Berechnung von Kovarianzellipsen

Nikolai Nawri

Die Wahrscheinlichkeitsdichte für zwei normalverteilte Variablen ist gegeben durch[1]

$$\Phi(\boldsymbol{x}) = k \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})^T B\,(\boldsymbol{x}-\boldsymbol{a})\right] \,, \tag{1}$$

wobei $\boldsymbol{x} = (x_1, x_2)^T$ der Vektor der Wertepaare und $\boldsymbol{a} = (a_1, a_2)^T$ der Vektor der entsprechenden Erwartungswerte ist. Die (2×2)-Matrix B ist die Inverse der Kovarianzmatrix

$$C = \begin{pmatrix} \sigma_1^2 & \mathrm{cov}(x_1,x_2) \\ \mathrm{cov}(x_1,x_2) & \sigma_2^2 \end{pmatrix} \,, \tag{2}$$

wobei σ_1 und σ_2 die Standardabweichungen der Größen x_1 und x_2 sind. Matrix B ist dann gegeben durch

$$B = C^{-1} = \frac{1}{\sigma_1^2\sigma_2^2 - \mathrm{cov}(x_1,x_2)^2} \begin{pmatrix} \sigma_2^2 & -\mathrm{cov}(x_1,x_2) \\ -\mathrm{cov}(x_1,x_2) & \sigma_1^2 \end{pmatrix} \,. \tag{3}$$

Mit dem Korrelationskoeffizienten

$$\rho = \frac{\mathrm{cov}(x_1,x_2)}{\sigma_1\sigma_2} \tag{4}$$

lässt sich B auch schreiben als

$$B = \frac{1}{1-\rho^2} \begin{pmatrix} \sigma_1^{-2} & -\frac{\rho}{\sigma_1\sigma_2} \\ -\frac{\rho}{\sigma_1\sigma_2} & \sigma_2^{-2} \end{pmatrix} \,. \tag{5}$$

Auf Kurven mit gleicher Wahrscheinlichkeitsdichte Φ muss der Exponent in (1) konstant sein. Es gilt dann

$$(\boldsymbol{x}-\boldsymbol{a})^T B\,(\boldsymbol{x}-\boldsymbol{a}) = \frac{1}{1-\rho^2}\left[\frac{(x_1-a_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-a_1)(x_2-a_2)}{\sigma_1\sigma_2} + \frac{(x_2-a_2)^2}{\sigma_2^2}\right] = c \,, \tag{6}$$

mit einer Konstanten c. Setzt man $c = 1$, ergibt sich

$$\frac{1}{1-\rho^2}\left[\frac{(x_1-a_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-a_1)(x_2-a_2)}{\sigma_1\sigma_2} + \frac{(x_2-a_2)^2}{\sigma_2^2}\right] = 1 \,. \tag{7}$$

Das ist die spezielle Form einer Ellipse mit Mittelpunkt (a_1, a_2), die in Bezug auf das Koordinatensystem (x_1, x_2) um einen Winkel α gedreht ist (siehe Abbildung 1). Durch eine Verschiebung des Koordinatenursprungs auf den Mittelpunkt der Ellipse, $x'_1 = x_1 - a_1$ und $x'_2 = x_2 - a_2$, erhält man

$$\frac{1}{1-\rho^2}\left[\frac{x_1'^2}{\sigma_1^2} - 2\rho\frac{x'_1x'_2}{\sigma_1\sigma_2} + \frac{x_2'^2}{\sigma_2^2}\right] = 1 \,. \tag{8}$$

Durch eine Drehung um den Winkel α,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} \,, \tag{9}$$

