alpha, \beta \in \mathbb{R}$?

OPGAVE 35. Lad B være en $m \times n$ matrix og $b \in \mathbb{R}^m$. For vektorer $b, c \in \mathbb{R}^n$ defineres $b \leq c$ hvis det for alle indices $i \in \{1, \ldots, n\}$ gælder $b_i \leq c_i$. Antag, at mængden

$$C = \{x \in \mathbb{R}^n \mid Bx \leq b\}$$

er et affint underrum. Vis at så findes der en vektor $c \in \mathbb{R}^m$, så at

$$C = \{x \in \mathbb{R}^n \mid Bx = c\}$$

(med samme matrix B).

OPGAVE 36. Lad A være en $m \times n$ matrix. Lad $x_0 \in \mathbb{R}^n$ og $b = Ax_0$.

Lad C betegne den affine mængde

$$C = \{x \in \mathbb{R}^n \mid Ax = b\}\,.$$

Find det parallelle underrum til C og angiv dimensionen af C (udtrykt ved hjælp af m, n og rangen for A).

OPGAVE 37. Lad C være en konveks mængde i $\mathbb{R}^n$, og lad $x \in \mathbb{R}^n$ med egenskaben $x \notin \operatorname{aff}(C)$.

Lad $S = \operatorname{conv}(C \cup \{x\})$.

Vis, at $\dim S = 1 + \dim C$.

OPGAVE 38. Lad $S \subseteq \mathbb{R}^n$ være kompakt. Så er $\operatorname{conv}(S)$ kompakt.

OPGAVE 39. Lad $a_1, \cdots, a_m \in \mathbb{R}^n$. Lad

$$K = \{\lambda_1 a_1 + \cdots + \lambda_m a_m \mid \lambda_i \geq 0\,,\ i = 1, \ldots, m\}\,.$$

Så er K konveks og lukket.

OPGAVE 40. Vis at når C, D er konvekse mængder i $\mathbb{R}^n$, så er $C \times D$ en konveks mængde i $\mathbb{R}^{2n}$ og

$$\operatorname{ri}(C \times D) = \operatorname{ri}(C) \times \operatorname{ri}(D)\,.$$

Illustrer i tilfældet $n = 1$.

OPGAVE 41. Lad C være en konveks mængde i $\mathbb{R}^n$ og lad K være defineret ved

$$K = \{\lambda c \mid c \in C\,,\ \lambda \geq 0\}\,.$$

Illustrer for $n = 2$, – både i tilfældet, hvor $0 \notin C$ og hvor $0 \in C$.

1. Vis, at K er en konveks mængde. (K kaldes den konvekse kegle frembragt af C).
2. Vis, at $\dim K$ enten er lig med: $\dim C$ eller med: $1 + \dim C$.
3. Antag, at $0 \notin C$ og at C er kompakt. Vis, at

   $$\operatorname{ri} K = \{\lambda > 0\,,\ c \in \operatorname{ri}(C))\}\,.$$

   Vis, at K er lukket.
4. Er udsagnene i spørgsmål 4 korrekte, hvis vi dropper én af betingelserne $0 \notin C$ eller er C kompakt, – gælder det f.eks. at hvis $0 \notin C$ og C er lukket, så er K lukket ?

OPGAVE 42. Lad C være en kompakt og konveks mængde i $\mathbb{R}^n$. Lad

$$D = \{\lambda c \mid 0 \leq \lambda \leq 1\,,\ c \in C\}\,.$$

Vis, at D er kompakt og konveks. Illustrer.

OPGAVE 43. Gælder følgende: Hvis C, D er lukkede, konvekse mængder i $\mathbb{R}^n$, så er $C + D$ lukket ?

OPGAVE 44. Lad $K, P \subseteq \mathbb{R}^n$ være kompakte konvekse mængder i $\mathbb{R}^n$.

Vis, at $K \times P$ er kompakt og konveks i $\mathbb{R}^{2n}$ og at $K + P$ er kompakt og konveks i $\mathbb{R}^n$.

OPGAVE 45. Lad C være en lukket konveks mængde i $\mathbb{R}^n$, så at $0 \in \operatorname{int}(C)$.

Lad for ethvert $x \in \mathbb{R}^n$:

$$\gamma(x|C) = \inf\{\lambda \geq 0 \mid x \in \lambda C\}\,.$$

Vis, at $\gamma(\cdot|C)$ er konveks, positiv homogen og lukket.

Vis, at $\operatorname{dom}\gamma(\cdot|C) = \mathbb{R}^n$, at

$$C = \{x \mid \gamma(x|C) \leq 1\}$$

og

$$\operatorname{int} C = \{x \mid \gamma(x|C) \leq 1\}\,.$$

($\gamma(\cdot|C)$ kaldes *gauge-funktionen* for C).

Bestem $\gamma(\cdot|K)$ hvor $K \subseteq \mathbb{R}^n$ er den lukkede kugle $\overline{K(0,1)}$.

OPGAVE 46. Find Gauge-funktionerne for C og D givet ved

$$C = \{(x_1, \cdots, x_n) \in \mathbb{R}^n \mid |x_i| \leq 1\,,\ i = 1, \cdots, n\}$$
$$D = \{(x_1, \cdots, x_n) \in \mathbb{R}^n \mid |x_1| + \cdots + |x_n| \leq 1\}.$$

OPGAVE 47. Lad C være en ikke-tom konveks mængde, så at

$$\{\lambda c \mid \lambda > 0\,,\ c \in C\} = \mathbb{R}^n\,.$$

Vis, at $0 \in \operatorname{int}(C)$.

2. Konvekse funktioner

OPGAVE 1. Betragt følgende udvidede reelle funktioner på $\mathbb{R}$:

(1) $f_1(x) = e^x, \ x \in \mathbb{R}$.

(2) $f_2(x) = \begin{cases} e^x, & x \leq 0, \\ +\infty, & x > 0. \end{cases}$

(3) $f_3(x) = \begin{cases} e^x, & x < 0, \\ +\infty & , x \geq 0. \end{cases}$

(4) $f_4(x) = \begin{cases} e^x, & x < 0, \\ 2, & x = 0 \\ +\infty, & x > 0. \end{cases}$

(5) $f_5(x) = \begin{cases} e^x, & x < 0, \\ \frac{1}{2}, & x = 0, \\ +\infty, & x > 0. \end{cases}$

(6) $f_6(x) = \begin{cases} e^x, & x \leq 0, \\ 1, & x > 0. \end{cases}$

(7) $f_7(x) = \begin{cases} e^x, & x \leq 0, \\ 2, & x > 0. \end{cases}.$

(8) $f_8(x) = \begin{cases} e^x, & x \leq 0, \\ x+2, & x > 0. \end{cases}$

(9) $f_9(x) = \begin{cases} e^x, & x \leq 0, \\ x+1, & x > 0. \end{cases}$

(10) $f_{10}(x) = \begin{cases} e^x, & x \leq 0, \\ x, & x > 0. \end{cases}$

Hvilke af funktionerne er konvekse?

OPGAVE 2. Lad $f_{\alpha,\beta} : \mathbb{R} \to \mathbf{R}$ være givet ved

$$f_{\alpha,\beta}(x) = \begin{cases} e^x, & x \leq 0, \\ \alpha x + \beta, & x > 0. \end{cases}$$

Bestem de $(\alpha, \beta) \in \mathbb{R}^2$ for hvilke $f_{\alpha,\beta}$ er konveks.

OPGAVE 3. Lad f være en konveks funktion på intervallet $[0, 1]$. Vis ved et eksempel, at f ikke nødvendigvis er kontinuert på hele intervallet.

OPGAVE 4. Vis, at $x \to \log(1 + e^x)$ er konveks på $\mathbb{R}$.

OPGAVE 5. Lad $f : [a, b] \to \mathbb{R}$ være en kontinuert funktion, der er konveks på $]a, b[$.
Vis, at f er konveks på $[a, b]$.

OPGAVE 6. Lad

$$f(x) = -\sqrt{1 - x^2}\,, \quad |x| \leq 1\,.$$

Vis, at f er konveks på $[-1, 1]$ og tegn grafen for f.

OPGAVE 7. Lad $0 \leq a < b \leq d$. Vis, at f er defineret ved

$$f(\lambda) = \log\left(\frac{1-e^{-\lambda d}}{e^{-\lambda a}-e^{-\lambda b}}\right), \quad \lambda > 0$$

er konveks på $]0,+\infty[$:

(Denne opgave er måske lidt drilagtig, – men den er et specialtilfælde af en funktion der er forekommet i praksis, – nemlig i Esben Høg's speciale i statistik, sommeren 1982).

OPGAVE 8. Lad I være et interval på den reelle akse og $f : I \to \mathbb{R}$ en konveks funktion.

Vis, at

$$f(\lambda_1 x_1 + \cdots + \lambda_k x_k) \leq \lambda_1 f(x_1) + \cdots + \lambda_k f(x_k)$$

når $x_1, \cdots, x_k \in I$, $\lambda_1, \cdots, \lambda_k \in [0,1]$ og $\lambda_1 + \cdots + \lambda_k = 1$.

Denne ulighed kaldes Jensen's ulighed efter en dansker J.L.W.V. Jensen, som viste en række klassiske uligheder som specialtilfælde af ovenstående. Resultatet er beskrevet i afhandlingerne:

J.L.W.V. Jensen:

Om konvekse funktioner og uligheder mellem middelværdier.
Nyt Tidsskrift Math. B 16 (1905), p. 49–68.

J.L.W.V. Jensen:

Sur les fonctions convexe et les inegalités entre les valeurs moyennes,
Acta Math. 30 (1906), p. 175–193.

