Berechnung von Kovarianzellipsen

Nikolaj Nawri

8. Dezember 2004

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

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

wobei $\vec{x} = (x_1, x_2)^T$ der Vektor der Wertepaare und $\vec{a} = (a_1, a_2)^T$ der Vektor der entsprechenden Erwartungswerte ist. Die (2x2) - Matrix B ist die Inverse der Kovarianzmatrix

$$C = \begin{pmatrix} {\sigma^2}_1 & cov(x_1, x_2) \\ cov(x_1, x_2) & {\sigma^2}_2 \end{pmatrix} \quad .$$

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

$$B = C^{-1} = \frac{1}{\sigma_1^2\sigma_2^2 - cov(x_1, x_2)^2} \begin{pmatrix} {\sigma^2}_1 & cov(x_1, x_2) \\ cov(x_1, x_2) & {\sigma^2}_2 \end{pmatrix} \quad .$$

Mit dem Korrelationskoeffizient

$$\varrho = \frac{cov(x_1, x_2)}{\sigma_1\sigma_2} \tag{2}$$

lässt sich B auch schreiben als

$$B = \frac{1}{1-\varrho} \begin{pmatrix} {\sigma^{-2}}_1 & -\frac{\varrho}{\sigma_1\sigma_2} \\ -\frac{\varrho}{\sigma_1\sigma_2} & \sigma_2^{-2} \end{pmatrix} \quad .$$

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

$$(\vec{x}-\vec{a})^T B(\vec{x}-\vec{a}) = \frac{1}{1-\varrho^2}\left[\frac{(x_1-a_1)^2}{\sigma_1^2} - 2\varrho\frac{(x_1-a_1)(x_2-a_2)}{\sigma_1\sigma_2} + \frac{(x_2-a_2)^2}{\sigma_2^2}\right] = c \quad .$$

Setzt man $c = 1$, ergibt sich

[1] siehe: L. Fahrmeir et al. - Multivariate statistische Verfahren (1996)

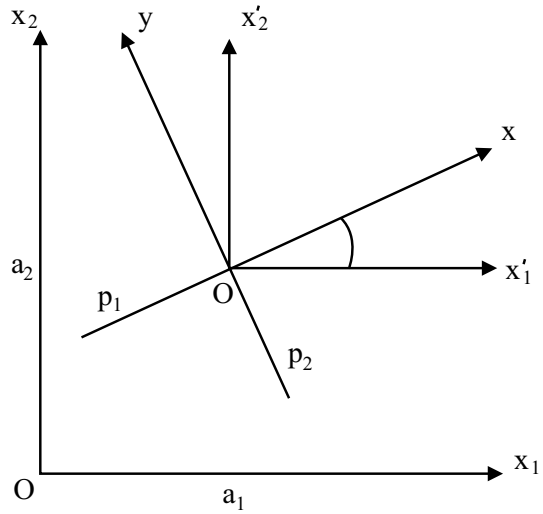


Abbildung 1: Die verwendeten Koordinatensysteme

$$\frac{1}{1-\varrho^2}\left[\frac{(x_1-a_1)^2}{\sigma_1^2}-2\varrho\frac{(x_1-a_1)(x_2-a_2)}{\sigma_1\sigma_2}+\frac{(x_2-a_2)^2}{\sigma_2^2}\right]=c \quad . \tag{3}$$

Das ist die spezielle Form einer Ellipse mit dem Mittelpunkt (a_1, a_2), die in Bezug auf die Koordinatenachse (x_1, x_2) um einen Winkel α gedreht ist. Durch eine Verschiebung des Koordinatenursprungs auf den Mittelpunkt der Ellipse, mit $x_1' = x_1 - a_1$ und $x_2' = x_2 - a_2$, erhält man

$$\frac{1}{1-\varrho^2}\left[\frac{x_1'^2}{\sigma_1^2}-2\varrho\frac{x_1'^2x_2'^2}{\sigma_1^2\sigma_2^2}+\frac{x_2'^2}{\sigma_2^2}\right]=1 \quad . \tag{4}$$

Durch eine Drehung um den Winkel α gelangt man dann in das Hauptachsensystem der Kovarianzellipse. Die dazugehörige Koordinatentransformation lautet

