Kirchberger's sætning om separation af to mængder

Maria Larissa Ziino

I denne artikel fremføres to sætninger af henholdsvis den østrigske matematiker Eduard Helly og den tyske matematiker Paul Kirchberger, som hver især skrev deres disputats i starten af 1900-tallet. Jeg vil her vise Kirchberger's sætning ved at benytte mig af Helly's sætning.
Disse to sætninger kan faktisk betragtes som værende duale til hinanden. Den mest fremtrædende forskel på Helly's og Kirchberger's sætning er, at Helly's sætning beskriver betingelserne for, at et antal mængder har noget tilfælles mens at Kirchberger's sætning beskriver betingelserne for at kunne separere to mængder.
Kirchberger's ide kan beskrives ved følgende eksempel: Forestil dig en mark med et antal får og et antal bukke der står stille. Hvordan ved man om fårene kan adskilles fra bukkene ved et lige hegn?
Sætningen giver, at hvis hver gruppe af 4 får og bukke kan adskilles af et lige hegn, så kan de alle adskilles af et lige hegn. Den generelle formulering af sætningen ses i Sætning 13.

Indledende begreber

Det rum jeg arbejder i her kaldes $\mathbb{E}^n$ og det defineres ved følgende

Definition 1 Det n-dimensionale Euklidiske rum $\mathbb{E}^n$ er givet ved det lineære rum $\mathbb{R}^n$ udstyret med det indre produkt $\langle x, y\rangle = \sum_{i=1}^n \alpha_i\beta_i$, hvor $x = (\alpha_1, \dots, \alpha_n)$ og $y = (\beta_1, \dots, \beta_n)$.
Nulvektoren i $\mathbb{E}^n$ defineres som $o = (0, \dots, 0)$.

Konveksitetsteorien giver følgende, hvor jeg starter ud med at minde om begrebet affin kombination.

Definition 2 Lad $a_0, \dots, a_k$ være punkter i $\mathbb{E}^n$. Disse siges at være affint afhængige hvis der findes et talsæt $(\alpha_0, \dots, \alpha_k) \in \mathbb{R}^{k+1} \backslash (0, \dots, 0)$, så $\alpha_0 a_0 + \cdots + \alpha_k a_k = o$, og $\alpha_0 + \cdots + \alpha_k = 0$.

Sætning 3 *Enhver delmængde af $\mathbb{E}^n$ bestående af mindst $n+1$ forskellige punkter er lineært afhængig. Enhver delmængde af $\mathbb{E}^n$ bestående af mindst $n+2$ forskellige punkter er affint afhængig.*

Begreberne konveks kombination, konveks mængde og det konvekse hylster defineres nu.

Definition 4 Lad $k \in \mathbb{N}_0$; $x_0, \dots, x_k \in \mathbb{E}^n$; $\lambda_0, \dots, \lambda_k \in \mathbb{E}$. 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 $x_0, \dots, x_k$.

Definition 5 En delmængde S af $\mathbb{E}^n$ siges at være konveks hvis der for enhver konveks kombination $\lambda_0 s_0 + \cdots + \lambda_k s_k$ af elementer fra S gælder, at $\lambda_0 s_0 + \cdots \lambda_k s_k \in S$.

Definition 6 Lad $S \subseteq \mathbb{E}^n$. Ved det konvekse hylster - $\mathrm{conv} S$ - af S forstås den mindste konvekse delmængde af $\mathbb{E}^n$, som indeholder S.

Sætning 7 *Lad $S \subseteq \mathbb{E}^n$, da består* $\mathrm{conv} S$ *netop af alle konvekse kombinationer af elementer fra S, dvs.*

$$\mathrm{conv} S = \left\{ \sum_{i=0}^{k} \lambda_i x_i \,\middle|\, k \in \mathbb{N}_0, \lambda_i \geq 0, \lambda_0 + \cdots + \lambda_k = 1, x_i \in S \right\}.$$

Sætning 8 (Carathéodory) *For en ikke-tom delmængde S af $\mathbb{E}^n$ kan ethvert $x \in$* $\mathrm{conv} S$ *udtrykkes ved en konveks kombination af $n+1$ eller færre punkter fra S.*