OPGAVE 9. Lad $f : \mathbb{R} \to \mathbb{R}$ være en kontinuert afbildning. Gælder følgende:

1. $L \subseteq \mathbb{R}$, L begrænset $\Rightarrow f(L)$ begrænset?
2. $L \subseteq \mathbb{R}$, L lukket $\Rightarrow f(L)$ lukket?
3. $L \subseteq \mathbb{R}$, L lukket og begrænset $\Rightarrow f(L)$ lukket og begrænset?

OPGAVE 10. Lad $A : \mathbb{R}^n \to \mathbb{R}^m$ være en lineær afbildning. Gælder følgende:

1. $L \subseteq \mathbb{R}^n$, L begrænset $\Rightarrow A(L)$ begrænset?
2. $L \subseteq \mathbb{R}^n$, L lukket $\Rightarrow A(L)$ lukket?
3. $L \subseteq \mathbb{R}^n$, L kompakt $\Rightarrow A(L)$ kompakt?

OPGAVE 11. Lad A, B være ikke-tomme delmængder af $]-\infty, \infty]$. Vis, at

$$\sup(A+B) = \sup A + \sup B$$
$$\sup A = -\inf(-A).$$

(Husk på, at $\sup A$ er karakteriseret ved

1. $a \leq \sup A$, $\forall a \in A$,
2. $a \leq k$, $\forall a \in A \Rightarrow \sup A \leq k$).

Lad $f : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}$ være begrænsede funktioner. Vis, at

$$(*) \qquad \sup\{f(x)+g(x) \mid x \in \mathbb{R}\} \leq \sup\{f(x) \mid x \in \mathbb{R}\} + \sup\{g(x) \mid x \in \mathbb{R}\}.$$

Giv et eksempel, der viser, at man kan have $<$ i $(*)$.

OPGAVE 12. Vis, at $f, g : \mathbb{R}^n \to \mathbb{R}$ er konvekse, hvor

$$
\begin{aligned}
f(x_1, \cdots, x_n) &= \max\{|x_1|, \cdots, |x_n|\} \\
g(x_1, \cdots, x_n) &= |x_1| + \cdots + |x_n| \,.
\end{aligned}
$$

OPGAVE 13. Lad $\mathcal{C} \subseteq \mathbb{R}^n$ være konveks og $f : \mathcal{C} \to \mathbb{R}$ være en konveks funktion. Antag, at der findes $x_0 \in \operatorname{ri} C$, så at

$$
f(x_0) = \sup_{x \in C} f(x) \,.
$$

Vis, at f er konstant på C.

OPGAVE 14. Lad $f : \mathbb{R}^n \to]-\infty, \infty]$ være givet ved

$$
f(x) = f(x_1, \cdots, x_n) = \begin{cases} -(x_1 \cdot x_2 \ldots x_n)^{1/n}, & \text{hvis} \quad x_i > 0 \,, \forall i \\ \infty \,, & \text{ellers} \,. \end{cases}
$$

Vis, at f er konveks ved at vise følgende:

1. Hvis $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{R}$, så er

$$
(\alpha_1 + \cdots + \alpha_n)^2 \leq n(\alpha_1^2 + \cdots + \alpha_n^2) \,.
$$

2. Hvis $A = (\alpha_{ij})_1^n = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_1^n$, så er

$$
\langle Ay, y \rangle = \frac{1}{n^2} f(x) \cdot \left[\left(\sum_{j=1}^{n} \frac{y_j}{x_j} \right)^2 - n \sum_{j=1}^{n} \left(\frac{y_j}{x_j} \right)^2 \right]
$$

   når $x_1 > 0$, $x_2 > 0, \ldots, x_n > 0$.
3. f er konveks.

OPGAVE 15. Vis, at følgende funktioner er konvekse:

1.
$$
f(x, y) = \begin{cases} \log(e^x + e^y) \,, & x^2 + y^2 < 1 \\ \infty \,, & x^2 + y^2 \geq 1 \,. \end{cases}
$$

2.
$$
g(x, y, z) = \begin{cases} \frac{1}{x+y} + e^{x-z} \,, & x > 0 \,, \, y > 0 \,, \, z \in \mathbb{R} \\ \infty \,, & \text{ellers} \,. \end{cases}
$$

OPGAVE 16.

Hvis f er en konveks funktion defineret på en konveks mængde $\mathcal{C}$ og x_0 er et lokalt minimumspunkt for f, så er x_0 et globalt minimumspunkt for f, dvs. hvis der findes en omegn U om x_0, så at

$$
f(x_0) \leq f(x) \,, \quad \forall x \in U
$$

så er

$$f(x_0) \leq f(x)\,, \quad \forall x \in \mathbb{R}^n\,.$$

OPGAVE 17. (Eksamen, Sommeren 1980, Opgave 5).
Med $f : \mathbb{R}^2 \to \mathbb{R}$ betegnes den ved

$$f(x) = |x_1 - x_2|\,, \quad \forall x = (x_1, x_2) \in \mathbb{R}^2$$

definerede funktion.

1. Vis, at f er konveks.
2. Bestem subdifferentialet $\partial f(a)$ i ethvert punkt $a = (a_1, a_2) \in \mathbb{R}^2$.

OPGAVE 18. Find subdifferentialet for f i ethvert punkt for følgende funktioner:

1. $f(x) = \langle Ax, x\rangle$, $x \in \mathbb{R}^n$, – hvor A er en positiv semidefinit, symmetrisk $n \times n$ matrix.
2. $f(x, y) = \sqrt{x^2 + y^2}$, $(x, y) \in \mathbb{R}^2$.

OPGAVE 19. (Eksamen, Sommeren 1982, Opgave 3).
Med $f : \mathbb{R}^2 \to \mathbb{R}$ betegnes den ved

$$f(x) = |x_1 + 2x_2|\,, \quad x = (x_1, x_2) \in \mathbb{R}^2$$

definerede funktion.

1. Vis, at f er konveks.
2. Bestem subdifferentialet $\partial f(a)$ for ethvert $a = (a_1, a_2) \in \mathbb{R}^2$.

OPGAVE 20. (Eksamen, Vinteren 80/81, Opgave 3).
Med a og b betegnes to indbyrdes forskellige punkter af $\mathbb{R}^n$, og med $f : \mathbb{R}^n \to \mathbb{R}$ den ved

$$f(x) = |x - a| + |x - b|\,, \quad \forall x \in \mathbb{R}^n$$

definerede funktion.

1. Gør rede for, at f er konveks.
2. Find subdifferentialet $\partial f(a)$.
3. Gør rede for, at f har en mindste værdi γ og bestem γ samt minimumsmængden $\{z \in \mathbb{R}^n \mid f(z) = \gamma\}$ (man kan f.eks. benytte, at for ethvert $x \in \mathbb{R}^n$ har $f(x)$ en simpel geometrisk betydning)?

OPGAVE 21. (Eksamen, Sommeren 1981, Opgave 4).
Med $f : \mathbb{R}^3 \to \mathbb{R}$ betegnes den ved

$$f(x) = |x_1 + x_2|\,, \quad \forall x = (x_1, x_2, x_3) \in \mathbb{R}^3$$

definerede funktion.

1. Vis, at f er konveks.
2. Bestem subdifferentialet $\partial f(a)$ af f i ethvert punkt $a = (a_1, a_2, a_3) \in \mathbb{R}^3$.

OPGAVE 22. (Eksamen, Sommeren 1978, Opgave 4).
Gør rede for, at den ved

$$f(x) = |x| + \langle x, x\rangle\,, \quad \forall x \in \mathbb{R}^n$$

definerede funktion $f : \mathbb{R}^n \to \mathbb{R}$ er konveks, og bestem subdifferentialerne $\partial f(0)$ og $\partial f(a)$, hvor $a = (1, 1, \cdots, 1)$.

OPGAVE 23. (Eksamen, Vinteren 83/84, Nr. 1)

1. Lad $f : \mathbb{R}^2 \to [0, \infty[$ være konveks. Vis, at funktionen h givet ved
$$h(x, y) = f(x, y) + f(x, y)^2\,, \quad \forall (x, y) \in \mathbb{R}^2$$
er konveks.
2. Lad g være den konvekse funktion på $\mathbb{R}^2$:
$$g(x, y) = |x| + |y| + (|x| + |y|)^2\,, \quad \forall (x, y) \in \mathbb{R}^2\,.$$
Find subdifferentialet $\partial g(0, 0)$ af g i punktet $(0, 0)$.

OPGAVE 24. (Eksamen, Sommeren 1979, Nr. 1).
Med $f : \mathbb{R} \to]-\infty, \infty[$ betegnes en konveks funktion. Undersøg, om der gælder: Hvis

$$f(\lambda) \to 0 \quad \text{for} \quad \lambda \to \infty$$

da er

$$f(\lambda) \geq 0\,, \quad \forall \lambda \in \mathbb{R}\,.$$

OPGAVE 25. Definer funksjonen F for alle $x > 0$ ved

$$f(x) = \int_0^x e^{xt^2} dt$$

(a) Finn uttrykk for $F'(x)$ og $F''(x)$.
(b) Påvis at F er strengt konveks over $(0, \infty)$.

OPGAVE 26.

(a) Hvis $f(x_1, \cdots, x_n)$ er konkav, for hvilke verdier av konstanterne a og b er $af(x_1, \cdots, x_n) + b$ konkav?
(b) Hvis $f(x_1, \cdots, x_n)$ er konkav og bare antar positive verdier, afgjør om funksjonerne
$$h(x_1, \cdots, x_n) = \ln f(x_1, \cdots, x_n) \quad \text{og} \quad g(x_1, \cdots, x_n) = e^{f(x_1, \cdots, x_n)}$$
er konkave/kvasikonkave.

