Den flerdimensionale normalfordeling

1 Stokastiske vektorer

Ved en stokastisk vektor skal vi forstå en vektor, hvor de enkelte komponenter er sædvanlige stokastiske variable.

For den stokastiske vektor $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ definerer vi middelværdivektoren som

$$\mathrm{E}[\boldsymbol{Y}] = (\mathrm{E}[Y_1], \ldots, \mathrm{E}[Y_n]) \ .$$

I sammenhæng med matricer vil vi opfatte $\boldsymbol{Y}$ og $\mathrm{E}[\boldsymbol{Y}]$ som søjlevektorer.

Lad os betragte to stokastiske vektorer $\boldsymbol{X} = (X_1, \ldots, X_m)$ og $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$. Alle kovarianser mellem $\boldsymbol{X}$'s komponenter og $\boldsymbol{Y}$'s komponenter kan naturligt samles i en $m \times n$ matrix, den såkaldte kovariansmatrix:

$$\begin{aligned}
\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y}) &= \{\mathrm{Cov}(X_i, Y_j)\} \\
&= \{\mathrm{E}[(X_i - \mathrm{E}[X_i])(Y_j - \mathrm{E}[Y_j])]\} \\
&= \mathrm{E}\left[(\boldsymbol{X} - \mathrm{E}[\boldsymbol{X}])(\boldsymbol{Y} - \mathrm{E}[\boldsymbol{Y}])^\top\right] \ .
\end{aligned}$$

I det sidste udtryk skal vi ved middelværdien af en matrix naturligvis forstå matricen bestående af alle elementernes middelværdier.

Sætter vi $\boldsymbol{X} = \boldsymbol{Y}$ i $\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y})$, får vi den såkaldte variansmatrix for $\boldsymbol{Y}$ (også kaldet varians-kovariansmatrix, kovariansmatrix, dispersionsmatrix):

$$\mathrm{Var}(\boldsymbol{Y}) = \mathrm{Cov}(\boldsymbol{Y}, \boldsymbol{Y}) = \{\mathrm{Cov}(Y_i, Y_j)\} \ .$$

Bemærk, at $\mathrm{Var}(\boldsymbol{Y})$ har $\mathrm{Var}(Y_i)\,,\ i = 1, \ldots, n$, i hoveddiagonalen. Bemærk endvidere, at $\mathrm{Var}(\boldsymbol{Y})$ er symmetrisk, da $\mathrm{Cov}(Y_i, Y_j) = \mathrm{Cov}(Y_j, Y_i)$.

At to stokastiske vektorer er uafhængige, betyder at en vilkårlig komponent i den ene vektor er uafhængig af en vilkårlig komponent i den anden vektor. (To forskellige komponenter i samme vektor kan være afhængige eller uafhængige.) Bemærk, at

$$\boldsymbol{X}, \boldsymbol{Y} \text{ uafhængige} \quad \Rightarrow \quad \mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y}) = O.$$

Lad os betragte en lineær transformation T af den stokastiske vektor $\boldsymbol{Y}$ repræsenteret af matricen A, dvs. $T : R^n \to R^m,\ T(\boldsymbol{Y}) = A\boldsymbol{Y}$. Da middelværdioperatoren er lineær, har vi umiddelbart, at $\mathrm{E}[A\boldsymbol{Y}] = A\,\mathrm{E}[\boldsymbol{Y}]$. For variansmatricen får vi, at

$$\begin{aligned}
\mathrm{Var}(A\boldsymbol{Y}) &= \mathrm{E}\left[(A\boldsymbol{Y} - A\mathrm{E}[\boldsymbol{Y}])(A\boldsymbol{Y} - A\mathrm{E}[\boldsymbol{Y}])^\top\right] \\
&= \mathrm{E}\left[A(\boldsymbol{Y} - \mathrm{E}[\boldsymbol{Y}])(\boldsymbol{Y} - \mathrm{E}[\boldsymbol{Y}])^\top A^\top\right] \\
&= A\,\mathrm{E}\left[(\boldsymbol{Y} - \mathrm{E}[\boldsymbol{Y}])(\boldsymbol{Y} - \mathrm{E}[\boldsymbol{Y}])^\top\right] A^\top \\
&= A\,\mathrm{Var}(\boldsymbol{Y})\,A^\top \ .
\end{aligned}$$

$\mathrm{Var}\,(A\boldsymbol{Y})$ er altså en $m \times m$ matrix.