Helly's sætning

Helly's sætning vises ved at gøre brug af en sætning af Johann Radon som er givet ved følgende

Sætning 9 (Radon's sætning) *Lad* $S = \{x_1, \ldots, x_r\}$ *være en vilkårlig endelig mængde af* r *indbyrdes forskellige punkter i* $\mathbb{E}^n$*. Hvis* $r \geq n+2$ *kan* S *opdeles i to disjunkte delmængder,* S_1 *og* S_2*, så der gælder at* $\mathrm{conv} S_1 \cap \mathrm{conv} S_2 \neq \varnothing$*.*

Bevis. Da $r \geq n+2$ giver Sætning 3 at S er affint afhængig. Dvs. der eksisterer skalarer $\alpha_1, \ldots, \alpha_r$, hvor ikke alle er 0, således at

$$\sum_{i=1}^{r} \alpha_i x_i = o \quad \text{og} \quad \sum_{i=1}^{r} \alpha_i = 0.$$

Ifølge ovenstående må mindst to af disse α_i'er have modsatte fortegn, så vi kan uden tab af generalitet antage, at de første k er større end eller lig nul og at de sidste $r-k$ er skarpt mindre end nul. Da α_i'erne summer til nul kan vi definere α som

$$\alpha = \alpha_1 + \cdots + \alpha_k = -(\alpha_{k+1} + \cdots + \alpha_r).$$

Så er $\alpha > 0$ og vi kan definere x som

$$x = \sum_{i=1}^{k} \frac{\alpha_i}{\alpha} x_i = \sum_{i=k+1}^{r} \left(-\frac{\alpha_i}{\alpha}\right) x_i.$$

Da $\sum_{i=1}^{k} \frac{\alpha_i}{\alpha} = 1$ er x en konveks kombination af $x_1, \ldots, x_k$, så af Sætning 7, får vi at $x \in \mathrm{conv}\{x_1, \ldots, x_k\}$. På samme måde får vi at $x \in \mathrm{conv}\{x_{k+1}, \ldots, x_r\}$.
Ved at lade $S_1 = \{x_1, \ldots, x_k\}$ og $S_2 = \{x_{k+1}, \ldots, x_r\}$ har vi at $\mathrm{conv} S_1 \cap \mathrm{conv} S_2 \neq \varnothing$. Da punkterne $\{x_1, \ldots, x_r\}$ alle er parvis forskellige har vi yderligere at $S_1 \cap S_2 = \varnothing$. □

Sætning 10 (Helly's sætning) *Lad $\mathcal{F} = \{B_1, \ldots, B_r\}$ være en familie bestående af r konvekse mængder i $\mathbb{E}^n$ med $r \geq n+1$. Hvis enhver delfamilie bestående af $n+1$ mængder fra $\mathcal{F}$ har ikke-tom fællesmængde så har alle r mængder ikke-tom fællesmængde.*

Bevis. Sætningen vises ved induktion efter $r-1$. Sætningen er triviel for induktionsstarten $r = n+1$ og derfor klart opfyldt. Antag at $r \geq n+2$ og at sætningen holder for $r-1$ konvekse mængder. Vi vil vise at den gælder for r.
For hvert $i \in \{1, \ldots, r\}$, eksisterer der et punkt x_i for hvilket der ifølge induktionsantagelsen gælder at

$$x_i \in B_1 \cap \cdots \cap B_{i-1} \cap B_{i+1} \cap \cdots \cap B_r. \tag{1}$$

Mængden af alle disse punkter betegnes $S = \{x_1, \ldots, x_r\}$. Hvis der er gentagelser i S, dvs. $x_i = x_j$ for et par i, j hvor $i \neq j$, så har vi både at

$$\begin{array}{rrcl} & x_i & \in & B_1 \cap \cdots \cap B_{i-1} \cap B_{i+1} \cap \cdots \cap B_r \\ \text{og at} & x_j & \in & B_1 \cap \cdots \cap B_{j-1} \cap B_{j+1} \cap \cdots \cap B_r. \end{array}$$