OPGAVE 27. Lad f være en konveks funktion defineret på en konveks delmængde $\mathcal{C} \subseteq \mathbb{R}^n$ som har ikke tomt indre og lad x_0 være et indre punkt i $\mathcal{C}$. Vi skal angive et bevis for at f er kontinuert i x_0. Lad $\varepsilon > 0$:

1. Vis ved at separere $(x_0, f(x_0) - \frac{\varepsilon}{2})$ fra $\operatorname{epi} f$ at der findes en affin funktion a med $a \leq f$, så at $a(x_0) > f(x_0) - \varepsilon$.

Slut heraf, at der findes en omegn U af x_0, så at

$$y \in U \Rightarrow f(y) > f(x_0) - \varepsilon$$

.

2. Gør rede for, at der for $i = 1, \cdots, n$ findes $\delta_i > 0$, så at

$$f(x_0 + \lambda e_i) < f(x_0) + \varepsilon\,, \quad \text{når} \quad -\delta_i \le \lambda \le \delta_i\,.$$

   (e_i betegner den i'te naturlige basisvektor).
   Gør rede for, at

$$0 \text{ er et indre punkt af } (\mathrm{conv}\{\pm\,\delta_1 e_1, \cdots, \pm\,\delta_n e_n\})$$

   og slut heraf, at der findes en omegn om $x_0 : V$, så at

$$y \in V \Rightarrow f(y) < f(x_0) + \varepsilon$$

   .

3. Vis, at f er kontinuert i punktet x_0.

OPGAVE 28.
Lad $f : \mathbb{R}^n \to \mathbb{R}$ være en konveks funktion.
Vis, at f er (det punktvise) supremum af sine affine minoranter.

3. Dualitet og separation

OPGAVE 1. For en vilkårlig mængde M i $\mathbb{R}^n$ sætter vi

$$\operatorname{polar}(M) := \{y \in \mathbb{R}^n \mid \forall x \in M : x \cdot y \leq 1\}.$$

Mængden polar(M) kaldes M's *polare* mængde.

1° Vis, at $\operatorname{polar}(M) = \bigcap_{x \in M} J(x, 1)$

2° Vis, at $y \in \operatorname{polar}(M) \Leftrightarrow M \subseteq J(y, 1)$.

3° Vis, at polar(M) er en afsluttet konveks mængde.

4° Vis, at $\operatorname{polar}((\{0\} \cup \operatorname{cl}(\operatorname{conv} M))) = \operatorname{polar}(M)$.

5° Vis, at $M_1 \subseteq M_2 \Rightarrow \operatorname{polar}(M_1) \supseteq \operatorname{polar}(M_2)$.

6° Vis, at $\operatorname{polar}(K(0, r)) = \overline{K(0, \frac{1}{r})}$.

7° Vis, at hvis M er begrænset, så er 0 et indre punkt i polar(M).

8° Vis, at hvis 0 er et indre punkt i M, så er polar(M) begrænset.

9° Vis, at hvis P er en konveks polytop, så er polar(P) et konvekst polyeder.

10° Vis, at $\operatorname{polar}(L) = L^{\perp}$ når L er et lineært underrum.

11° Vis, at polar(K) er K's normalkegle når K er en konveks kegle.

OPGAVE 2. Tegn den polare mængde (jf. Opgave 1) til følgende mængder i $\mathbb{R}^2$:

$$\begin{aligned}
M_1 &= \{(1,1), (-1,1), (-1,-1), (1,-1)\}.\\
M_2 &= \operatorname{conv} M_1.\\
M_3 &= \{(1,0), (0,1), (-1,0), (0,-1)\}.\\
M_4 &= \operatorname{conv} M_3.\\
M_5 &= \{(x_1, x_2) \mid x_1 \geq 0, x_2 \geq 0\}.\\
M_6 &= \{(x_1, x_2) \mid x_2 \geq 0\}.\\
M_7 &= \{(x_1, x_2) \mid 2x_1 + x_2 = 0\}.\\
M_8 &= \{(1,1)\}.\\
M_9 &= [(0,0), (1,1)].\\
M_{10} &= \{(x_1, x_2) \mid x_1 - x_2 \leq 2, x_1 \geq -2, x_2 \geq -1\}.
\end{aligned}$$

OPGAVE 3. I Opgave 1 definerede vi M's polare mængde polar(M) ved

$$\operatorname{polar}(M) := \{y \in \mathbb{R}^n \mid \forall x \in M : x \cdot y \leq 1\}.$$

Ved M's *bipolare* mængde forstås mængden polar(polar(M)).

1° Vis, at $\operatorname{polar}(\operatorname{polar}(M)) = \bigcap\{J(y, 1) \mid M \subseteq J(y, 1)\}$.

2° Vis, at $\operatorname{polar}(\operatorname{polar}(M)) = \operatorname{cl}\operatorname{conv}(\{0\} \cup M)$. -Dette resultat kaldes *Bipolarsætningen.*

OPGAVE 4. Lad $x^* = (x_1^*, \cdots, x_n^*) \in \mathbb{R}^n$ og lad

$$\mathbb{R}^n_+ = \{(x_1, \cdots, x_n) \in \mathbb{R}^n \mid x_i \geq 0\,,\ i = 1, \cdots, n\}\,.$$

Antag, at afbildningen $t \to \langle x^*, t\rangle$ er nedad begrænset på $\mathbb{R}^n_+$.

Vis, at $x^* \in \mathbb{R}^n_+$.

OPGAVE 5. Lad

$$B_1 = \{(x, y) \in \mathbb{R}^2 \mid 0 < x, 0 < y\,,\ 1 \leq xy\}$$
$$B_2 = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x\,,\ 0 \leq y\}\,.$$

Lad $x^* = (x_1^*, x_2^*) \in \mathbb{R}^2$. Angiv $\sup\limits_{c \in B_i} \langle c, x^*\rangle$ og $\inf\limits_{c \in B_i} \langle c, x^*\rangle$ for $i = 1, 2$.

(Det kan være nødvendigt at graduere svaret efter egenskaber ved x^*).

OPGAVE 6. Lad H være en hyperplan i $\mathbb{R}^n$ og K en ikke-tom konveks mængde, så at

$$K \cap H = \varnothing\,.$$

Vis, at K er indeholdt i et af de ved H bestemte åbne halvrum.

OPGAVE 7. Lad $b, d \in \mathbb{R}^n$ være forskellige fra nul. Lad $\beta, \gamma \in \mathbb{R}$. Lad H og M betegne hyperplanerne

$$H = \{x \in \mathbb{R}^n \mid \langle x, b\rangle = \beta\}$$
$$M = \{x \in \mathbb{R}^n \mid \langle x, d\rangle = \gamma\}\,.$$

Vis $H = M$ hvis og kun hvis der findes $\alpha \in \mathbb{R}$, så at $b = \alpha d$ og $\beta = \alpha\gamma$.

(Vink til den vanskelige implikation: Hvis $\mathcal{U}$ er det parallelle underrum til H, så er $\mathcal{U}^\perp = \operatorname{span}(b)$, – og analogt for M.)

OPGAVE 8. Lad som sædvanlig B betegne den lukkede enhedskugle i $\mathbb{R}^2$. Lad $(x, y) \in \mathbb{R}^2$. Find den minimale afstand fra (x, y) til punkter i følgende mængder:

$$B$$
$$a + \varepsilon B\,,\ (a \in \mathbb{R}^2\,,\ \varepsilon > 0)$$
$$\{(s, t) \in \mathbb{R}^2 \mid |s| \leq 1\,,\ |t| \leq 1\}$$

og angiv det punkt, hvori den minimale afstand realiseres.

OPGAVE 9. Hvis hyperplanen H separerer mængderne S og T i $\mathbb{R}^n$, så separerer H også $\operatorname{conv}(S)$ og $\operatorname{conv}(T)$.

OPGAVE 10.

1. Hvis B som sædvanlig betegner den lukkede enhedskugle i $\mathbb{R}^n$ og D en delmængde af $\mathbb{R}^n$, så gælder for alle $b \in \mathbb{R}^n$, $\varepsilon > 0$:

$$\sup\{\langle x, b\rangle \mid x \in D + \varepsilon B\} = \sup\{\langle y, b\rangle \mid y \in D\} + \varepsilon\|b\|\,.$$

2. Hvis C_1 og C_2 er konvekse mængder i $\mathbb{R}^n$, hvorom der gælder, at der findes et $\varepsilon > 0$, så at $C_1 + \varepsilon B$ og $C_2 + \varepsilon B$ kan separeres, så kan C_1 og C_2 separeres stærkt.
3. Hvis C_1 og C_2 er konvekse mængder i $\mathbb{R}^n$, så at C_1 og C_2 kan separeres stærkt, så findes $\varepsilon > 0$, så at $C_1 + \varepsilon B$ og $C_2 + \varepsilon B$ kan separeres.

OPGAVE 11. Lad $S \subseteq \mathbb{R}^n$ være en begrænset mængde. Lad $x^* \in \mathbb{R}^n$. Vis, at

$$\sup\{\langle x^*, s\rangle \mid s \in S\} = \sup\{\langle x^*, s\rangle \mid s \in c\ell(S)\}$$
$$\int\{\langle x^*, s\rangle \mid s \in S\} = \inf\{\langle x^*, s\rangle \mid s \in c\ell(S)\}\,.$$

OPGAVE 12. Lad C og D være konvekse mængder i $\mathbb{R}^n$. Er følgende påstande korrekte:

1. Hvis C og D kan separeres egentligt, så kan $\overline{C}$ og $\overline{D}$ separeres egentligt ?
2. Hvis C og D kan separeres stærkt, så kan $\overline{C}$ og $\overline{D}$ separeres stærkt ?

OPGAVE 13. Lad C_1 og C_2 være ikke-tomme konvekse mængder i $\mathbb{R}^n$. Antag, at $0 \notin \overline{(C_1 - C_2)}$.