Betragter vi specielt en linearkombination af $\boldsymbol{Y}$'s komponenter $a_1 Y_1 + \ldots + a_n Y_n$, der kort kan skrives $\boldsymbol{a}^\top \boldsymbol{Y}$, får vi $\mathrm{E}\left[\boldsymbol{a}^\top \boldsymbol{Y}\right] = \boldsymbol{a}^\top \mathrm{E}\,[\boldsymbol{Y}]$ og $\mathrm{Var}\left(\boldsymbol{a}^\top \boldsymbol{Y}\right) = \boldsymbol{a}^\top \mathrm{Var}\,(\boldsymbol{Y})\,\boldsymbol{a}$. $\mathrm{Var}\left(\boldsymbol{a}^\top \boldsymbol{Y}\right)$ er et reelt tal (1×1 matrix).

En ortogonal transformation er en lineær transformation, der repræsenteres af en ortogonal matrix[1], dvs. en matrix hvor søjlevektorerne er ortonormerede. For en sådan matrix gælder, at systemet af rækkevektorer ligeledes er ortonormeret, at den inverse matrix er lig den transponerede, og at den tilhørende determinant antager værdien 1 eller -1.

Lad $\boldsymbol{Z} = C\boldsymbol{Y}$ definere en ortogonal transformation. C er altså en ortogonal matrix. Der gælder, at

$$\sum_i Z_i^2 = \boldsymbol{Z}^\top \boldsymbol{Z} = (C\boldsymbol{Y})^\top C\boldsymbol{Y} = \boldsymbol{Y}^\top C^\top C\boldsymbol{Y} = \boldsymbol{Y}^\top \boldsymbol{Y} = \sum_i Y_i^2\,,$$

dvs. kvadratsummen af en stokastisk vektors komponenter er invariant over for ortogonaltransformation.

Betragt nu en stokastisk kvadratisk form $\boldsymbol{Y}^\top A\boldsymbol{Y}$, hvor $\mathrm{E}\,[\boldsymbol{Y}] = \boldsymbol{\mu}$ og $\mathrm{Var}\,(\boldsymbol{Y}) = \Sigma$. Middelværdien af $\boldsymbol{Y}^\top A\boldsymbol{Y}$ udregnes:

$$\begin{aligned}
\mathrm{E}\left[\boldsymbol{Y}^\top A\boldsymbol{Y}\right] &= \mathrm{E}\left[\sum_i \sum_j a_{ij} Y_i Y_j\right] = \sum_i \sum_j a_{ij} \mathrm{E}\,[Y_i Y_j] \\
&= \sum_i \sum_j a_{ij} \left(\mathrm{Cov}\,(Y_i, Y_j) + \mathrm{E}\,[Y_i]\,\mathrm{E}\,[Y_j]\right) \\
&= \sum_i \sum_j a_{ij}\sigma_{ij} + \sum_i \sum_j a_{ij}\mu_i\mu_j \\
&= \sum_i \sum_j a_{ij}\sigma_{ji} + \boldsymbol{\mu}^\top A\boldsymbol{\mu} = \sum_i (A\Sigma)_{ii} + \boldsymbol{\mu}^\top A\boldsymbol{\mu} \\
&= \mathrm{tr}\,A\Sigma + \boldsymbol{\mu}^\top A\boldsymbol{\mu}
\end{aligned}$$

Når $\mathrm{Var}\,(\boldsymbol{Y}) = \sigma^2 I$ forenkles udtrykket til

$$\mathrm{E}\left[\boldsymbol{Y}^\top A\boldsymbol{Y}\right] = \sigma^2 \mathrm{tr}\,A + \boldsymbol{\mu}^\top A\boldsymbol{\mu}\ ,$$

og når yderligere $\mathrm{E}\,[\boldsymbol{Y}] = \boldsymbol{0}$ til

$$\mathrm{E}\left[\boldsymbol{Y}^\top A\boldsymbol{Y}\right] = \sigma^2 \mathrm{tr}\,A\ .$$

Den momentfrembringende funktion hørende til en stokastisk vektor defineres som

$$M_{\boldsymbol{Y}}(t_1, \ldots, t_n) = \mathrm{E}\left[\exp(\boldsymbol{t}^\top \boldsymbol{Y})\right].$$

[1]Den traditionelle betegnelse, en mere præcis betegnelse ville være ortonormal matrix.

Ofte skrives $M_{\boldsymbol{Y}}(\boldsymbol{t})$ som en forkortelse for $M_{\boldsymbol{Y}}(t_1, \ldots, t_n)$.

De sædvanlige egenskaber ved momentfrembringende funktion har også gyldighed her, dvs.

$$M_{A\boldsymbol{Y}+\boldsymbol{b}}(\boldsymbol{t}) = \exp(\boldsymbol{t}^\top \boldsymbol{b}) M_{\boldsymbol{Y}}(A^\top \boldsymbol{t}),$$

og for $\boldsymbol{X}$ og $\boldsymbol{Y}$ uafhængige

$$M_{\boldsymbol{X}+\boldsymbol{Y}}(\boldsymbol{t}) = M_{\boldsymbol{X}}(\boldsymbol{t}) M_{\boldsymbol{Y}}(\boldsymbol{t}),$$

jf. opgave 4 og 5.