Da $x_i = x_j$ har vi derfor, at de begge ligger i $B_1 \cap \cdots \cap B_r$ og derfor er det ønskede opnået i dette tilfælde.
Antag derfor, at x_i'erne er parvis forskellige. Da $r \geq n+2$ kan vi benytte Radon's sætning 9 til at finde to disjunkte delmængder af S, S_1 og S_2, der opfylder at $\operatorname{conv} S_1 \cap \operatorname{conv} S_2 \neq \varnothing$. Uden tab af generalitet kan vi lade $S_1 = \{x_1, \ldots, x_k\}$ og $S_2 = \{x_{k+1}, \ldots, x_r\}$, hvor $1 \leq k < r$. Lad derfor

$$x \in \operatorname{conv}\{x_1, \ldots, x_k\} \cap \operatorname{conv}\{x_{k+1}, \ldots, x_r\}. \tag{2}$$

Vi ønsker at vise at $x \in \bigcap_{i=1}^{r} B_i$.
På grund af (1) har vi for alle $i = 1, \ldots, k$ at $x_i \in B_{k+1} \cap \cdots \cap B_r$.

Vi minder om, at B_i er konveks for alle $i = k+1, \ldots, r$. Vi har da, at det konvekse hylster af x_i'erne, for $i = 1, \ldots, k$, ligger i snittet af B_i'erne, for $i = k+1, \ldots, r$, dvs.

$$\operatorname{conv}\{x_1, \ldots, x_k\} \subseteq B_{k+1} \cap \cdots \cap B_r.$$

Tilsvarende gælder for alle $i = k+1, \ldots, r$ at $x_i \in B_1 \cap \cdots \cap B_k$. Da B_i er konveks for alle $i = 1, \ldots, k$ er $\operatorname{conv}\{x_{k+1}, \ldots, x_r\} \subseteq B_1 \cap \cdots \cap B_k$. Nu har vi det ønskede da (2) giver at

$$x \in B_{k+1} \cap \cdots \cap B_r \qquad \text{og} \qquad x \in B_1 \cap \cdots \cap B_k. \qquad \square$$

Kirchberger's sætning

Da Kirchberger skrev sin Ph.D.-afhandling i 1902 viste han blandt andet den efterfølgende sætning, men han viste også Separationssætningen og et specialtilfælde af Carathéodory's sætning, nemlig Sætning 8 for endelige mængder. Idet selve Carathéodory's sætning først blev vist af Carathéodory ca. 5 år senere (hvorfra sætningen har fået navnet Carathéodory's sætning) er det spændende at studere Kirchberger's 30 sider lange disputats [1].
For at kunne vise nedenstående sætning vil jeg minde om begreberne halvrum i $\mathbb{E}^{n+1}$ og hyperplan i henholdsvis $\mathbb{E}^n$ og $\mathbb{R} \times \mathbb{E}^n$.

Definition 11 Lad $y \in \mathbb{E}^n \backslash (0, \ldots, 0)$ og $\alpha \in \mathbb{E}$. Hyperplanen $H(y, \alpha)$ i $\mathbb{E}^n$ er defineret ved

$$H(y, \alpha) = \{x \in \mathbb{E}^n : \langle y, x \rangle = \alpha\}.$$

Definition 12 En hyperplan i $\mathbb{R} \times \mathbb{E}^n$ er for en vektor $(s, a) \in \mathbb{R} \times \mathbb{E}^n$ og et $t \in \mathbb{R}$, hvor $(s, a) \neq (0, o)$, givet ved

$$H = \{(\alpha, x) \in \mathbb{R} \times \mathbb{E}^n : (\alpha, x) \cdot (s, a) = t\},$$

hvor (s, a) er normalvektoren til H. Det øvre åbne halvrum er givet ved

$$\{(\alpha, x) \in \mathbb{R} \times \mathbb{E}^n : (\alpha, x) \cdot (s, a) > t\},$$

som er nedad begrænset af H. Det nedre åbne halvrum er givet ved

$$\{(\alpha, x) \in \mathbb{R} \times \mathbb{E}^n : (\alpha, x) \cdot (s, a) < t\},$$

som er opad begrænset af H.

