

Multivariate Verfahren

2. Multivariate Verteilungen

Moritz Berger

Institut für Statistik, LMU München

Sommersemester 2019

Multivariate Zufallsvariablen

Wir betrachten Zufallsvektoren

$$\mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix} \in \mathbb{R}^p,$$

mit Erwartungswertvektor

$$\mathbb{E}(\mathbf{X}) = \boldsymbol{\mu} = \mathbb{E} \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix} = \begin{pmatrix} \mathbb{E}(X_1) \\ \vdots \\ \mathbb{E}(X_p) \end{pmatrix},$$

bzw. Erwartungswertmatrix

$$\mathbb{E}(\mathbf{X}) = (\mathbb{E}(X_{ij})) .$$

Multivariate Zufallsvariablen

Die Zusammenhangsstruktur in \mathbf{X} wird durch die Kovarianz

$$\text{cov}(\mathbf{X}) = \mathbf{\Sigma} = (\sigma_{ij})$$

erfasst, wobei

$$\sigma_{ij} = \text{cov}(X_i, X_j) = \mathbb{E}(X_i - \mathbb{E}(X_i))\mathbb{E}(X_j - \mathbb{E}(X_j)) ,$$

$$\sigma_{ii} = \text{cov}(X_i, X_i) = \text{var}(X_i) = \sigma_i^2 .$$

Multivariate Zufallsvariablen

Für zwei Vektoren $\mathbf{X} = (X_1, \dots, X_p)^\top$ und $\mathbf{Y} = (Y_1, \dots, Y_q)$ gilt

$$\boldsymbol{\Sigma}_{XY} = \text{cov}(\mathbf{X}, \mathbf{Y}) = \begin{pmatrix} \text{cov}(X_1, Y_1) & \dots & \text{cov}(X_1, Y_q) \\ \vdots & \vdots & \vdots \\ \text{cov}(X_p, Y_1) & \dots & \text{cov}(X_p, Y_q) \end{pmatrix}$$

→ $\boldsymbol{\Sigma}_{XY}$ ist i.A. nicht symmetrisch!

Rechenregeln

$$\textcircled{1} \quad \mathbb{E}(\mathbf{AX} + \mathbf{b}) = \mathbf{A}\mathbb{E}(\mathbf{X}) + \mathbf{b}$$

$$\textcircled{2} \quad \text{cov}(\mathbf{X}) = E \left((\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^\top \right) =$$

$$\textcircled{3} \quad \text{cov}(\mathbf{AX} + \mathbf{b}) = \mathbf{A} \text{cov}(\mathbf{X}) \mathbf{A}^\top$$

$$\textcircled{4} \quad \text{cov}(\mathbf{X}, \mathbf{Y}) = \mathbb{E} \left((\mathbf{X} - \mathbb{E}(\mathbf{X}))(\mathbf{Y} - \mathbb{E}(\mathbf{Y}))^\top \right)$$

$$\textcircled{5} \quad \text{cov}(\mathbf{AX} + \mathbf{a}, \mathbf{BY} + \mathbf{b}) = \mathbf{A} \text{cov}(\mathbf{X}, \mathbf{Y}) \mathbf{B}^\top$$

Beachte: \mathbf{A} , \mathbf{B} , \mathbf{a} und \mathbf{b} sind konstant.

Zerlegung der Kovarianzmatrix

Es gilt:

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{\frac{1}{2}} \boldsymbol{\Sigma}^{\frac{1}{2}} \quad \text{bzw.} \quad \boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{\frac{1}{2}} \boldsymbol{\Sigma}^{\frac{T}{2}}$$

und

$$\boldsymbol{\Sigma}^{-1} = \boldsymbol{\Sigma}^{-\frac{T}{2}} \boldsymbol{\Sigma}^{-\frac{1}{2}}$$

Daraus resultiert:

$$\text{cov}(\boldsymbol{\Sigma}^{-\frac{1}{2}} \mathbf{X}) =$$

$\Rightarrow \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{X} - \mathbb{E}(\mathbf{X}))$ hat Erwartungswert $\mathbf{0}$ and Varianz \mathbf{I} .