2 Den flerdimensionale normalfordeling

En stokastisk vektor $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ siges, at være n-dimensional normalfordelt, når $Y_1, \ldots, Y_n$ har den simultane tæthed

$$f_{\boldsymbol{Y}}(y_1, \ldots, y_n) = (2\pi)^{-\frac{n}{2}} (\det \Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} (\boldsymbol{y} - \boldsymbol{\mu})^\top \Sigma^{-1} (\boldsymbol{y} - \boldsymbol{\mu}) \right) ,$$

hvor $\boldsymbol{\mu} \in R^n$ og $\Sigma \in R^{n \times n}$. $\boldsymbol{\mu}$ kan vælges vilkårligt, mens Σ skal være en positiv definit matrix[2].

At $\boldsymbol{Y}$ er n-dimensional normalfordelt med parametrene $\boldsymbol{\mu}$ og Σ, skrives kort $\boldsymbol{Y} \sim N_n(\boldsymbol{\mu}, \Sigma)$.

Vi vil først vise, at vi ved en passende affin transformation af $\boldsymbol{Y}$ kan opnå, at komponenterne i den transformerede stokastiske vektor er uafhængige og standardiserede normalfordelte.

Sæt $\boldsymbol{Y} = \Sigma^{\frac{1}{2}} \boldsymbol{X} + \boldsymbol{\mu}$ (se opgave 7, 8 og 9 vedr. $\Sigma^{\frac{1}{2}}$). Jacobideterminanten hørende til denne transformation ses umiddelbart at være $\det \Sigma^{\frac{1}{2}}$. Tæthedsfunktionen for $\boldsymbol{X}$ bliver derfor

$$\begin{aligned} f_{\boldsymbol{X}}(x_1, \ldots, x_n) &= (2\pi)^{-\frac{n}{2}} (\det \Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \boldsymbol{x}^\top \Sigma^{\frac{1}{2}} \Sigma^{-1} \Sigma^{\frac{1}{2}} \boldsymbol{x} \right) \left| \det \Sigma^{\frac{1}{2}} \right| \\ &= (2\pi)^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \boldsymbol{x}^\top \boldsymbol{x} \right) = (2\pi)^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} {x_i}^2 \right) \\ &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{{x_i}^2}{2}} , \end{aligned}$$

hvoraf ses, at $X_i \sim N(0, 1)$, $i = 1, \ldots, n$, og at X_i'erne er uafhængige.

[2]En symmetrisk $n \times n$ matrix er positiv definit, når $\forall \boldsymbol{x} \in R^n \backslash \{\boldsymbol{0}\} : \boldsymbol{x}^\top A \boldsymbol{x} > 0$. A er positiv semidefinit, når $\forall \boldsymbol{x} \in R^n \backslash \{\boldsymbol{0}\} : \boldsymbol{x}^\top A \boldsymbol{x} \geq 0$, og der findes (mindst) et $\boldsymbol{x} \neq \boldsymbol{0}$, så $\boldsymbol{x}^\top A \boldsymbol{x} = 0$. Analogt defineres negativ definit og negativ semidefinit.

Herefter kan vi let udregne $\mathrm{E}[\boldsymbol{Y}]$ og $\mathrm{Var}(\boldsymbol{Y})$:

$$\begin{aligned}\mathrm{E}[\boldsymbol{Y}] &= \Sigma^{\frac{1}{2}}\mathrm{E}[\boldsymbol{X}] + \boldsymbol{\mu} = \Sigma^{\frac{1}{2}}\,\mathbf{0} + \boldsymbol{\mu} = \boldsymbol{\mu}\ ,\\ \mathrm{Var}(\boldsymbol{Y}) &= \Sigma^{\frac{1}{2}}\mathrm{Var}(\boldsymbol{X})\,(\Sigma^{\frac{1}{2}})^\top = \Sigma^{\frac{1}{2}} I_n \Sigma^{\frac{1}{2}} = \Sigma\ .\end{aligned}$$

Parametrene $\boldsymbol{\mu}$ og Σ er altså henholdsvis middelværdivektor og variansmatrix for $\boldsymbol{Y}$.

