Bivariate Normalverteilung

Die Dichtefunktion der gemeinsamen Verteilung für korrelierte normalverteilte Zufallsvariable ist nicht unmittelbar verständlich. Die Definition dieser multivariaten Normalverteilung als Ausgangspunkt für eine Bearbeitung zu nehmen, ist schlechter Stil, der auch bei vielen anderen Themen beobachtbar ist. Er verleitet zu dem Eindruck, Mathematik sei nur etwas für Eingeweihte.

Seien X und Y unabhängig und $\mathcal{N}(0,1)$-verteilt.
Wir basteln uns auf naheliegende Weise zwei korrelierte Zufallsvariable:

$$\begin{aligned} X^* &= X \\ Y^* &= \rho X + \sqrt{1-\rho^2}\, Y \end{aligned}$$

Die Abhängigkeiten für $-1 \le \rho \le 1$ treten deutlich hervor, betrachte $\rho = 1$, $\rho = 0$.
(Für standardisierte Zufallsvariable X^*, Y^* mit der Korrelation ρ gilt stets dieser Zusammenhang.)

$$\begin{aligned} V(X^*) = V(Y^*) &= 1 \qquad \text{beachte} \quad V(aX+bY) = a^2V(X) + b^2V(Y) \\ Cov(X^*,Y^*) &= \rho \qquad Cov(X, aX+bY) = aCov(X,X) + bCov(X,Y) = aV(X) \end{aligned}$$

Die Kovarianz stimmt hier mit dem Korrelationskoeffizienten überein.

Um die gemeinsame Verteilung zu ermitteln, ist das Integral

$$P(X^* \le x^*, Y^* \le y^*) = \iint\limits_{\substack{x \le x^* \\ \rho x + \sqrt{1-\rho^2}\, y \le y^*}} f(x,y)\ dx\,dy \qquad \text{mit} \qquad f(x,y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}$$

auszuwerten.

Dies gelingt mit der Transformation (Umkehrabbildung) $\qquad (x, \frac{y-\rho x}{\sqrt{1-\rho^2}}) \longleftarrow (x, y)$

Allgemein gilt für eine Transformation (siehe Verschiedenes, krummlinige Koordinaten)

$$(u,v) \longrightarrow (\,x(u,v),\ y(u,v)\,)$$

$$dA^* = \left| \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} \right| du\,dv \qquad \text{und in diesem Fall} \qquad \left| \begin{vmatrix} 1 & 0 \\ \dots & \frac{1}{\sqrt{1-\rho^2}} \end{vmatrix} \right| = \frac{1}{\sqrt{1-\rho^2}}.$$

$$P(X^* \le x^*, Y^* \le y^*) = \int_{-\infty}^{y^*}\int_{-\infty}^{x^*} \frac{1}{\sqrt{1-\rho^2}} f(x, \frac{y-\rho x}{\sqrt{1-\rho^2}})\ dx\,dy \qquad \text{mit} \qquad f(x,y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}$$

Einsetzen und leichte Umformungen führen zur gesuchten Dichte:

$$g(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)}$$

Da x und y vertauscht werden können, ist sie symmetrisch zur Winkelhalbierenden.
Die Höhenlinien $x^2 - 2\rho xy + y^2 = const$ sind Ellipsen.

Bivariate Normalverteilung Fortsetzung

Die Matrizenschreibweise der Dichte mit Hilfe der Kovarianzmatrix sollte nun auch einsichtig sein.

$$g(x,y) = \frac{1}{2\pi\sqrt{\mid \Sigma \mid}}\, e^{-\frac{1}{2}(x|y)\,\Sigma^{-1}\binom{x}{y}} \qquad *$$