Spektralzerlegung

Es gilt:

$$\mathbf{\Sigma} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{\top} \quad \text{und} \quad \mathbf{\Sigma}^{-1} = \mathbf{P}\mathbf{\Lambda}^{-1}\mathbf{P}^{\top},$$

wobei

$\mathbf{P} = (\mathbf{p}_1, \dots, \mathbf{p}_p)^{\top}$: Matrix der *orthonormalen* Eigenvektoren und

$\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_p)$: Diagonalmatrix der *geordneten* Eigenwerte.

Stetige Zufallsvektoren

Ein Zufallsvektor \mathbf{X} heißt stetig verteilt, falls eine Dichte

$$f(\mathbf{X}) = f(X_1, \dots, X_p) \geq 0$$

existiert mit Verteilungsfunktion

$$F(\mathbf{X}) = F(X_1, \dots, X_p) = \int_{-\infty}^{X_p} \dots \int_{-\infty}^{X_1} f(u_1, \dots, u_p) du_1, \dots, du_p,$$

wobei $0 \leq F(\mathbf{X}) \leq 1$.

Multivariate Normalverteilung

Der Zufallsvektor $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ heißt multivariat normalverteilt mit Dichtefunktion

$$f(\mathbf{X}) = \frac{1}{(2\pi)^{p/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \right\}.$$

- $\boldsymbol{\mu} \in \mathbb{R}^p$: Erwartungswert
- $\boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$: Kovarianzmatrix

→ Die Kovarianzmatrix $\boldsymbol{\Sigma}$ ist *symmetrisch* und *positiv definit*.

Damit existiert auch $\boldsymbol{\Sigma}^{-1}$.

Zusammenhang zur Spektralzerlegung

Es gilt:

$$\begin{aligned}(\mathbf{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu}) &= (\mathbf{X} - \boldsymbol{\mu})^\top \mathbf{P} \boldsymbol{\Lambda}^{-1} \mathbf{P}^\top (\mathbf{X} - \boldsymbol{\mu}) = \\ &= \mathbf{Y} \boldsymbol{\Lambda}^{-1} \mathbf{Y}^\top = \sum_{i=1}^p \frac{y_i^2}{\lambda_i} = \sum_{i=1}^p \left(\frac{y_i}{\sqrt{\lambda_i}} \right)^2\end{aligned}$$

Die Isodensiten sind damit durch folgende Gleichung bestimmt:

$$(\mathbf{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu}) = \sum_{i=1}^p \left(\frac{y_i}{\sqrt{\lambda_i}} \right)^2 = c$$

→ Dies stellt in den y -Werten ein Ellipsoid mit den Hauptachsenlängen $\sqrt{\lambda_i}$ dar!

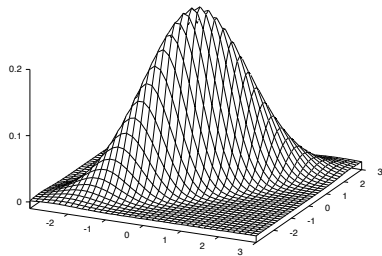
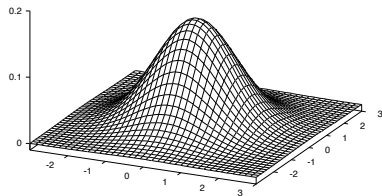
Bivariate Normalverteilung

$$f(X_1, X_2) =$$

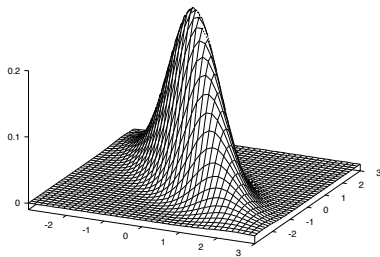
mit

$$\Sigma =$$

Beispiele



Beispiele



Eigenschaften

① $\mathbf{AX} + \mathbf{b} \sim N(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top)$