Vis, at så findes der en hyperplan, der separerer C_1 og C_2 stærkt.

OPGAVE 14. Lad $K \subseteq \mathbb{R}^n$ være en kegle, dvs.

$$\lambda > 0\,,\ k \in K \Rightarrow \lambda k \in K\,.$$

Lad $b \in \mathbb{R}^n$ og antag, at der findes $\alpha \in \mathbb{R}$, så at

$$\langle k, b\rangle \leq \alpha\,, \quad \text{for alle } k \in K\,.$$

Så gælder, at

$$\langle k, b\rangle \leq 0 \quad \text{for alle } k \in K\,.$$

OPGAVE 15. (Farkas lemma).

Lad $a_0, a_1, \cdots, a_m \in \mathbb{R}^n$. Så er følgende udsagn ækvivalente:

1. Der findes $\lambda_1, \cdots, \lambda_m \in \mathbb{R}$, alle ikke-negative så at $a_0 = \lambda_1 a_1 + \cdots + \lambda_m a_m$.
2. Uligheden $\langle a_0, x\rangle \leq 0$ er en konsekvens af ulighederne $\langle a_i, x\rangle \leq 0$, $i = 1, \cdots, m$, dvs.

$$\langle x, a_i\rangle \leq 0\,, \quad i = 1, \cdots, m \Rightarrow \langle x, a_0\rangle \leq 0\,.$$

(Vink: Den ene vej kan f.eks. vises ved at bruge opgave 14 ovenfor og separere a_0 fra mængden

$$\{\lambda_1 a_1 + \cdots + \lambda_m a_m \mid \lambda_i \geq 0\,,\ i = 1, \cdots, m\}$$

der er lukket og konveks ifølge opgave K39.

OPGAVE 16. Lad A være en reel $m \times n$ matrix. Vis, at hvis billedmængden $A(\mathbb{R}^n)$ og den “positive orthant”

$$\{(x_1, \cdots, x_m) \in \mathbb{R}^m \mid x_i > 0\,,\ i = 1, \cdots, m\}$$

er disjunkte, så findes der en vektor $p \in \mathbb{R}^m \setminus \{0\}$, så at alle p’s koordinater er ikke-negative, og så at $A^t p = 0$ (hvor A^t er den transponerede matrix til A).

Vis derefter, at præcis ét af følgende udsagn er sandt:

1. Der findes $y \in \mathbb{R}^n$, så at samtlige koordinater i Ay er positive.
2. Der findes en vektor $x \in \mathbb{R}^m \setminus \{0\}$, så at x har ikke-negative koordinater, og så at $A^t x = 0$.

OPGAVE 17. Lad L være et underrum af $\mathbb{R}^n$. Vis, at enten indeholder L en vektor med lutter positive (dvs > 0 her) koordinater, eller også indeholder $L^\perp$ en vektor, som er forskellig fra nul og som har lutter ikke-negative koordinater.

Tolk resultatet i $\mathbb{R}^2$ og $\mathbb{R}^3$.

OPGAVE 18. Lad A være en $m \times n$ matrix og $b \in \mathbb{R}^m$. Vis, at præcis et af følgende udsagn er sandt:

1. Der findes $x \in \mathbb{R}^n$ med $x \geq 0$, så at $Ax = b$.
2. Der findes $y \in \mathbb{R}^m$ med $A^t y \geq 0$, $\langle b, y \rangle < 0$,

(hvor A^t betegner den transponerede matrix til A, – og $z \geq 0$ betyder, at alle z's koordinater er ikke-negative).

OPGAVE 19. Lad A være en reel $m \times n$ matrix og $b \in \mathbb{R}^m$.

1. Følgende udsagn er ækvivalente:
   (i) $\exists x \geq 0 : Ax \leq b$
   (ii) $\exists x \geq 0, \exists u \geq 0 : Ax + u = \{A, I\}\binom{x}{u} = b$.
2. Der indtræffer ét og kun ét af tilfældene:
   (i) $\exists x \geq 0 : Ax \leq b$
   (ii) $\exists z \geq 0 : A^t z \geq 0$, $\langle b, z \rangle < 0$.

(Vink: Brug ovenstående opgave 18 på $\{A, I\} : m \times (m+n)$–matrix).

OPGAVE 20. (Dualitetssætningen i lineær programmering).

Lad A være en $m \times n$ matrix, $b \in \mathbb{R}^m$ og $c \in \mathbb{R}^n$. Betragt problemerne:

(P) Maksimer $\langle c, x \rangle$ under bibetingelserne $Ax \leq b$, $0 \leq x$.
(P*) Minimer $\langle b, y \rangle$ under bibetingelserne $A^t y \geq c$, $0 \leq y$.

Vi skal nu vise, at $\max(P) = \min(P^*)$ under forudsætning af, at der findes $x_0 \in \mathbb{R}^n$ og $y_0 \in \mathbb{R}^m$, så at

$$\begin{aligned} 0 \leq x_0\,, \quad & Ax_0 \leq b \\ 0 \leq y_0\,, \quad & c \leq A^t y_0\,. \end{aligned}$$

1. Vis, at $\sup(P) \leq \inf(P^*)$.
2. Gør rede for følgende: For at vise resten er det nok at vise, at der findes $x \geq 0$, $y \geq 0$ med $Ax \leq b$, $A^t y \geq c$ og $\langle c, x \rangle = \langle b, y \rangle$.
3. Lad
$$B = \left\{ \begin{matrix} A & 0 \\ 0 & -A^t \\ -c & b \end{matrix} \right\} \quad d = \left\{ \begin{matrix} b \\ -c \\ 0 \end{matrix} \right\}.$$
   Vis ved hjælp af opgave DS19, at hvis der ikke findes x og y som opfylder kravene i 2., så findes $z \geq 0$, så at $B^t z \geq 0$, $\langle d, z \rangle < 0$.
4. Lad z som i 3.: $z = (z_1, z_2, \theta)$, $\theta \in \mathbb{R}$. Vis, at $\theta > 0$ – brug så $\frac{z_1}{\theta}$ og $\frac{z_2}{\theta}$ til at fuldende argumentet.

4. Optimering under Bibetingelser

OPGAVE 1. Betragt funktionerne $f_0 : \mathbf{R}^3 \to \mathbf{R}$ og $f_1 : \mathbf{R}^3 \to \mathbf{R}$ givet ved

$$f_0(x, y, z) = x + y + z,$$
$$f_1(x, y, z) = x^2 + y^2 + z^2 - 25.$$

1° Undersøg, om f_0 har en størsteværdi og/eller mindsteværdi under bibetingelsen $f_1 = 0$.
2° Opstil og løs Lagrange ligningerne.
3° Anvend (om muligt) Lagranges sætning til at finde de eventuelle globale maksimumsspunkter/minimumsspunkter under bibetingelsen $f_1 = 0$.
4° Overvej, om opgaven kunne være løst uden anvendelse af Lagranges sætning.

OPGAVE 2. Udfør samme program som i Opgave 1 på hvert af nedenstående sæt af funktioner. I $6° - 8°$ drejer det sig om maksimering og minimering af f_0 under bibetingelserne $f_1 = 0$, $f_2 = 0$.

1° $f_0(x, y, z) = x^2 + y^2 + z^2$,
$f_1(x, y, z) = x + y + z - 25$.
2° $f_0(x, y, z) = x^2 + y^2 + z^2$,
$f_1(x, y, z) = x - y + z - 1$.
3° $f_0(x, y) = y^2 - 4xy + 4x^2$,
$f_1(x, y) = x^2 + y^2 - 1$.
4° $f_0(x, y) = 2x^2 + xy - y^2 + y$,
$f_1(x, y) = 2x + 3y - 1$.
5° $f_0(x, y, z) = x + 2y - 3z$,
$f_1(x, y, z) = 4x^2 + y^2 - z$.
6° $f_0(x, y, z) = x^2 + y^2 + z^2$,
$f_1(x, y, z) = x - y - 1$,
$f_2(x, y, z) = y^2 - z^2 - 1$.
7° $f_0(x, y, z) = z - x^2 - y^2$,
$f_1(x, y, z) = x + y + z - 1$,
$f_2(x, y, z) = x^2 + y^2 - 4$.
8° $f_0(x_1, x_2, x_3, x_4) = x_1x_2x_3x_4$,
$f_1(x_1, x_2, x_3, x_4) = x_1 - x_3 - 2$,
$f_2(x_1, x_2, x_3, x_4) = x_2^2 + x_4 - 4$.

OPGAVE 3. Der skal fremstilles et lukket kasse, som skal rumme 2 m^3. Prisen pr. kvadratmeter af de materialer, som kassen skal laves af, er 1 kr. for siderne, 2 kr. for bunden og 1,5 kr. for toppen. Hvilke dimensioner af kassen vil gøre prisen mindst mulig?

OPGAVE 4. Diskutér fortolkningen af Lagrange multiplikatorerne for eksemplerne i Opgave 3.

OPGAVE 5. Lad C være en ikke-tom åben konveks delmængde af $\mathbf{R}^n$, og lad f være en C^2 funktion på C. Man kan bevise, at f er en konveks funktion, hvis og kun hvis den symmetriske matrix

$$\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{i,j=1,\dots,n}$$

er positivt semidefinit for alle $x \in C$, dvs. den tilhørende kvadratiske form er ≥ 0.