Sætning 13 (Kirchberger's sætning) *Lad P og Q være ikke-tomme endelige delmængder af $\mathbb{E}^n$. Så kan P og Q separeres strengt af en hyperplan hvis og kun hvis der for enhver delmængde T bestående af $n+2$ eller færre punkter fra $P \cup Q$ findes en hyperplan der separerer $T \cap P$ og $T \cap Q$ strengt.*

Bevis. Hvis P og Q kan separeres strengt af en hyperplan så vil der specielt, for hvert valg af T, gælde at $T \cap P$ og $T \cap Q$ kan separeres af hyperplanen.
For at vise den anden vej antager vi, at der for enhver delmængde T bestående af $n+2$ eller færre punkter fra $P \cup Q$ findes en hyperplan H, der separerer $T \cap P$ og $T \cap Q$ strengt.
Definer for hvert $p \in P$ og $q \in Q$ de åbne halvrum

$$\begin{aligned}
F_p &= \{(\alpha, x) \in \mathbb{R} \times \mathbb{E}^n : (\alpha, x) \cdot (-1, p) < 0\} \\
&= \{(\alpha, x) \in \mathbb{R} \times \mathbb{E}^n : -\alpha + \langle x, p \rangle < 0\} \\
&= \{(\alpha, x) \in \mathbb{R} \times \mathbb{E}^n : \langle x, p \rangle < \alpha\} \\
G_q &= \{(\alpha, x) \in \mathbb{R} \times \mathbb{E}^n : (\alpha, x) \cdot (-1, q) > 0\} \\
&= \{(\alpha, x) \in \mathbb{R} \times \mathbb{E}^n : -\alpha + \langle x, q \rangle > 0\} \\
&= \{(\alpha, x) \in \mathbb{R} \times \mathbb{E}^n : \langle x, q \rangle > \alpha\}.
\end{aligned}$$

Der findes per antagelse en hyperplan

$$H(u,\beta) = \{y \in \mathbb{E}^n : \langle u, y\rangle = \beta\}$$

i $\mathbb{E}^n$, således at $T \cap P$ og $T \cap Q$ separeres af denne hyperplan $H(u,\beta)$. Uden tab af generalitet kan vi antage at $T \cap P$ er opad begrænset af $H(u,\beta)$ og at $T \cap Q$ er nedad begrænset af $H(u,\beta)$. Dvs.

$$\begin{array}{rcll} \langle u,p\rangle & < & \beta & \text{for alle} \quad p \in T \cap P \\ \langle u,q\rangle & > & \beta & \text{for alle} \quad q \in T \cap Q. \end{array}$$

Vi har derfor at $(\beta,u) \in F_p$ for alle $p \in T \cap P$ og at $(\beta,u) \in G_q$ for alle $q \in T \cap Q$. Dvs.

$$(\beta,u) \in (\bigcap_{p\in T\cap P} F_p) \cap (\bigcap_{q \in T\cap Q} G_q).$$

Så da T består af $n+2$ eller færre punkter fra $P \cup Q$ findes der for vilkårlige $n+2$ eller færre halvrum F_p og G_q i $\mathbb{R}\times\mathbb{E}^n$ et punkt som de har tilfælles. Helly's sætning 10 giver, da F_p og G_q er konvekse for alle $p \in P$ og alle $q \in Q$, at der findes et punkt (δ, z) som alle halvrum F_p og G_q har tilfælles, for $p \in P$ og $q \in Q$.
Dvs.

$$\begin{array}{rrcll} & \langle z,p\rangle & < & \delta & \text{for alle} \quad p \in P \\ \text{og} & \langle z,q\rangle & > & \delta & \text{for alle} \quad q \in Q, \end{array}$$

derfor vil hyperplanen

$$H'(z,\delta) = \{y \in \mathbb{E}^n : \langle z, y\rangle = \delta\}$$

separere P og Q strengt i $\mathbb{E}^n$. □

Litteratur

[1] Kirchberger, P.: 1903. Über Tchebychefsche Annäherungsmethoden, *Math. Ann.*, **57**, 509-540.

[2] Lay, S. R.: 1982. *Convex Sets and their Applications*, John Wiley & Sons, New York.