Insbesondere gilt: $X_i \sim N(\mu_i, \sigma_i^2)$, denn

② $\boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu}) \sim N_p(\mathbf{0}, \mathbf{I})$

③ $(\mathbf{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2(p)$

Bedingte Normalverteilung

Sei

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Y} \\ \mathbf{X} \end{pmatrix} \sim N \left(\boldsymbol{\mu}_Z = \begin{pmatrix} \boldsymbol{\mu}_Y \\ \boldsymbol{\mu}_X \end{pmatrix}, \boldsymbol{\Sigma}_Z = \begin{pmatrix} \boldsymbol{\Sigma}_{YY} & \boldsymbol{\Sigma}_{YX} \\ \boldsymbol{\Sigma}_{XY} & \boldsymbol{\Sigma}_{XX} \end{pmatrix} \right),$$

dann gilt für die bedingte Verteilung

$$\mathbf{Y}|\mathbf{X} \sim N \left(\boldsymbol{\mu}_{Y|X}, \boldsymbol{\Sigma}_{Y|X} \right),$$

mit

$$\boldsymbol{\mu}_{Y|X} = \boldsymbol{\mu}_Y + \boldsymbol{\Sigma}_{YX} \boldsymbol{\Sigma}_{XX}^{-1} (\mathbf{X} - \boldsymbol{\mu}_X),$$

$$\boldsymbol{\Sigma}_{Y|X} = \boldsymbol{\Sigma}_{YY} - \boldsymbol{\Sigma}_{YX} \boldsymbol{\Sigma}_{XX}^{-1} \boldsymbol{\Sigma}_{XY}.$$

Wishart-Verteilung

Seien $\mathbf{X}_1, \dots, \mathbf{X}_m \stackrel{i.i.d.}{\sim} N_p(\mathbf{0}, \mathbf{\Sigma})$, dann ist die Matrix

$$\mathbf{M} = \sum_{i=1}^m \mathbf{X}_i \mathbf{X}_i^{\top} = \mathbf{X}^{\top} \mathbf{X} \in \mathbb{R}^{p \times p}$$

wishart-verteilt mit Parametern $\mathbf{\Sigma}$ und m .

- $\mathbf{M} \sim W_p(\mathbf{\Sigma}, m)$
- Falls $p = 1$, so gilt $\mathbf{M} = \sum_{i=1}^m X_i^2 \sim \chi^2(m)$, wobei die Komponenten $X_i \sim N(0, \sigma^2)$

Hotellings T^2 -Verteilung

Seien $\mathbf{d} \sim N_p(\mathbf{0}, \mathbf{I})$ und $\mathbf{M} \sim W_p(\mathbf{I}, m)$ unabhängig, dann ist

$$u = m\mathbf{d}^\top \mathbf{M}^{-1} \mathbf{d}$$

Hottellings T^2 -verteilt mit Parametern p und m .

- $u \sim T^2(p, m)$
- Falls $p = 1$, so ergibt sich die $F(1, m)$ -Verteilung
- Seien $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ und $\mathbf{M} \sim W_p(\boldsymbol{\Sigma}, m)$, dann erhält man:

Wilks' Λ -Verteilung

Seien $\mathbf{A} \sim W_p(\mathbf{I}, m)$ und $\mathbf{B} \sim W_p(\mathbf{I}, n)$ unabhängig, dann ist

$$\Lambda = \frac{\det(\mathbf{A})}{\det(\mathbf{A} + \mathbf{B})}$$

Wilks' Λ -verteilt mit Parametern p , m und n .

- $\Lambda \sim \Lambda(p, m, n)$
- Falls $p = 1$, so gilt $A \sim \chi^2(m)$ und $B \sim \chi^2(n)$ und damit erhält man: $\Lambda \sim B(m/2, n/2)$