Det burde være bevist i Matematik 1, at en symmetrisk matrix er positivt semidefinit, hvis og kun hvis alle dens egenværdier er ≥ 0. Benyt dette til at undersøge, om følgende funktioner er konvekse:

1° $f(x_1, x_2) = e^{x_1 - x_2} + e^{x_2 - x_1}, \quad (x_1, x_2) \in \mathbf{R}^2$.
2° $f(x_1, x_2, x_3) = 2x_1^2 + x_2^2 - 2x_1x_2 - x_1 - x_2 + x_3, \quad (x_1, x_2, x_3) \in \mathbf{R}^3$.
3° $f(x_1, x_2) = x_1^2 + x_2^3, \quad (x_1, x_2) \in \mathbf{R}^2$.
4° $f(x_1, x_2) = x_1^2 + x_2^3, \quad (x_1, x_2) \in \overline{\mathbf{R}}_+^2$.

OPGAVE 6. Følgende eksempel er hentet fra en amerikansk lærebog. Diskutér eksemplet!

Consider the problem

$$\begin{aligned} \text{minimize} \quad & x_1x_2 + x_2x_3 + x_3x_1 \\ \text{subject to} \quad & x_1 + x_2 + x_3 = 3. \end{aligned}$$

The necessary conditions become

$$\begin{aligned} x_2 + x_3 + \lambda &= 0 \\ x_1 \qquad + x_3 + \lambda &= 0 \\ x_1 + x_2 \qquad + \lambda &= 0. \end{aligned}$$

The three equations together with the one constraint equation give four equations that can be solved for the four unknowns x_1, x_2, x_3, λ. Solution yields $x_1 = x_2 = x_3 = 1,\ \lambda = -2$.

OPGAVE 7. En funktion $f : A \to \mathbf{R}$, hvor $A \subseteq \mathbf{R}^n$, siges at være *inf-kompakt*, dersom der findes et $a \in \mathbf{R}$, således at mængden

$$\{x \in A \mid f(x) \leq a\}$$

er ikke-tom og kompakt. Bevis, at hvis $f : A \to \mathbf{R}$ er kontinuert og inf-kompakt, så har f en mindsteværdi på A.

OPGAVE 8. Definer funksjonen f for alle x, y ved

$$f(x, y) = 3 \sin 2x + 3y - 3y^2 + 8$$

(a) Vis at f er konkav over rektanglet $R = \{(x, y) : 0 < x < \pi/2,\ 9 < y < 1\}$.
(b) Finn maksimum av f over R.
(c) Vis at f har minst ett lokalt maksimumspunkt utenfor R.

OPGAVE 9. Minimér funktionen

$$f_0(x_1, x_2) = e^{x_1 - x_2} + e^{x_2 - x_1}, \quad (x_1, x_2) \in \mathbf{R}^2,$$

under bibetingelserne

$$\begin{aligned} f_1(x_1, x_2) &= x_1^2 - x_2 + 1 \leq 0, \\ f_2(x_1, x_2) &= 2 - x_2 \leq 0. \end{aligned}$$

OPGAVE 10. Minimér funktionen

$$f_0(x_1, x_2, x_3) = 2x_1^2 + x_2^2 - 2x_1x_2 - x_1 - x_2 + x_3, \quad (x_1, x_2, x_3) \in \mathbf{R}^3,$$

under bibetingelserne

$$0 \le x_1, \quad 0 \le x_2 \le 1, \quad x_1^2 + x_2^2 \le x_3 + 1.$$

OPGAVE 11. Minimér funktionen

$$f_0(x_1, x_2, x_3) = x_1^2 + x_2^2 - x_1x_2 + x_3, \quad (x_1, x_2, x_3) \in \mathbb{R}^3,$$

under bibetingelserne

$$0 \le x_1, \quad 1 \le x_2, \quad x_1^2 + x_2 - x_3 \le 2.$$

OPGAVE 12. La f være definert for alle (x, y) ved

$$f(x, y) = 2 - (x-1)^2 - e^{y^2} \tag{1}$$

(a) Beregn Hesse-matrisen til f og påvis at f er konkav.
(b) Finn den største verdien f antar.
(c) Betrakt det ikke-lineære programmeringsproblemet

$$\max f(x, y) \ \text{ når } \ x^2 + y^2 \le a \ \ (a \text{ positiv konstant}) \tag{2}$$

Still opp Kuhn-Tucker-betingelserne for løsning av dette problemet.
(d) Finn løsningen av problemet (2).

OPGAVE 13. La D være mengden av punkter (x, y) i xy-planet som er slik at $-1 < x < 1$ og $-1 < y < 1$, og la

$$f(x, y) = \frac{1}{12}(x-y)^4 - (x-y)^2 - (x+y)^2 .$$

(a) Undersøk om f er konveks eller konkav i D.
(b) Finn maksimum av f i D.

OPGAVE 14. Betrakt problemet

$$(*) \qquad \text{maksimer } (2x + y) \ \text{ når } \begin{cases} (x+1)^2 + y^2 \le 4 \\ x^2 + (y+1)^2 \le 4 \end{cases} \quad x \ge 0\,,\ y \ge 0$$

(a) La S være mengden av alle (x, y) som tilfredsstiller alle de fire bibetingelsene. Skraver S i xy-planet, og tegn inn noen nivålinjer for $f(x, y) = 2x + y$.
(b) Løs problemet $(*)$ ved å bruke et geometrisk resonnement.
(c) Vis, at $(*)$ er et konkavt maksimeringsproblem. Still opp Kuhn-Tucker-betingelsene. Verifiser at det punktet (x_0, y_0) du fant i (b) tilfredstiller Kuhn-Tucker-betingelsene.
(d) Anta at betingelsen $x^2 + (y+1)^2 \le 4$ i $(*)$ erstattes med $x^2 + (y+1)^2 \le 4.1$. Anslå endringen i den maksimale verdien av $2x + y$.

OPGAVE 15. Betrakt problemet

$$\text{maksimer } (x_1 + x_2 + x_3 + x_4) \text{ når } \begin{cases} (1)\ \frac{1}{3}x_1 + \frac{1}{3}x_2 + \frac{1}{8}x_3 + \frac{1}{8}x_4 & \leq 3 \\ (2)\ x_1x_2x_3x_4 & \leq 144 \\ (3)\ x_1, x_2, x_3, x_4 \text{ alle} & \geq 0 \end{cases}$$

(a) Still opp Kuhn-Tucker-betingelsene for at $(x_1^0, x_2^0, x_3^0, x_4^0)$ skal løse dette problemet, idet du lar λ og μ være multiplikatorer tilordnet henholdsvis (1) og (2). (Bruk Setning 4.25. Føringsbetingelsen trenger du ikke undersøke.)
(b) Forutsett at problemet har løsning og finn denne. (Vink: Se på de to tilfellene (I): (2) er inaktiv, (II): (2) er aktiv.)
(c) Anta at høyresiden i bibetingelse (1) endres fra 3 til 3.1. Hva blir tilnærmet den tilhørende endringen i den maksimale verdien av $x_1+x_2+x_3+x_4$? Hva blir den tilsvarende virkningen av en liten endring i høyresiden i bibetingelse (2) ?

OPGAVE 16. Betrakt problemet

$$\text{maksimer } x^2ye^{-x-y} \text{ når } x \geq 1\,,\ y \geq 1\,,\ x + y \geq 4\,.$$

(a) Still opp Kuhn-Tucker-betingelsene for problemet. (En trenger ikke se på føringsbetingelsen.)
(b) Finn alle løsningene av disse betingelsene.
(c) (Kan sløyfes). Har du i (b) funnet maksimum i problemet ?

OPGAVE 17. Løs følgende problem med hjelp av Kuhn-Tucker betingelser:

$$\text{minimer } f(x, y) = e^{x+y} + e^y + 2x + y \text{ når } x \geq -1\,,\ y \geq -1\,,\ x + y \geq 0\,.$$

OPGAVE 18.

(a) La A være en symmetrisk $n \times n$ matrise med $|A| \neq 0$, la B være en $1 \times n$-matrise, og la X være en $n \times 1$-matrise. Betrakt uttstykket

$$(*) \qquad \left(X + \frac{1}{2}A^{-1}B'\right)' A \left(X + \frac{1}{2}A^{-1}B'\right) - \frac{1}{4}BA^{-1}B'$$

Multipliser ut uttrykket, og forenkle det mest mulig.
(b) Anta at A er symmetrisk og positiv definitt (dvs. at $Y'AY > 0$ for alle $n\times 1$-matriser $Y \neq 0$). Finn med hjelp av $(*)$ den matrisen X som minimerer uttrykket $X'AX + BX$.

OPGAVE 19. Betrakt problemet

$$\text{maks } \left[\ln(x^2 + 2y) - \frac{1}{2}x^2 - y\right] \text{ når } y \geq 2/x\,,\ x \geq 1\,,\ y \geq 1\,.$$

(a) Still opp Kuhn-Tucker-betingelsene for problemet. Se bort fra føringsbetingelsen.
(b) Løs problemet.

OPGAVE 20. Betragt problemet

$$\text{maks } (4z - x^2 - y^2 - z^2) \ \text{ når } \begin{cases} z & \leq xy \\ x^2 + y^2 + z^2 & \leq 3 \end{cases}$$

(a) Still opp Kuhn-Tucker-betingelsene for løsning av pro-blemet. Se bort fra føringsbetingelsene.
(b) Begrunn at maksimeringsproblemet har en løsning, og finn løsningen.
(c) Hvilken endring får en tilnærmet av maksimumsverdien av $4z - x^2 - y^2 - z^2$ hvis den første bibetingelsen endres til $z \leq xy + 0.1$?

OPGAVE 21. La f være en funksjon av to variable, gitt ved

$$f(x, y) = -x^4 - cx^2 + 6xy - 6y^2$$

der c er en konstant.

(a) For hvilke verdier av c vil f være konkav i hele planet ?

I resten av denne oppgaven skal vi studere problemet

$$(*) \qquad \text{maks}\,(-x^4 - y^4 - 4x^2 + 6xy - 6y^2 + ax + by) \ \text{ når } \ x + y^2 \leq 1 \text{ og } y \geq -1.$$

Her er a og b konstanter.