[1] siehe: Fahrmeir *et al.*, 1996: Multivariate statistische Verfahren

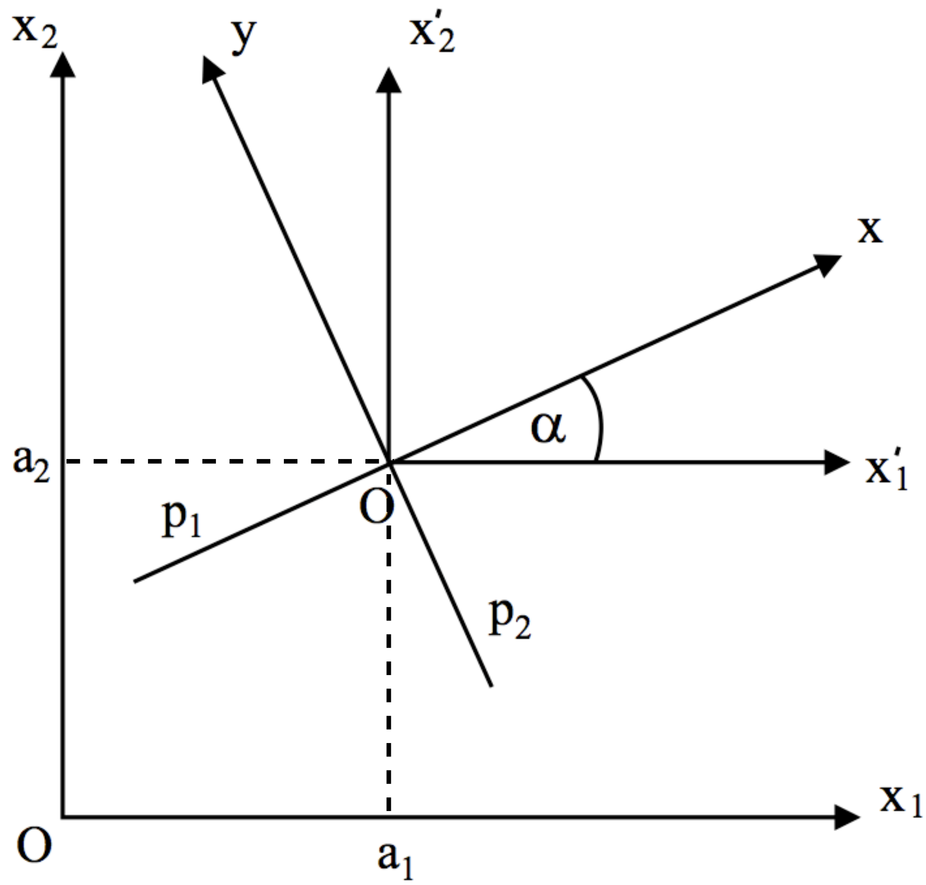

Abbildung 1: Die verwendeten Koordinatensysteme.

kann Ellipsengleichung (8) in das Hauptachsensystem transformiert werden. Die Transformationsmatrix

$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \tag{10}$$

ist quadratisch, und hat daher die Eigenschaft $A^{-1} = A^T$. Die Rücktransformation ist also gegeben durch

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad . \tag{11}$$

In den neuen Koordinaten (x, y) lautet die Ellipsengleichung (8) dann

$$\begin{aligned} \frac{1}{1-\rho^2}\left[x^2\left(\frac{\cos^2\alpha}{\sigma_1^2} - \frac{2\rho\sin\alpha\cos\alpha}{\sigma_1\sigma_2} + \frac{\sin^2\alpha}{\sigma_2^2}\right) + y^2\left(\frac{\sin^2\alpha}{\sigma_1^2} + \frac{2\rho\sin\alpha\cos\alpha}{\sigma_1\sigma_2} + \frac{\cos^2\alpha}{\sigma_2^2}\right)\right] \\ + \frac{2xy}{1-\rho^2}\left(\frac{\rho\sin^2\alpha}{\sigma_1\sigma_2} - \frac{\rho\cos^2\alpha}{\sigma_1\sigma_2} - \frac{\sin\alpha\cos\alpha}{\sigma_1^2} + \frac{\sin\alpha\cos\alpha}{\sigma_2^2}\right) = 1 \ . \end{aligned} \tag{12}$$

Andererseits ist in diesem Koordinatensystem eine Ellipse allgemein gegeben durch

$$\frac{x^2}{p_1^2} + \frac{y^2}{p_2^2} = 1 \ , \tag{13}$$

mit den Halbachsen p_1 und p_2. Durch Koeffizientenvergleich zwischen (12) und (13) erhält man