Den momentfrembringende funktion for $\boldsymbol{Y}$ kan vi bestemme ved følgende udregninger, hvor vi igen benytter transformationen $\boldsymbol{Y} = \Sigma^{\frac{1}{2}}\boldsymbol{X} + \boldsymbol{\mu}$:

$$\begin{aligned}M_{\boldsymbol{Y}}(\boldsymbol{t}) &= \mathrm{E}\left[\exp(\boldsymbol{t}^\top\boldsymbol{Y})\right]\\ &= \int_{R^n} \exp(\boldsymbol{t}^\top\boldsymbol{y})(2\pi)^{-\frac{n}{2}}(\det\Sigma)^{-\frac{1}{2}}\exp\left(-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})^\top\Sigma^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right)d\Omega\\ &= \int_{R^n} \exp(\boldsymbol{t}^\top(\Sigma^{\frac{1}{2}}\boldsymbol{x}+\boldsymbol{\mu}))(2\pi)^{-\frac{n}{2}}\exp\left(-\frac{1}{2}\boldsymbol{x}^\top\boldsymbol{x}\right)d\Omega\\ &= \exp(\boldsymbol{t}^\top\boldsymbol{\mu})\int_{R^n}(2\pi)^{-\frac{n}{2}}\exp\left(-\frac{1}{2}(\boldsymbol{x}^\top\boldsymbol{x} - 2\boldsymbol{t}^\top\Sigma^{\frac{1}{2}}\boldsymbol{x})\right)d\Omega\\ &= \exp\left(\boldsymbol{t}^\top\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}^\top\Sigma\boldsymbol{t}\right)\int_{R^n}(2\pi)^{-\frac{n}{2}}\exp\left(-\frac{1}{2}(\boldsymbol{x}^\top\boldsymbol{x} - 2\boldsymbol{t}^\top\Sigma^{\frac{1}{2}}\boldsymbol{x} + \boldsymbol{t}^\top\Sigma\boldsymbol{t})\right)d\Omega\\ &= \exp\left(\boldsymbol{t}^\top\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}^\top\Sigma\boldsymbol{t}\right)\int_{R^n}(2\pi)^{-\frac{n}{2}}\exp\left(-\frac{1}{2}(\boldsymbol{x} - \Sigma^{\frac{1}{2}}\boldsymbol{t})^\top(\boldsymbol{x} - \Sigma^{\frac{1}{2}}\boldsymbol{t})\right)d\Omega\\ &= \exp\left(\boldsymbol{t}^\top\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}^\top\Sigma\boldsymbol{t}\right)\ ,\end{aligned}$$

idet funktionen under det sidste integraltegn ses at være tæthedsfunktionen for en n-dimensional normalfordeling med middelværdivektor $\Sigma^{\frac{1}{2}}\boldsymbol{t}$ og variansmatrix I_n.

I stedet for at definere den flerdimensionale normalfordeling ud fra tæthedsfunktionen kan vi benytte følgende alternative definition:

$\boldsymbol{Y}$ er n-dimensional normalfordelt, når og kun når der for alle $\boldsymbol{l} \neq \mathbf{0}$ gælder, at $\boldsymbol{l}^\top\boldsymbol{Y}$ er (endimensional) normalfordelt. Benytter vi symbolerne $\boldsymbol{\mu}$ og Σ som ovenfor, skal der altså gælde at $\boldsymbol{l}^\top\boldsymbol{Y} \sim N(\boldsymbol{l}^\top\boldsymbol{\mu}, \boldsymbol{l}^\top\Sigma\,\boldsymbol{l})$.

I denne definition kan vi umiddelbart medtage det singulære tilfælde, hvor Σ kun er positiv semidefinit svarende til, at der findes (mindst) et $\boldsymbol{l} \neq \mathbf{0}$, så $\boldsymbol{l}^\top\Sigma\,\boldsymbol{l} = 0$. I det singulære tilfælde vil der altså optræde linearkombination(er) $\boldsymbol{l}^\top\boldsymbol{Y}$ med varians 0.

Fra tidligere kender vi den momentfrembringende funktion for en normalfordelt stokastisk variabel $X \sim N(\mu, \sigma^2)$. Der gælder $M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$. Følgelig har vi

$$M_{\boldsymbol{l}^\top\boldsymbol{Y}}(t) = \exp\left(\boldsymbol{l}^\top\boldsymbol{\mu}t + \frac{1}{2}\boldsymbol{l}^\top\Sigma\,\boldsymbol{l}t^2\right)\ .$$

Ved benyttelse af momentfrembringende funktioner kan det let eftervises, at de to forskellige definitioner på flerdimensional normalfordeling er ækvivalente, når der ses bort fra det singulære tilfælde, se opgave 11.