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

$$\sqrt{\mid \Sigma \mid} = \sqrt{1-\rho^2}$$

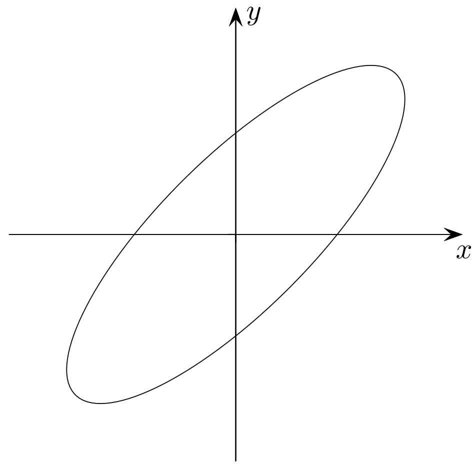
$$\Sigma^{-1} = \frac{1}{1-\rho^2}\begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}$$

Die Höhenlinien von $g(x,y)$ sind Ellipsen (siehe Hauptachsentransformation) mit der Gleichung:

$$(x \mid y)\,\Sigma^{-1}\binom{x}{y} = const$$

Die Eigenwerte von Σ^{-1} lauten: $\frac{1}{1+\rho}$, $\frac{1}{1-\rho}$ und

die Eigenvektoren $\binom{1}{1}$, $\binom{-1}{1}$. Hieraus ist ersichtlich:
Je größer ρ, umso schmaler sind die Ellipsen.
Ihre Halbachsen liegen auf den Winkelhalbierenden.



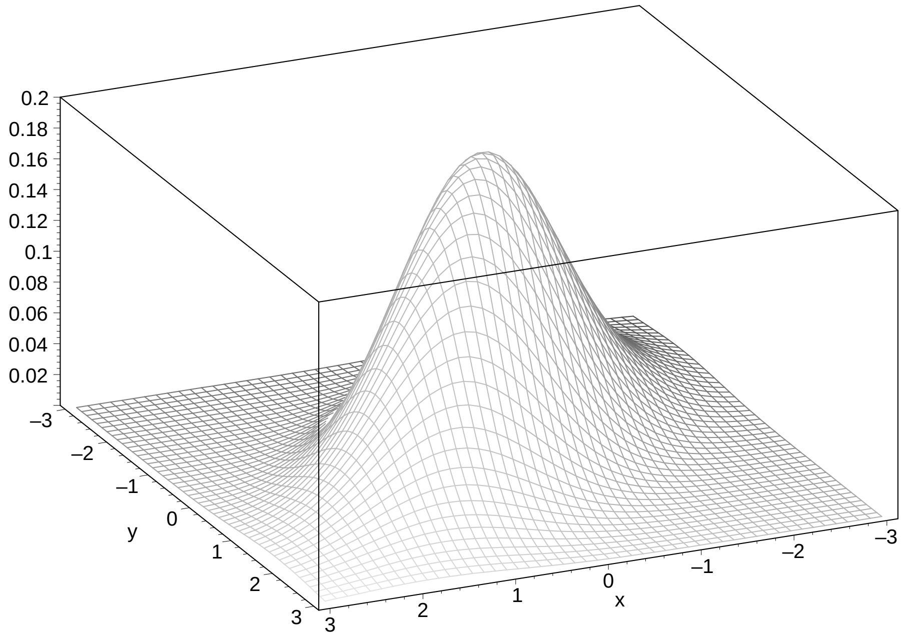
Seien nun X und Y $\mathcal{N}(0,1)$-verteilt, ρ der Korrelationskoeffizient, $g(x,y)$ die Dichte der gemeinsamen Verteilung. Beim Übergang zu den Zufallsvariablen $\sigma_x X$ und $\sigma_y Y$ ändert sich die Dichte $g(x,y)$ gemäß $h(x,y) = \frac{1}{\sigma_x\sigma_y}\, g(\frac{x}{\sigma_x}, \frac{y}{\sigma_y})$. Der Korrelationskoeffizient bleibt erhalten, beachte $\rho(X,Y) = \rho(aX, bY)$.