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

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

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{5}$$

In den neuen Koordinaten x und y lautet dann die Ellipsengleichung (4)

$$\frac{1}{1-\varrho^2}\left[\frac{(x\cos\alpha - y\sin\alpha)^2}{\sigma_1^2} + \frac{(x\sin\alpha + y\cos\alpha)^2}{\sigma_2^2}\right) -$$

$$-\frac{2\varrho(x\cos\alpha - y\sin\alpha)(x\sin\alpha + y\cos\alpha)}{\sigma_1^2\sigma_2^2(1-\varrho^2)} =$$

$$= \frac{1}{1-\varrho^2}\left[x^2\left(\frac{\cos^2\alpha}{\sigma_1^2} - \frac{2\varrho\sin\alpha\cos\alpha}{\sigma_1\sigma_2} + \frac{\sin^2\alpha}{\sigma_2^2}\right) + y^2\left(\frac{sin^2\alpha}{\sigma_1^2} + \frac{2\varrho\sin\alpha\cos\alpha}{\sigma_1\sigma_2} + \frac{\cos^2\alpha}{\sigma_2^2}\right)\right] +$$

$$+\frac{2xy}{1-\varrho^2}\underbrace{\left(\frac{\varrho\sin^2\alpha}{\sigma_1\sigma_2} - \frac{\varrho\cos^2\alpha}{\sigma_1\sigma_2} - \frac{\sin\alpha\cos\alpha}{\sigma_1^2} + \frac{sin\alpha\cos\alpha}{\sigma_2^2}\right)}_{\stackrel{!}{=}0} = 1 \quad .$$

Andererseits ist in diesem Koordinatensystem eine Ellipse allgemein gegeben durch

$$\frac{x^2}{p_1^2} + \frac{y^2}{p_2^2} = 1 \tag{6}$$

mit den Halbachsen p_1 und p_2. Durch Koeffizientenvergleich erhält man dann

$$p_1^2 = \frac{1-\varrho^2}{\left(\frac{\cos^2\alpha}{\sigma_1^2} - \frac{2\varrho\sin\alpha\cos\alpha}{\sigma_1\sigma_2} + \frac{\sin^2\alpha}{\sigma_2^2}\right)} \tag{7}$$

$$p_2^2 = \frac{1-\varrho^2}{\left(\frac{\sin^2\alpha}{\sigma_1^2} + \frac{2\varrho\sin\alpha\cos\alpha}{\sigma_1\sigma_2} + \frac{\cos^2\alpha}{\sigma_2^2}\right)} \tag{8}$$

$$\alpha = \frac{1}{2}\arctan\left(\frac{2\varrho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2}\right) \quad . \tag{9}$$

Die Ellipse mit den beschriebenen Eigenschaften heißt Kovarianzellipse einer Normalverteilung zweier Variablen. In Abbildung 2 sind drei Ellipsen dargestellt, die alle die gleichen Standartabweichungen σ_1 und σ_2 haben, so dass alle Kovarianellipsen innerhalb eines Rechtecks liegen, das durch σ_1 und σ_2 und den Punkt (a_1, a_2) bestimmt ist. Sie berühren dieses Rechteck in vier Punkten. Durch die Drehung um den Winkel α gelangt man dann in das Hauptachsensystem der Ellipsen. Abhänigig von dem ϱ liegt der Winkel zwischen $-90°$ und $90°$. Ist ϱ negativ, ist der Winkel α auch negativ und die Kovarianzellipse ist nach links geneigt.

Da der Korrelationskoeffizient (2) nichts anderes ist als die auf die Standardabweichung der beiden Variablen normierte Kovarianz, geht die Kovarianz