(b) Still opp Kuhn-Tucker-betingelsene for et punkt (x, y) skal løse $(*)$.
(c) Finn nødvendige og tilstrekkelige betingelser på a og b for at maksimumspunktet i $(*)$ skal være $(x, y) = (0, -1)$.

OPGAVE 22. La f og g være funksjoner av to variable, gitt ved

$$f(x, y) = ax + by - 6x^2 - 5xy - 5y^2\,, \quad g(x, y) = 3 - (x^2 + y^2)^2$$

der a og b er konstanter.

(a) Vis at både f og g er konkave i hele planet.

I resten av denne oppgaven skal vi studere problemet

$$(*) \qquad \text{maks}\,(f(x, y) + g(x, y)) \ \text{ når } \ x \geq 1 \text{ og } y \leq \sqrt{x}$$

(b) Still opp Kuhn-Tucker-betingelsene for at et punkt (x, y) skal løse $(*)$.
(c) Finn nødvendige og tilstrekkelige betingelser på a og b for at maksimumspunktet i $(*)$ skal være $(x, y) = (1, 0)$.
(d) La $a = 300$. For hvilken verdi av b vil $(*)$ ha løsningen $(x, y) = (4, 2)$?

OPGAVE 23. Bekrakt problemet

$$(*) \qquad \begin{array}{l} \text{maksimer } a \ln(z + 1) - z - 2x - y \\ \text{når } x \geq 0\,,\ y \geq 0\,,\ z \geq 0 \text{ og } z^2 \leq x + y \end{array}$$

der a er en positiv konstant.

(a) Still opp Kuhn-Tucker-betingelsene for løsning av $(*)$.
(b) Finn løsningene av Kuhn-Tucker-betingelsene. Se på tilfellene $a > 1$ og $a \leq 1$ hver for seg.
(c) Begrunn hvorfor det eller de punktene du fant i (b) virkelig gir maksimum i problemet $(*)$.

OPGAVE 24. Betragt problemet

$$(*) \qquad \text{maksimer } \; y - e^{-x} \; \text{ når } \; e^x + e^y \leq 6\,, \; x - y \leq 0$$

(a) Still opp Kuhn-Tucker-betingelsene for løsning av $(*)$.
(b) Finn løsningen av problemet.

OPGAVE 25. Betrakt problemet

$$\text{maksimer } xz + yz \; \text{ når } \; x^2 + y^2 + z^2 \leq 1$$

(a) Still opp de nødvendige Kuhn-Tucker-betingelsene for løsning av problemet.
(b) Finn løsningen av problemet.

OPGAVE 26. Lad $h : \mathbb{R}^2 \to \mathbb{R}$ betegne den konvekse funktion

$$h(x, y) = \max\{x + 2y\,,\, x - 2y\}\,, \quad (x, y) \in \mathbb{R}^2\,.$$

1. Find $\partial h((x, y))$, $\forall (x, y) \in \mathbb{R}^2$
2. Hvad er $\inf\limits_{\mathbb{R}^2} h$?

OPGAVE 27. (Eksamen, Vinteren 83/84, Opgave 2).
Lad

$$f(x, y, z) = x^2 + y^2 + z^2 + xz + yz + xy$$

for alle $(x, y, z) \in \mathbb{R}^3$.

1. Vis, at f er konveks på $\mathbb{R}^3$.
2. Lad A betegne den konvekse mængde

$$A = \{(x, y, z) \in \mathbb{R}^3 \mid 0 \leq x; y \leq 0; 1 \leq 2x + y + 4z\}\,.$$

Find minimum for f over mængden A og angiv samtlige punkter i A, hvori f antager sit minimum over A.

OPGAVE 28. Lad $g(x, y, z) = x^2 + 2y^2 + 3z^2$ og lad

$$C = \{(x, y, z) \mid x^2 + y^2 + z^2 \leq 4\}$$

Find den optimale værdi og samtlige optimale løsninger til programmet

$$(\mathrm{P}) \qquad \text{Minimer } g(\underline{x})$$

under betingelserne

$$\underline{x} \in C\,, \quad |x| + y - z \leq -1$$

hvor $\underline{x} = (x, y, z)$.

OPGAVE 29. Lad (P) være et standard minimeringsproblem givet ved $C, X \in \mathbb{R}^n$, $B \in \mathbb{R}^m$ og $A \in \mathbb{M}(m, n)$ som

$$\min C^t X \text{ under bibetingelserne } \; AX \geq B,\; X \geq 0.$$

Lad X være en mulig løsning til (P), og lad Y være en mulig løsning til det duale problem $(P^\star)$. Vis, at følgende to betingelser er ensbetydende:

(1) X er en optimal løsning til (P), og Y er en optimal løsning til $(P^\star)$.
(2) Der gælder

$$x_j > 0 \Rightarrow \sum_{i=1}^{m} a_{ij} y_i = c_j, \quad j = 1, \ldots, n,$$

og

$$y_i > 0 \Rightarrow \sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = 1, \ldots, m.$$

(Udsagnet (1) $\Leftrightarrow$ (2) kaldes sætningen om "complementary slackness". Slackness $\sim$ spillerum.)

Opgave 30. Lad (P) være et standard minimeringsproblem (se opgave 29 for formuleringen), og lad X være en mulig løsning til (P). Vis, at følgende to betingelser er ensbetydende:

(1) X er en optimal løsning til (P).
(2) Der findes Y, således at

(a) $$y_i \geq 0, \quad i = 1, \ldots, m,$$

(b) $$\sum_{i=1}^{m} a_{ij} y_i \leq c_j, \quad j = 1, \ldots, n,$$

(c) $$x_j > 0 \Rightarrow \sum_{i=1}^{m} a_{ij} y_i = c_j, \quad j = 1, \ldots, n,$$

(d) $$\sum_{j=1}^{n} a_{ij} x_j > b_i \Rightarrow y_i = 0, \quad i = 1, \ldots, m.$$

Gør endvidere rede for, at hvis Y opfylder (a)-(d) i (2), så er Y en optimal løsning til $(P^\star)$.

Opgave 31. Lad (P) være et standard minimeringsproblem med

$$C^t = (-18 \quad 7 \quad -12 \quad -5 \quad 0 \quad -8),$$
$$A = \begin{pmatrix} -2 & 6 & -2 & -7 & -3 & -8 \\ 3 & 1 & -4 & 3 & -1 & -2 \\ -8 & 3 & -5 & 2 & 0 & -2 \\ -4 & 0 & -8 & -7 & 1 & -3 \\ -5 & -2 & 3 & -6 & 2 & 1 \end{pmatrix},$$
$$B^t = (-1 \quad 2 \quad -4 \quad -1 \quad -5).$$

Ned fra himlen er dalet X_0, hvor

$$X_0^t = (2 \quad 4 \quad 0 \quad 0 \quad 7 \quad 0).$$

Vis, at X_0 er en optimal løsning til (P).

Vink: Man kan bruge Opgave 31. Betingelserne (c) og (d) giver tilsammen et lineært ligningssystem med ubekendte $y_1, \ldots, y_m$. Dette ligningssystem har forhåbentlig netop én løsning. (Hvorfor forhåbentlig?) Brug løsningen!

Opgave 32. Lad (P) være et standard minimeringsproblem med

$$C^t = (-8 \quad 9 \quad -12 \quad -4 \quad 11),$$
$$A = \begin{pmatrix} -2 & 3 & -4 & -1 & -3 \\ -1 & -7 & -3 & 2 & -1 \\ -5 & -4 & 6 & -2 & -5 \end{pmatrix},$$
$$B^t = (-1 \quad -1 \quad -22).$$

Lad videre X_0 være givet ved

$$X_0^t = (0 \quad 2 \quad 0 \quad 7 \quad 0).$$

Undersøg, om X_0 er en optimal løsning til (P).

Opgave 33. (Følsomhedsanalyse.) Vi betragter et standard minimeringsproblem (P). Vi antager, at (P) har en optimal løsning X_0. Det duale problem $(P^\star)$ har da også en optimal løsning Y_0. (Hvorfor?) Sammen med (P) betragter vi også familien af "perturberede" problemer

(P_T) Minimér C^tX ub. $AX \geq B + T$ og $X \geq 0$,

hvor $T^t = (t_1 \cdots t_m)$. Vi vil interessere os for ændringer i den optimale værdi C^tX_0 når (P) erstattes med (P_T). Vi antager følgende er opfyldt:

(#) Der findes $\varepsilon > 0$, således at Y_0 er optimal løsning til alle $(P_T^\star)$ med $|t_i| < \varepsilon$ for $i = 1, \ldots, m$.

Bevis, at ændringen i den optimale værdi er T^tY_0 når $|t_i| < \varepsilon$ for $i = 1, \ldots, m$.

Opgave 34. Minimér

$$-2x_1 + 3x_2 + 12x_3 + 15x_5$$

under bibetingelserne

$$\begin{aligned} -2x_1 + x_2 + 3x_3 + x_4 + 10x_5 &\geq 2, \\ -x_1 + x_2 + 4x_3 - 2x_4 - 6x_5 &\geq 1, \\ x_1, x_2, x_3, x_4, x_5 &\geq 0. \end{aligned}$$

Løs derefter den tilsvarende maksimeringsopave.

Opgave 35. Minimér hver af funktionerne

$$x_1 + x_2, \quad x_1 - x_2, \quad -x_1 + x_2, \quad 2x_1 - x_2$$

under bibetingelserne

$$2x_1 + x_2 \geq 2, \quad -x_1 - 2x_2 \leq -2, \quad 2x_1 - x_2 \geq -1, \quad x_1, x_2 \geq 0.$$

Opskriv dernæst de duale programmer. Lav for hver af de fire objektfunktioner en skitse af $M^\star$. Sammenhold med teorien.