Die Kovarianzmatrix lautet nun:

$$\Sigma = \begin{pmatrix} \sigma_x^2 & Cov(X,Y) \\ Cov(X,Y) & \sigma_y^2 \end{pmatrix}, \qquad Cov(X,Y) = \rho\sigma_x\sigma_y, \quad \sqrt{\mid \Sigma \mid} = \sqrt{\sigma_x^2\sigma_y^2 - Cov(X,Y)^2} = \sigma_x\sigma_y\sqrt{1-\rho^2}$$

$$\Sigma^{-1} = \frac{1}{\sigma_x^2\sigma_y^2(1-\rho^2)}\begin{pmatrix} \sigma_y^2 & -Cov(X,Y) \\ -Cov(X,Y) & \sigma_x^2 \end{pmatrix} = \frac{1}{1-\rho^2}\begin{pmatrix} \frac{1}{\sigma_x^2} & -\frac{\rho}{\sigma_x\sigma_y} \\ -\frac{\rho}{\sigma_x\sigma_y} & \frac{1}{\sigma_y^2} \end{pmatrix}$$

Den allgemeinen Fall (gehe von $*$ aus) erhalten wir schließlich durch eine Verschiebung, so dass der Ursprung in $(\mu_x \mid \mu_y)$ übergeht.

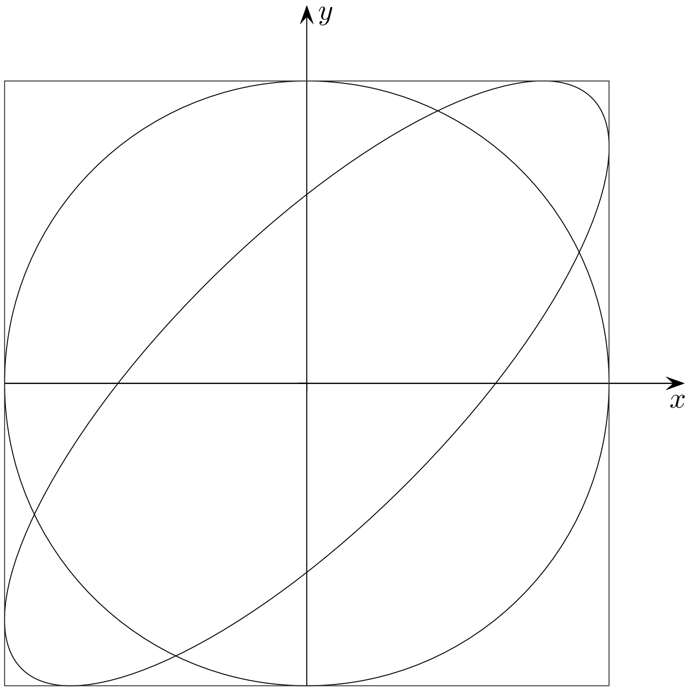

0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
–3
–2
–1
y
0
1
2
3
3
2
1
0
x
–1
–2
–3

Bilder von Kreisen

Bei der Abbildung

$$
\begin{aligned}
x^* &= x \\
y^* &= \rho x + \sqrt{1-\rho^2}\, y
\end{aligned}
$$

gehen Kreise in Ellipsen über.
Das kann direkt eingesehen werden.



Hierzu ermitteln wir das Bild der Kreispunktmenge $(x \mid \pm\sqrt{r^2 - x^2})$, $\;-r \le x \le r$.

$$
\begin{aligned}
y^* &= \rho x^* \pm \sqrt{1-\rho^2}\,\sqrt{r^2 - {x^*}^2} \\
&\cdots \qquad \text{umgestellt, quadriert, zusammengefasst, umbenannt in } x,y \\
x^2 - 2\rho xy + y^2 &= r^2(1-\rho^2)
\end{aligned}
$$

Das ist die Gleichung der um 45° gedrehten Ellipse $\frac{x^2}{1+\rho} + \frac{y^2}{1-\rho} = r^2$.
Aus ihr sind die Hauptachsen $\; a = r\sqrt{1+\rho},\; b = r\sqrt{1-\rho}$ zu ersehen.
Die Umkehrabbildung (inverse Drehmatrix) der Drehung lautet:

$$
\begin{aligned}
x^* &= \frac{\sqrt{2}}{2}\,(x+y) \\
y^* &= \frac{\sqrt{2}}{2}\,(-x+y)
\end{aligned}
$$