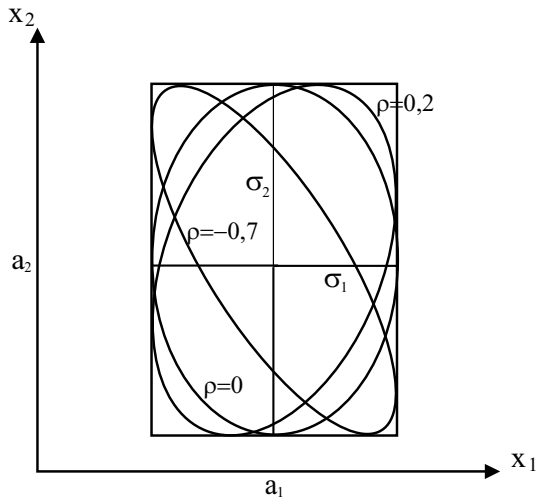


Abbildung 2: Beispiel

gegen Null ($\varrho = 0$), wenn zwischen den beiden Variablen kein Zusammenhang besteht. Die Punktwolke füllt also näherungsweise einen kreisförmigen Bereich aus. Der Drehwinkel α ist 0°. Korrelieren dagegen die Variablen sehr stark miteinander, ergibt sich eine Gerade ($\varrho = \pm 1$), die in dem Rechteck dann eine der beiden Diagonalen darstellt.

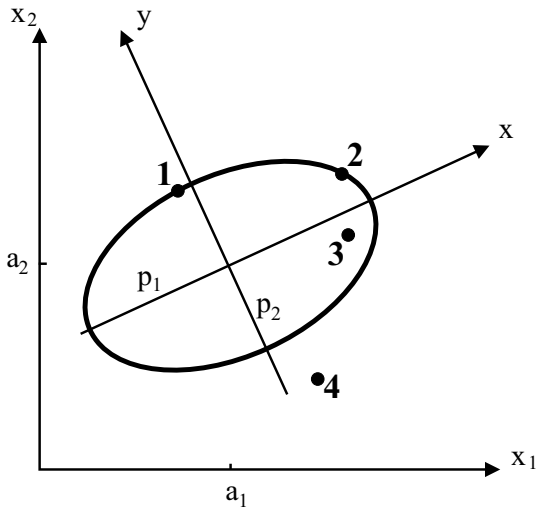


Abbildung 3: Relative Wahrscheinlichkeit für verschiedene Punkte nach einer Gaußverteilung zweier Variablen

Aus Gl. 4 wird klar, dass auch andere Linien gleicher Wahrscheinlichkeit

($c \neq 1$) Ellipsen sind. Sie sind ähnlich und konzentrisch mit der Kovarianzellipse und liegen innerhalb (außerhalb) für kleinere (größere) Wahrscheinlichkeiten. (Siehe Abbildung 3) Die Punkte 1 und 2, die auf der Ellipse liegen, stehen für die gleiche Wahrscheinlichkeit ($P_1 = P_2 = P_e$), obwohl der Abstand des Punktes 1 zum Mittelpunkt geringer ist, als der Abstand von Punkt 2 zum Mittelpunkt. Dagegen ist Punkt 3 wahrscheinlicher und Punkt 4 unwahrscheinlicher ($P_4 > P_e, P_4 < P_e$).

Als Parametrisierung für die Ellipsen in dem (x,y)-Koordinatensystem wählt man nun

$$\begin{aligned} x &= r(\Phi)\cos\Phi \\ y &= r(\Phi)\sin\Phi \end{aligned}$$

mit $0 \leq \Phi \leq 2\pi$ und $r(\Phi) \geq 0$. Dieser Ansatz ergibt mit der Ellipsengleichung (6):

$$r^2(\Phi)\left(\frac{\cos^2\Phi}{p_1^2} + \frac{\sin^2\Phi}{p_2^2}\right) = 1 \tag{10}$$

oder

$$r(\Phi) = \left(\frac{\cos^2\Phi}{p_1^2} + \frac{\sin^2\Phi}{p_2^2}\right)^{-\frac{1}{2}} \quad . \tag{11}$$

Mit der Koordinatentransformation (5) findet dann eine Drehung um $-\Phi$ in das (x_1', x_2')-Koordinatensystem statt, und mit $x_1 = x_1' + a_2, x_2 = x_2' + a_2$ erhält man dann die Ellipsengleichung (3) in der Kurvendarstellung (12, 13) mit Parameter Φ:

$$x_1(\Phi) = r(\Phi)(\cos\alpha\cos\Phi - \sin\alpha\sin\Phi) + a_1 \tag{12}$$

$$x_2(\Phi) = r(\Phi)(\sin\alpha\cos\Phi - \cos\alpha\sin\Phi) + a_2 \tag{13}$$