Den flerdimensionale normlfordeling

3 Egenskaber ved normalfordelte stokastiske vektorer

Lad os betragte en affin transformation $A\boldsymbol{Y}+\boldsymbol{b}$ af $\boldsymbol{Y} \sim N_n(\boldsymbol{\mu}, \Sigma)$, hvor Σ er positiv definit, og hvor A er $q \times n$ med $\operatorname{rang} A = q \leq n$. Der gælder

$$\begin{aligned}\forall \boldsymbol{l} \neq \mathbf{0} \ : \ \boldsymbol{l}^\top(A\boldsymbol{Y}+\boldsymbol{b}) &= (\boldsymbol{l}^\top A)\boldsymbol{Y} + \boldsymbol{l}^\top \boldsymbol{b} \\ &\sim N((\boldsymbol{l}^\top A)\boldsymbol{\mu} + \boldsymbol{l}^\top \boldsymbol{b}, (\boldsymbol{l}^\top A)\Sigma(l^\top A)^\top) \\ &= N(\boldsymbol{l}^\top(A\boldsymbol{\mu}+\boldsymbol{b}), \boldsymbol{l}^\top A\Sigma A^\top \boldsymbol{l}) \ ,\end{aligned}$$

som viser, at $A\boldsymbol{Y}+\boldsymbol{b} \sim N_q(A\boldsymbol{\mu}+\boldsymbol{b}, A\Sigma A^\top)$, idet $A\Sigma A^\top$ er positiv definit, jf. opgave 12.

Heraf følger, at en vilkårlig linearkombination $\boldsymbol{a}^\top \boldsymbol{Y}$ af $\boldsymbol{Y}$'s komponenter er normalfordelt.

Ved et passende valg af A kan vi udtage en vilkårlig delvektor af $\boldsymbol{Y}$. Alle delvektorer af $\boldsymbol{Y}$, herunder også de enkelte komponenter, er altså normalfordelte.

For endimensionale normalfordelte variable gælder, at uafhængighed er ensbetydende med, at de variable er ukorrelerede. Betragter vi to normalfordelte stokastiske vektorer $\boldsymbol{Y}_1 \sim N_{n_1}(\boldsymbol{\mu}_1, \Sigma_1)$ og $\boldsymbol{Y}_2 \sim N_{n_2}(\boldsymbol{\mu}_2, \Sigma_2)$, vil vi tilsvarende vise, at uafhængighed mellem $\boldsymbol{Y}_1$ og $\boldsymbol{Y}_2$ er ensbetydende med, at $\operatorname{Cov}(\boldsymbol{Y}_1, \boldsymbol{Y}_2) = O$.

Det gælder generelt, at uafhængighed mellem stokastiske vektorer medfører, at de er ukorrelerede, jf. afsnit 1.

Antager vi omvendt, at $\operatorname{Cov}(\boldsymbol{Y}_1, \boldsymbol{Y}_2) = O$, bliver variansmatricen Σ for den stokastiske vektor $\boldsymbol{Y} = (\boldsymbol{Y}_1, \boldsymbol{Y}_2)$

$$\Sigma = \begin{bmatrix} \Sigma_1 & O \\ O & \Sigma_2 \end{bmatrix},$$

hvor Σ_1 og Σ_2 er variansmatricerne for henholdsvis $\boldsymbol{Y}_1$ og $\boldsymbol{Y}_2$. Heraf følger, at

$$\Sigma^{-1} = \begin{bmatrix} {\Sigma_1}^{-1} & O \\ O & {\Sigma_2}^{-1} \end{bmatrix}.$$

Bemærk, at

$$
\begin{aligned}
(\boldsymbol{y}-\boldsymbol{\mu})^\top \Sigma^{-1}(\boldsymbol{y}-\boldsymbol{\mu}) &= \begin{bmatrix} \boldsymbol{y}_1-\boldsymbol{\mu}_1 \\ \boldsymbol{y}_2-\boldsymbol{\mu}_2 \end{bmatrix}^\top \begin{bmatrix} {\Sigma_1}^{-1} & O \\ O & {\Sigma_2}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{y}_1-\boldsymbol{\mu}_1 \\ \boldsymbol{y}_2-\boldsymbol{\mu}_2 \end{bmatrix} \\
&= (\boldsymbol{y}_1-\boldsymbol{\mu}_1)^\top \Sigma_1^{-1}(\boldsymbol{y}_1-\boldsymbol{\mu}_1) + (\boldsymbol{y}_2-\boldsymbol{\mu}_2)^\top \Sigma_2^{-1}(\boldsymbol{y}_2-\boldsymbol{\mu}_2) \\
&= h_1(\boldsymbol{y}_1) + h_2(\boldsymbol{y}_2).
\end{aligned}
$$