OPGAVE 36. Betragt et kanonisk minimeringsproblem

$$(P) \qquad \text{Minimér } C^t X \text{ ub. } AX = B \text{ og } X \geq 0.$$

1° Find samtlige basisløsninger til (P) med

$$A = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 1 & 0 & 1 & 2 \end{pmatrix},$$
$$B = \begin{pmatrix} 6 \\ 4 \end{pmatrix},$$
$$C^t = (1 \quad 1 \quad 1 \quad 1).$$

Løs derefter (P).

2° Som 1° men nu med

$$C^t = (0 \quad 1 \quad 1 \quad 0).$$

3° Som 1° men nu med

$$B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

OPGAVE 37. Minimér $x_1 + x_2 + x_3 + x_4$ under bibetingelserne

$$\begin{aligned} x_1 + x_2 - 2x_3 + x_4 &= 3, \\ x_2 - x_3 + x_4 &\leq 2, \\ -x_1 + 3x_2 - 2x_3 &\geq -4, \\ x_1, x_2, x_3, x_4 &\geq 0. \end{aligned}$$

5 Kontrolteori og Variationsregning

OPGAVE 1. Betrakt variasjonsproblemet

$$\text{maks} \int_0^1 (4xt - \dot{x}^2)dt\,, \quad x(0) = 2\,, \quad x(1) = 2/3\,.$$

Finn Euler-likningen tilordnet problemet, og finn den eneste løsningen av den som tilfredsstiller begge randbetingelsene.

OPGAVE 2. Betrakt variasjonsproblemet

$$\text{minimer} \int_0^T e^{-rt}[g(\dot{x}) + c(t)x]dt\,, \quad x(0) = 0\,,\ x(T) = B$$

der g og c er gitte funksjoner, T, r og B er gitte positive tall, mens $x = x(t)$ er den ukjente funksjonen.

(a) Still opp Euler-likningen for dette problemet.
(b) Finn løsningen av problemet når $r > 0$, $g(\dot{x}) = \dot{x}^2$, $c(t) = 2$.
(c) Problemet ovenfor kan gis følgende tolkning: Man skal skaffe til veje B enheter av et produkt ved tidspunktet T. Produksjonen per tidsenhet er $\dot{x}$, $g(\dot{x})$ angir produksjonskostnadene per tidsenhet, $c(t)$ angir lagringskostnadene per tidsenhet og r er rentefoden. Denne tolkningen fungerer godt bare hvis løsningen av problemet har egenskapen $\dot{x}(t) \geq 0$ i $[0, T]$. Finn i tilfellet i (b) betingelser på parametrene som er nødvendige og tilskrekkelige for at $\dot{x}(t) \geq 0$ i $[0, T]$.

OPGAVE 3. Betrakt variasjonsproblemet

$$\min \int_0^T [px^2 + q\left(\frac{1}{b}(\dot{x} - ax)\right)^2]dt\,, \quad x(0) = x_0\,,\ x(T) = x_T$$

der p, q, a, b, T, x_0 og x_T er ikke-negative konstanter, $b \neq 0$, $q \neq 0$.

(a) Still opp Euler-likningen for problemet.
(b) Finn den generelle løsningen av Euler-likningen.
(c) Velg $p = 0$, $q = 1$, $a = 1$, $b = 1$, $T = 1$, $x_0 = 0$, $x_T = 1$. Finn løsningen av problemet i dette tilfellet.

OPGAVE 4. I teorien for økonomisk vekst forekommer variasjonsregningsproblemet

$$\text{maks} \int_0^T -e^{\dot{x} - ax}dt \text{ når } x(0) = C\,,\ x(T) = D\,.$$

Her er T, a, C og D gitte konstanter, $T > 0$ og $a > 0$.
" (a)" Still opp Euler-likningen for dette problemet.
" (b)" Finn løsningen med de gitte verdier av $x(0)$ og $x(T)$.

5. Betrakt variasjonsproblemet

$$\text{maks} \int_0^1 (2xe^{-t} - 2x\dot{x} - \dot{x}^2)dt\,, \quad x(0) = 0\,,\ x(1) = 1\,.$$

(a) Still opp Euler-likningen for problemet.
(b) Finn den løsningen av Euler-likningen som tilfredsstiller $x(0) = 0$, $x(1) = 1$.

OPGAVE 6.

(a) Betrakt variasjonsproblemet

$$\text{maks} \int_0^1 (-2\dot{x} - \dot{x}^2)e^{-t/10}dt\,, \quad x(0) = 1\,,\ x(1) = 0\,.$$

Finn Euler-likningen og vis at dens allmenne løsning kan skrives på formen $x(t) = Ae^{t/10} - t + B$. Løs så problemet (1).

(b) Et oljereservoar inneholder ved tidspunktet $t = 0$ $\bar{x}$ enheter olje. Vi ønsker å tappe ut all oljen i et gitt tidsrom $[0, T]$. Er $x(t)$ antall olje som er igjen ved tidspunktet t, så er $-\dot{x}(t)$ upptappingshastigheten (som er positiv når $x(t)$ avtar). Vi antar at verdensmarkedetes pris per liter olje er gitt og lik $ae^{\alpha t}$. Kostnadene per tidsenhet ved uttvinningen forutsetter vi er lik $(\dot{x}(t))^2e^{\beta t}$. Profitten per tidsenhet er da

$$\pi = -\dot{x}(t)ae^{\alpha t} - (\dot{x}(t))^2e^{\beta t}\,.$$

Her er a, α og β konstanter, $a > 0$.

Gi en økonomisk tolkning av variasjonsproblemet

$$\text{maks} \int_0^T [-\dot{x}(t)ae^{\alpha t} - (\dot{x}(t))^2e^{\beta t}]e^{-rt}dt \quad x(0) = \bar{x}\,,\ x(T) = 0\,,$$

der r er en positiv konstant.

(c) Still opp Euler-likningen til variasjonsproblemet, og vis at i optimum vil

$$\frac{\partial \pi}{\partial \dot{x}} = ce^{rt} \ \text{ for en konstant } c\,.$$

OPGAVE 7.

(a) Finn Euler-likningen til følgende variasjonsproblem:

$$\min \int_{t_0}^{t_1} (b(t)x + a(t)\dot{x}^2)dt\,, \quad x(t_0) = x_0\,,\ x(t_1) = x_1\,.$$

He er t_0, t_1, x_0 og x_1 konstanter, mens $a(t)$ og $b(t)$ er gitte positive, deriverbare funksjoner.

(b) Vis at den generelle løsningen til Euler-likningen kan skrives på formen

$$x(t) = \int \left(\frac{C}{a(t)} + \frac{1}{2a(t)} \int b(t)dt \right) dt + D\,,$$

der C og D er vilkårlige konstanter.

(c) Finn $x(t)$ når $a(t) = t$, $b(t) = t^2$, $t_0 = 1$, $t_1 = 3$, $x(1) = 0$, $x(3) = 2$.

OPGAVE 8. Betrakt variasjonsproblemet

$$(*) \qquad \text{maks} \int_0^1 [\ln(ax - \dot{x}]e^{-bt}dt\,, \quad x(0) = 1\,,\ x(1) = 1\,,$$

der a og b er positive konstanter.

(a) Still opp Euler-likningen tilordnet problemet.

(b) Finn den eneste mulige løsningen av problemet $(*)$.

OPGAVE 9. Betrakt variasjonsproblemet

$$\min \int_0^1 (x^2 + 2xt\dot{x} + \dot{x}^2)dt\,, \quad x(0) = 1\,,\ x(1) = 1\,.$$

(a) Finn Euler-likningen for problemet.
(b) Finn den eneste mulige løsningen på problemet.
(c) Vis at

$$\int_0^1 [(x(t))^2 + 2x(t) \cdot t \cdot \dot{x}(t) + (\dot{x}(t))^2]dt = 1 + \int_0^1 (\dot{x}(t))^2 dt$$

for alle tillatte funksjoner $x(t)$ i problemet.
(Hint: $\frac{d}{dt}(tx^2) = x^2 + 2tx\dot{x}$.)
(d) Kan vi av (c) slutte at løsningen i (b) faktisk gir minimum i problemet ?

OPGAVE 10. Betrakt variasjonsproblemet

$$\text{maksimer} \int_0^T U(\bar{b} - \dot{x}e^{rt})dt\,, \quad x(0) = x_0\,,\ x(T) = 0\,,$$

der $x = x(t)$ er den ukjente funksjonen T, $\bar{c}$, r og x_0 er positive konstanter og U er en gitt deriverbar funksjon.

(a) Still opp Euler-likningen tilordnet dette problemet.
(b) Sett $U(c) = \frac{1}{1-v}c^{1-v}$, der v er en konstant mellem 0 og 1. Løs Euler-likningen i dette tilfellet.

OPGAVE 11. Betrakt variasjonsproblemet

$$(*) \qquad \text{maks} \int_{=}^T [\ln(2\ddot{K} - \dot{K})]e^{-\frac{1}{4}t}dt\,, \quad K(0) = K_0\,,\ K(T) = K_T$$

(a) Still opp Euler-likningen tilordnet problemet, og påvis at den kan skrives på formen $a\dot{K} + b\dot{K} + cK = 0$, der a, b og c er konstanter.
(b) Finn den eneste mulige løsningen av problemet $(*)$.

OPGAVE 12. Betrakt problemet

$$\text{maksimer} \int_0^T (x(t) - (u(t))^2)dt$$

gitt at $\dot{x}(t) = x(t) + u(T)$, $x(0) = 0$, $u(t) \in (-\infty, \infty)$, $x(T)$ fri. Finn en optimal kontroll $u^*(t)$, med tilhørende bane $x^*(t)$.