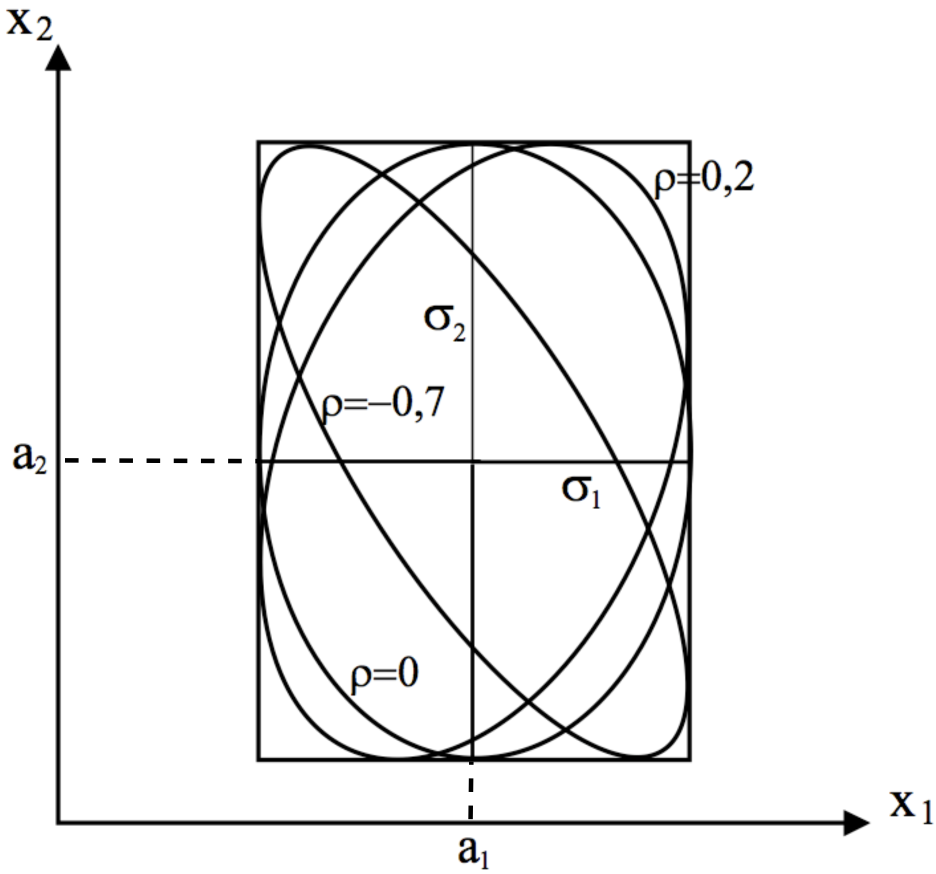

Abbildung 2: Kovarianzellipsen für gleiche Standardabweichungen und verschiedene Korrelationskoeffizienten.

dann

$$p_1^2 = \left(1-\rho^2\right)\left(\frac{\cos^2\alpha}{\sigma_1^2} - \frac{2\rho\sin\alpha\cos\alpha}{\sigma_1\sigma_2} + \frac{\sin^2\alpha}{\sigma_2^2}\right)^{-1} \tag{14}$$

$$p_2^2 = \left(1-\rho^2\right)\left(\frac{\sin^2\alpha}{\sigma_1^2} + \frac{2\rho\sin\alpha\cos\alpha}{\sigma_1\sigma_2} + \frac{\cos^2\alpha}{\sigma_2^2}\right)^{-1} \quad . \tag{15}$$

Ausserdem muss für beliebige (x, y) der gemischte Ausdruck in (12) verschwinden. Aus dieser Bedingung lässt sich dann der Rotationswinkel ableiten. Er ist gegeben durch

$$\alpha = \frac{1}{2}\arctan\left(\frac{2\rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2}\right) \quad . \tag{16}$$

Die Ellipse (13) heißt Kovarianzellipse einer Normalverteilung zweier Variablen. In Abbildung 2 sind drei Ellipsen mit gleichen Standardabweichungen σ_1 und σ_2 dargestellt. Sie liegen alle innerhalb eines Rechtecks, das durch die Standardabweichungen und die Erwartungswerte (a_1, a_2) bestimmt ist. Sie berühren dieses Rechteck in vier Punkten. Abhängig vom Korrelationskoeffizienten ρ variiert der Rotationswinkel α zwischen $\pm\frac{\pi}{2}$ und ist positiv für Rotationen im Uhrzeigersinn.

Für unabhängige Variablen geht der Korrelationskoeffizient gegen Null. Die Punktwolke füllt dann näherungsweise einen kreisförmigen Bereich aus. Korrelieren dagegen die Variablen sehr stark miteinander ($\rho \approx \pm 1$), ergibt sich eine Gerade, die diagonal innerhalb des Rechtecks liegt.