For $\boldsymbol{Y}$'s tæthedsfunktion får vi derfor, at

$$
\begin{aligned}
&f_{\boldsymbol{Y}}(y_{11},\ldots,y_{1n_1},y_{21},\ldots,y_{2n_2}) \\
&\quad= (2\pi)^{-\frac{n_1+n_2}{2}}(\det\Sigma_1\det\Sigma_2)^{-\frac{1}{2}}\exp\big(-\tfrac{1}{2}(h_1(\boldsymbol{y}_1)+h_2(\boldsymbol{y}_2))\big) \\
&\quad= (2\pi)^{-\frac{n_1}{2}}(\det\Sigma_1)^{-\frac{1}{2}}\exp\big(-\tfrac{1}{2}h_1(\boldsymbol{y}_1)\big)(2\pi)^{-\frac{n_2}{2}}(\det\Sigma_2)^{-\frac{1}{2}}\exp\big(-\tfrac{1}{2}h_2(\boldsymbol{y}_2)\big) \\
&\quad= f_{\boldsymbol{Y}_1}(y_{11},\ldots,y_{1n_1})f_{\boldsymbol{Y}_2}(y_{21},\ldots,y_{2n_2}),
\end{aligned}
$$

hvoraf det ses, at $\boldsymbol{Y}_1$ og $\boldsymbol{Y}_2$ er uafhængige.

Hvis vi betragter to lineære transformationer af samme vektor $\boldsymbol{Y}$, $A_1\boldsymbol{Y}$ og $A_2\boldsymbol{Y}$, hvor A_1 og A_2 har fuld rang, ses umiddelbart, at uafhængighed mellem $A_1\boldsymbol{Y}$ og $A_2\boldsymbol{Y}$ er ensbetydende med at $A_1\Sigma A_2^\top = O$, idet

$$
\mathrm{Cov}\,(A_1\boldsymbol{Y}, A_2\boldsymbol{Y}) = O \Leftrightarrow A_1\mathrm{Cov}\,(\boldsymbol{Y},\boldsymbol{Y})\,A_2^\top = O \Leftrightarrow A_1\Sigma A_2^\top = O.
$$

Bemærk, at med $A_1\Sigma A_2^\top = O$ er systemet af rækker i A_1 ortogonalt på systemet af rækker i A_2 med hensyn til det indre produkt $\langle \boldsymbol{u},\boldsymbol{v}\rangle = \boldsymbol{u}^\top\Sigma\boldsymbol{v}$. Matricen $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$ har derfor fuld rang, dvs. $\mathrm{rang}\,A = \mathrm{rang}\,A_1 + \mathrm{rang}\,A_2 \le n$.

Specielt gælder, at to linearkombinationer af $\boldsymbol{Y}$'s komponenter, ${\boldsymbol{a}_1}^\top\boldsymbol{Y}$ og ${\boldsymbol{a}_2}^\top\boldsymbol{Y}$, er uafhængige, når og kun når $\mathrm{Cov}\,\big({\boldsymbol{a}_1}^\top\boldsymbol{Y}, {\boldsymbol{a}_2}^\top\boldsymbol{Y}\big) = 0 \Leftrightarrow {\boldsymbol{a}_1}^\top\Sigma\boldsymbol{a}_2 = 0$.

Et vigtigt resultat er, at

$$
(\boldsymbol{Y}-\boldsymbol{\mu})^\top\Sigma^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}) \sim \chi^2(n)\ ,
$$

hvilket let eftervises, se opgave 14.

Lad os til slut vise, at de stokastiske variable $\bar{Y} = \frac{1}{n}\sum_{i=1}^n Y_i$ og $S^2 = \frac{1}{n-1}\sum_{i=1}^n (Y_i-\bar{Y})^2$, hvor $\bar{Y}$ og S^2 er baseret på en normalfordelt stikprøve med middelværdi μ og varians σ^2, er uafhængige. Bemærk, at $\boldsymbol{Y} \sim N_n(\mu\mathbf{1}, \sigma^2 I)$.

Når $\mathbf{Z}$ er en ortogonal transformation af $\boldsymbol{Y}$, dvs. $\boldsymbol{Z} = C\boldsymbol{Y}$ med C ortogonal, får vi

$$
\boldsymbol{Z} \sim N_n(\mu C\mathbf{1}, \sigma^2 CIC^\top) = N_n(\mu C\mathbf{1}, \sigma^2 I)\,.
$$

Komponenterne i $\boldsymbol{Z}$ er altså uafhængige.

Benytter vi specielt en ortogonal matrix, der har vektoren $(\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}})$ som første rækkevektor, får vi

$$Z_1 = \begin{bmatrix} \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \end{bmatrix} \boldsymbol{Y} = \frac{1}{\sqrt{n}} \sum_i Y_i = \sqrt{n}\, \bar{Y} \;\Leftrightarrow\; \bar{Y} = \frac{Z_1}{\sqrt{n}}$$

og

$$\begin{aligned}(n-1)S^2 &= \sum_i (Y_i - \bar{Y})^2 = \sum_i Y_i^2 - n\bar{Y}^2 = \sum_i Z_i^2 - n\left(\frac{Z_1}{\sqrt{n}}\right)^2 \\ &= Z_2^2 + \cdots + Z_n^2 \;\Leftrightarrow\; S^2 = \frac{1}{n-1}\left(Z_2^2 + \cdots + Z_n^2\right),\end{aligned}$$

hvoraf det umiddelbart fremgår, at $\bar{Y}$ og S^2 er uafhængige.