OPGAVE 13. Betrakt kontrollproblemet (fra økonomisk vekstteori)

$$\text{maks} \int_0^{10} (\sqrt{K} - u\sqrt{K})dt\,,\ \dot{K} = u\sqrt{K}\,,\ K(0) = 1\,,\ K(10)\text{ fri}\,,\ 0 \le u \le 1\,.$$

(a) Still opp maksimumsprinsippets betingelser for at $(K^*(t), u^*(t))$ skal løse problemet.
(b) Finn den eneste mulige løsningen av problemet og den tilhørende adjungerte funksjonen $p(t)$. (Vink: Vis at $p(t)$ er avtakende.)

OPGAVE 14. Løs kontrollproblemet

$$\text{maks} \int_0^2 (u^2 - x)dt \quad \text{når} \quad \dot{x} = u\,,\ 0 \le u \le 1\,,\ x(0) = 0\,,\ x(2) \text{ er fri}\,.$$

OPGAVE 15.

(a) Løs kontrollproblemet

$$\text{maks} \int_0^5 -3u^2 e^{-0.15t} dt\,, \qquad \dot{x} = u\,,\ x(0) = 0\,,\ x(5) = 1500\,,\ u \ge 0\,.$$

(b) En kommune ønsker å nydyrke et område i løpet av 5 år. La $x(t)$ være antall mål som er nydyrket ved tidspunktet t og la $u(t)$ være nydyrkningshastigheden slik at

$$\dot{x}(t) = u(t)\,.$$

La kostnadene per tidsenhet ved nydyrkning være gitt ved funksjonen $C(u, t)$. Den totale neddiskonterte kostnaden ved nydyrkningen i perioden fra $t = 0$ til $t = 5$ når renten er r, er da $\int_0^5 C(u, t)e^{-rt}dt$. Betrakt problemet

$$\min \int_0^5 C(u, t)e^{-rt} dt\,, \qquad \dot{x} = u\,,\ x(0) = 0\,,\ x(5) \ge 1500\,,\ u \ge 0\,.$$

skriv ned de betingelsene maksimumsprinsippet gir.

(c) Løs problemet når $r = 0$ og $C(u, t) = g(u)$, med $g(0) = 0$, $g(u) \ge 0$ og $g''(u) > 0$.

OPGAVE 16. Betrakt kontrollproblemet

$$\text{maks} \int_0^T (x^2 - x)dt \quad (T \text{ gitt positiv konstant}) \tag{1}$$

$$\dot{x} = u\,, \quad x(0) = 0\,, \quad x(T) \text{ fri} \tag{2}$$

$$u = u(t) \in [0, 1] \tag{3}$$

(a) Still opp de betingelsene maksimumsprinsippet gir.

(b) Klargjør hvilke muligeforløp den adjungerte funksjonen $p(t)$ kan ha, og finn de tilhørende forslag til optimale løsninger som maksimumsprinsippet gir. (Vink: $p(t)$ er konkav.) En kan vise (det skal ikke gjøres) at det finns en optimal løsning. Finn denne.

OPGAVE 17. Betrakt kontrollproblemet

$$\text{maks} \int_0^T (2x^2 e^{-2t} - ue^t)dt \quad (T \text{ gitt positiv konstant}) \tag{1}$$

$$(2) \qquad \dot{x} = ue^t\,, \quad x(0) = 1\,, \quad x(T) \text{ fri}$$

$$(3) \qquad u = u(t) \in [0,1]$$

(a) Finn den eneste mulige løsningen av problemet.
(b) Vis at om $(x^*(t), u^*(t))$ er løsningsparet i (a) og

$$V(T) = \int_0^T (2(x^*(t))^2 e^{-2t} - u^*(t)e^t)dt\,,\ H^*(T) = H(T, x^*(T), u^*(T), p(T))$$

(der H er Hamiltonfunksjonen, da er $V'(T) = H^*(T)$.

OPGAVE 18. La T være en gitt positiv konstant og betrakt kontrollproblemet

$$\text{maksimer} \int_0^T e^{-\beta t}\sqrt{u}dt \text{ når } \dot{x}(t) = \alpha x(t) - u(t),\ u(t) \geq 0,\ x(0) = 1,\ x(T) = 0,$$

der α og β er positive konstanter.

(a) Still opp de betingelsene sommaksimumsprinsippet gir, og løs problemet. (Du kan gå ut fra at det finnes en optimal løsning.)
(b) Under hvilke betingelser vil den optimale kontrollfunksjonen u^* være konstant ?
(c) Hva skjer hvis terminalbetingelsen $X(T) = 0$ endres til $x(T) \geq 0$?

OPGAVE 19. Betrakt variasjonsproblemet

$$\text{maks} \int_0^T \left(\frac{1}{100}tx - \dot{x}^2\right) e^{-t/10}dt\,,\ x(0) = 0\,,\ x(T) = S$$

(a) Finn Euler-likningen tilordnet problemet og finn dens allmene løsning.
(b) Sett $T = 10$ og $S = 20$ og finn løsningen av problemet i dette tilfellet. Påvis at du virkelig har funet den optimale løsningen.
(c) Hvis løsningen du fant i (b) kalles $x^*(t)$, vis at $\dot{x}^*(t) > 1$ for alle $t \in [0, 10]$.
(d) Betrakt kontrollproblemet

$$\text{maks} \int_0^T \left(\frac{1}{100}tx - u^2\right) e^{-t/10}dt$$

Still opp maksimumsprinsippets betingelser for løsning av problemet.
(e) Hva blir den eneste mulige løsningen dersom vi forutsetter at den optimale kontrollen $u^*(t) > 1$ for alle $t \in [0, T]$? Påvis at dette bekrefter løsningen du fant i (b).
(f) Kan det tenkes at den optimale kontrollen i problemet i (d) kan anta verdien 1 i et interval ?

OPGAVE 20. Betrakt kontrollproblemet

$$\max \int : 0^T (ax - bu)dt\,,\ \dot{x} = x + u\,,\ x(0) = x_0\,,\ x(T) \text{ fri}\,,\ u \in [0, 12]$$

der alle konstanterne er positive.

(a) Løs problemet. (Vink: Du vil måtte skille mellom tilfellene $b < a(e^T - 1)$ og $b \geq a(e^T - 1)$.)

(b) La $J(x_0, T)$ betegne den optimale verdifunksjonen. Vis at $\partial J(x_0, T)/\partial x_0 = p(0)$, der $p(t)$ er den adjungerte funksjonen, og at $\partial J(x_0, T)/\partial T = H^*(T)$, når vi lar $H^*(t)$ betegne Hamiltonfunksjonen utregnet "langs" den optimale banen.

opgave 21. Betrakt kontrollproblemet

$$\text{maksimer} \int_0^1 (2x - x^2)dt$$
$$\text{når } \dot{x} = u\,, \quad x(0) = 0\,, \quad x(1) = 0\,, \quad u \in [-1, 1]$$

(a) Still opp maksimumsprinsippets betingelser for løsning av problemet.

(b) Forklar hvorfor et tillatt par $(x(t), u(t))$ som tilfredsstiller betingelsene i (a) må være en optimal løsning.

(c) Finn en optimal løsning. (Vink: Vis at $p(t)$ må være strengt avtagende.)

OPGAVE 22. Betrakt kontrollproblemet

$$\text{maksimer} \int_0^T (24x - ux)dt \ \text{ når } \ \dot{x} = ux^3\,,\ u \in [0, 1]\,,\ x(0) = 1\,,\ x(T) \text{ fri}\,.$$

(a) Still opp de nødvendige betingelsene som maksimumsprinsippet gor for at $/x^*(t), u^*(t))$ skal være en løsning av problemet.

(b) Finn den eneste mulige løsningen av problemet når $T = 1/3$. (Vink: Vis at $p(t)((x^*(t))^2$ må være strengt avtagende.)

(c) Undersøk problemet når $T = 1$.

OPGAVE 23. Betrakt kontrollproblemet

$$(K) \qquad \text{maksimer} \int_0^T x(2 - u)e^{-t}dt \ \text{ når } \begin{cases} \dot{x} = ux, & u \in [1/2, 3/4], \\ x(0) = 1, & x(T) \text{ fri} \end{cases}$$

der T er et gitt tall ≥ 1.

(a) Vis at hvis $(x(t), u(t))$ er et tallatt par i problemet (K), så har vi $x(t) \geq e^{t/2}$ for alle $t \in [o, T]$.

(b) Still opp de nødvendige betingelsene som maksimumsprinsippet gir for en løsning av (K).

(c) Løs problemet (K). (Vink: $p(t) - e^7 - t$ blir en strengt avtakende funksjon av t.)

(d) Sett nå $T = \infty$ og foreslå en løsning av (K) i dette tilfellet. Er tilstrekkelige betingelser oppfylt? (Vink: For å få et forslag til løsning med tilhørende adjungert funksjon, la $T \to \infty$ i resultaterne du fant i (c).)

OPGAVE 24. Betrakt kontrollproblemet

$$\max \int_0^1 (x(t))^2 dt\,,\ \dot{x} = 1 - (u(t))^2\,,\ x(0) = 4\,,\ u(t) \in [-1, 2]$$

(a) Skriv opp betingelsene maksimumsprinsippet gir for problemet.
(b) Finn den eneste mulige løsningen av problemet.

OPGAVE 25.

(a) Løs kontrollproblemet

$$\text{maks}\{\int_0^1 (x-u)dt + \frac{1}{2}x(1)\},\ \dot{x} = u\,,\ x(0) = 1/2\,,\ x(1)\text{ fri}\,,\ u \in [0,1]$$

(b) Erstatt kriteriefunksjonen med

$$\int_0^1 (x-u)dt - \frac{1}{4}(x(1)-2)^2$$

Løs problemet i (a) med denne kriteriefunksjon.