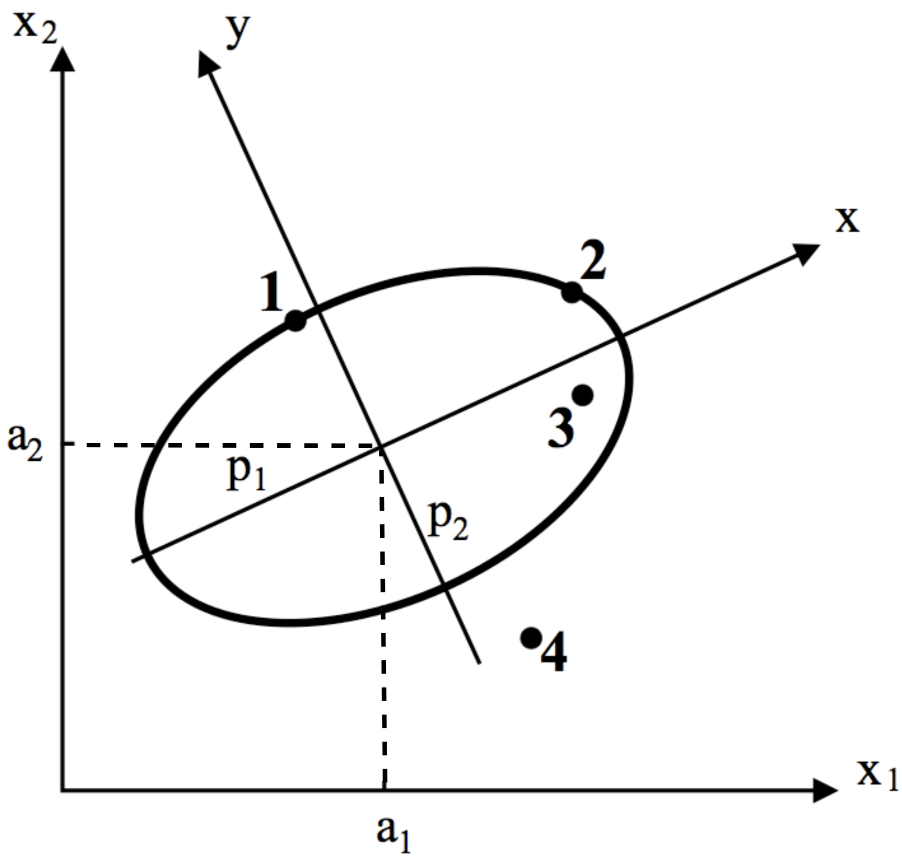

Abbildung 3: Relative Wahrscheinlichkeit für verschiedene Punkte nach einer Gaußverteilung zweier Variablen.

Isolinien mit anderen Wahrscheinlichkeitsdichten ($c \neq 1$) sind ebenfalls Ellipsen und konzentrisch zur Kovarianzellipse. Sie liegen innerhalb (ausserhalb) für größere (kleinere) Wahrscheinlichkeiten (siehe Abbildung 3). Die Punkte P_1 und P_2 haben die gleiche Wahrscheinlichkeit ($\Phi(P_1) = \Phi(P_2) = \Phi_0$), obwohl der Abstand von P_1 zum Mittelpunkt geringer ist als der Abstand von P_2 zum Mittelpunkt. Dagegen ist P_3 wahrscheinlicher ($\Phi(P_3) > \Phi_0$) und P_4 unwahrscheinlicher ($\Phi(P_4) < \Phi_0$).

Als Parametrisierung für die Ellipsen (13) in dem (x, y)-Koordinatensystem kann nun

$$x = r(\phi)\cos\phi \tag{17}$$
$$y = r(\phi)\sin\phi \tag{18}$$

gewählt werden, mit $0 \leq \phi \leq 2\pi$ und

$$r(\phi) = \left(\frac{\cos^2\phi}{p_1^2} + \frac{\sin^2\phi}{p_2^2}\right)^{-\frac{1}{2}} . \tag{19}$$

In parametrisierter Form sind die ursprünglichen Variablen dann gegeben durch

$$x_1(\phi) = r(\phi)(\cos\alpha\cos\phi - \sin\alpha\sin\phi) + a_1 \tag{20}$$
$$x_2(\phi) = r(\phi)(\sin\alpha\cos\phi - \cos\alpha\sin\phi) + a_2 . \tag{21}$$