4 En betinget fordeling

Lad $\boldsymbol{Y} \sim N_n(\boldsymbol{\mu}, \Sigma)$, og betragt delvektorene $\boldsymbol{Y}_1$ med n_1 komponenter og $\boldsymbol{Y}_2$ med n_2 komponenter, $n_1 + n_2 = n$. $\boldsymbol{Y}_1$ og $\boldsymbol{Y}_2$ er begge normalfordelte og i almindelighed afhængige. Vi kan skrive

$$(\boldsymbol{Y}_1, \boldsymbol{Y}_2) \sim N_{n_1+n_2}\left((\boldsymbol{\mu}_1, \boldsymbol{\mu}_2), \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right),$$

hvor betydningen af de benyttede betegnelser er åbenbar, fx står Σ_{12} for $\operatorname{Cov}(\boldsymbol{Y}_1, \boldsymbol{Y}_2)$.

Indfør $\boldsymbol{Z} = \boldsymbol{Y}_2 - \Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{Y}_1$, og betragt den stokastiske vektor $(\boldsymbol{Y}_1, \boldsymbol{Z})$, der fremkommer som en lineær transformation af $(\boldsymbol{Y_1}, \boldsymbol{Y_2})$. Transformationen er repræsenteret ved matricen $\begin{bmatrix} I_{n_1} & O \\ -\Sigma_{21}\Sigma_{11}^{-1} & I_{n_2} \end{bmatrix}$, som er $n \times n$ og regulær.

Ved udregning, jf. opgave 16, får vi, at

$$\operatorname{Var}((\boldsymbol{Y}_1, \boldsymbol{Z})) = \begin{bmatrix} \Sigma_{11} & O \\ O & \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{bmatrix},$$

der viser, at $\boldsymbol{Y}_1$ og $\boldsymbol{Z}$ er uafhængige.

For $\operatorname{Var}(\boldsymbol{Z})$ benyttes som regel betegnelsen $\Sigma_{22\cdot 1}$, altså

$$\Sigma_{22\cdot 1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}.$$

Bemærk, at

$$\mathrm{E}[\boldsymbol{Z}] = \mathrm{E}[\boldsymbol{Y}_2] - \Sigma_{21}\Sigma_{11}^{-1}\mathrm{E}[\boldsymbol{Y}_1] = \boldsymbol{\mu}_2 - \Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{\mu}_1,$$

hvorefter vi kan angive $\boldsymbol{Z}$'s fordeling som

$$\boldsymbol{Z} \sim N_{n_2}\left(\boldsymbol{\mu}_2 - \Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{\mu}_1, \Sigma_{22\cdot 1}\right).$$

Fra indførelsen af $\boldsymbol{Z}$ har vi, at

$$\boldsymbol{Y}_2 = \boldsymbol{Z} + \Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{Y}_1,$$

og ved at betinge med $\boldsymbol{Y}_1 = \boldsymbol{y}_1$, får vi

$$\boldsymbol{Y}_2|\boldsymbol{y}_1 = \boldsymbol{Z} + \Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{y}_1,$$

idet $\boldsymbol{Y}_1$ og $\boldsymbol{Z}$ er uafhængige. Middelværdivektor og variansmatrix for $\boldsymbol{Y}_2|\boldsymbol{y}_1$ bliver

$$\mathrm{E}\left[\boldsymbol{Y}_2|\boldsymbol{y}_1\right] = \boldsymbol{\mu}_2 + \Sigma_{21}\Sigma_{11}^{-1}(\boldsymbol{y}_1 - \boldsymbol{\mu}_1),$$

$$\mathrm{Var}\left(\boldsymbol{Y}_2|\boldsymbol{y}_1\right) = \mathrm{Var}\left(\boldsymbol{Z}\right) = \Sigma_{22\cdot 1}.$$

Fordelingen af $\boldsymbol{Y}_2|\boldsymbol{y}_1$ er altså

$$\boldsymbol{Y}_2|\boldsymbol{y}_1 \sim N_{n_2}\left(\boldsymbol{\mu}_2 + \Sigma_{21}\Sigma_{11}^{-1}(\boldsymbol{y}_1 - \boldsymbol{\mu}_1), \Sigma_{22\cdot 1}\right).$$

5 Opgaver

1. Vis, at $\mathrm{Cov}\left(\boldsymbol{X}, \boldsymbol{Y}\right) = \mathrm{E}\left[\boldsymbol{X}\boldsymbol{Y}^\top\right] - \mathrm{E}\left[\boldsymbol{X}\right]\mathrm{E}\left[\boldsymbol{Y}\right]^\top$.

2. Vis, at $\mathrm{Cov}\left(A\boldsymbol{X}, B\boldsymbol{Y}\right) = A\,\mathrm{Cov}\left(\boldsymbol{X}, \boldsymbol{Y}\right)B^\top$.

3. Vis, at $\mathrm{Var}\left(\boldsymbol{Y} + \boldsymbol{b}\right) = \mathrm{Var}\left(\boldsymbol{Y}\right)$.

4. Vis, at $M_{A\boldsymbol{Y}+\boldsymbol{b}}(\boldsymbol{t}) = \exp(\boldsymbol{t}^\top\boldsymbol{b})M_{\boldsymbol{Y}}(A^\top\boldsymbol{t})$.

5. Vis, at $\boldsymbol{X}$ og $\boldsymbol{Y}$ uafhængige $\Rightarrow$ $M_{\boldsymbol{X}+\boldsymbol{Y}}(\boldsymbol{t}) = M_{\boldsymbol{X}}(\boldsymbol{t})M_{\boldsymbol{Y}}(\boldsymbol{t})$.

6. Lad C være en ortogonal matrix, dvs. der gælder $C^\top C = I$. Vis, at systemet af rækkevektorer i C er ortonormeret, at $C^{-1} = C^\top$, og at $\det C = \pm 1$. Vis endvidere, at produktet af to ortogonale matricer giver en ortogonal matrix.

7. Lad A være en symmetrisk matrix. Vis, at A er positiv definit, når og kun når alle A's egenværdier er positive. Vink: Betragt en ortogonal substitution $\boldsymbol{x} = C\boldsymbol{y}$, hvor matricen C diagonaliserer A, dvs. $C^{-1}AC = \Lambda = \ulcorner\lambda_1 \dots \lambda_n\lrcorner$, og vis, at $\boldsymbol{x}^\top A\boldsymbol{x} = \sum_i \lambda_i y_i^2$.

8. Lad A være en positiv definit matrix, og lad C være en ortogonal matrix, som diagonaliserer A til Λ. Bestem først en matrix $\Lambda^{\frac{1}{2}}$, så $(\Lambda^{\frac{1}{2}})^2 = \Lambda$, og dernæst en matrix $A^{\frac{1}{2}}$, så $(A^{\frac{1}{2}})^2 = A$.

9. Vis, at $\det\Sigma^{\frac{1}{2}} = (\det\Sigma)^{\frac{1}{2}}$.

10. Kontroller at $f_{\boldsymbol{Y}}$ er en tæthedsfunktion, dvs. at $\int_{R^n} f_{\boldsymbol{Y}}(y_1, \ldots, y_n) dV = 1$. Vink: Benyt tæthedsfunktionen $f_{\boldsymbol{X}}(x_1, \ldots, x_n)$ for den transformerede stokastiske vektor $\boldsymbol{X}$.

11. Vis ved benyttelse af momentfrembringende funktioner, at de to definitioner på flerdimensional normalfordeling er ækvivalente, dvs. at definition 1 $\Rightarrow$ definition 2, og at definition 2 $\Rightarrow$ definition 1. (Der ses bort fra det singulære tilfælde.)

12. Vis, at når A er $n \times n$ positiv definit, og når C er $p \times n$ med $\operatorname{rang} C = p$, så er $CAC^\top$ positiv definit.

13. Eftervis, at $\boldsymbol{u}^\top \Sigma \boldsymbol{v}$, hvor Σ er positiv definit, definerer et indre produkt.

14. Vis, at $(\boldsymbol{Y} - \boldsymbol{\mu})^\top \Sigma^{-1} (\boldsymbol{Y} - \boldsymbol{\mu}) \sim \chi^2(n)$. Vink: Benyt transformationen $\boldsymbol{Y} = \Sigma^{\frac{1}{2}} \boldsymbol{X} + \boldsymbol{\mu}$.

15. Lad $Y_1, \ldots, Y_n$ være en stikprøve i en normalfordelt population. Giv et bevis for, at $S^2 \sim \frac{\sigma^2}{n-1} \chi^2(n-1)$. Vink: Betragt først $\boldsymbol{X} = \frac{1}{\sigma}(\boldsymbol{Y} - \boldsymbol{\mu})$ og dernæst $\boldsymbol{U} = C\boldsymbol{X}$ med samme C som i slutningen af afsnit 3.

16. Eftervis, at $\boldsymbol{Y_1}$ og $\boldsymbol{Z}$ defineret som i afsnit 4 er uafhængige. Vink: Husk, at $\text{Var}(A\boldsymbol{Y}) = A \text{Var}(\boldsymbol{Y}) A^\top$.

31.1.2011
Bo Rosbjerg