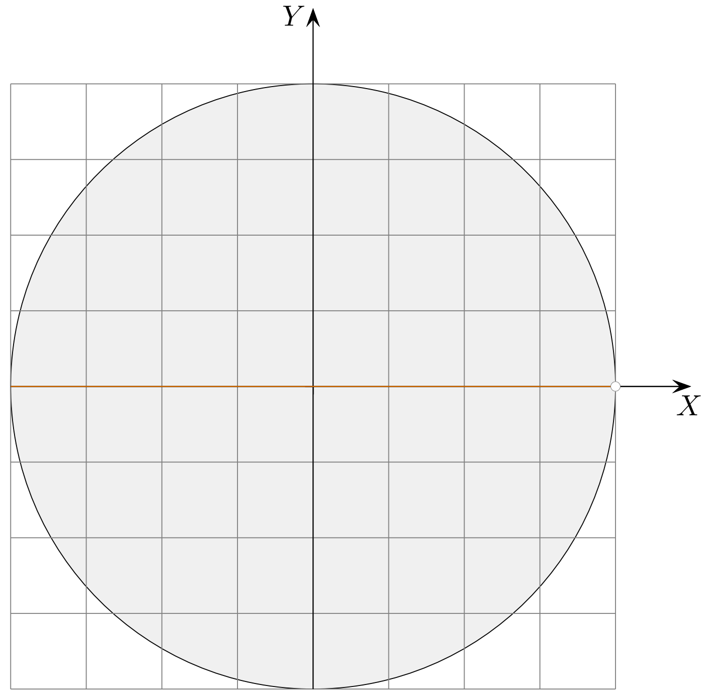
(Die Existenz des Quadrats ist noch nicht geklärt.)

Die Abbildung

$$
\begin{aligned}
x^* &= \rho y + \sqrt{1-\rho^2}\, x \\
y^* &= y
\end{aligned}
$$

der Kreispunktmenge $(\pm\sqrt{r^2 - y^2} \mid y)$, $\;-r \le y \le r$, erzeugt die gleiche gedrehte Ellipse.

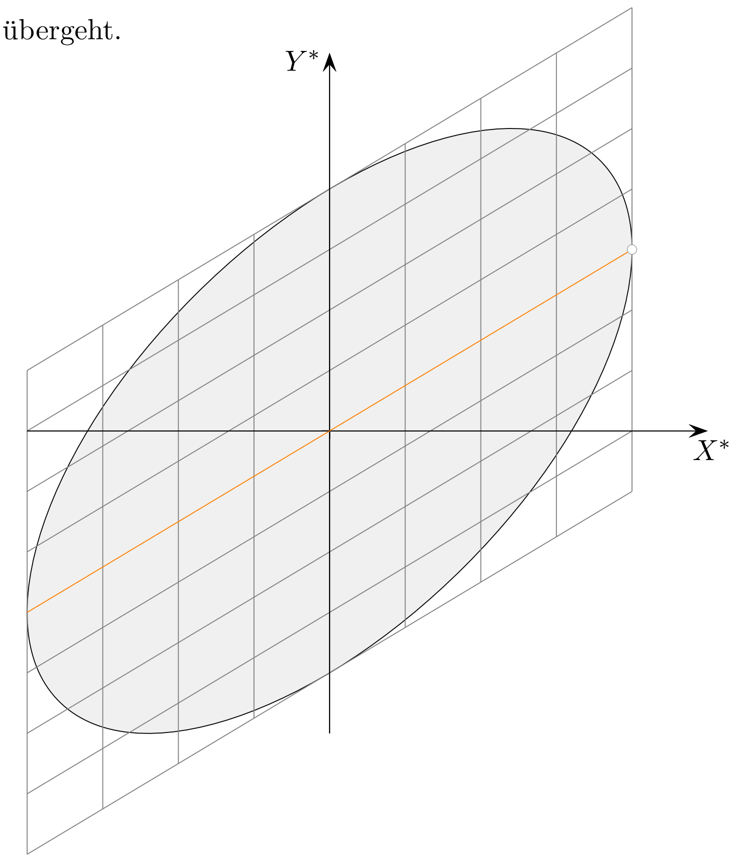
Geometrie der Abbildung



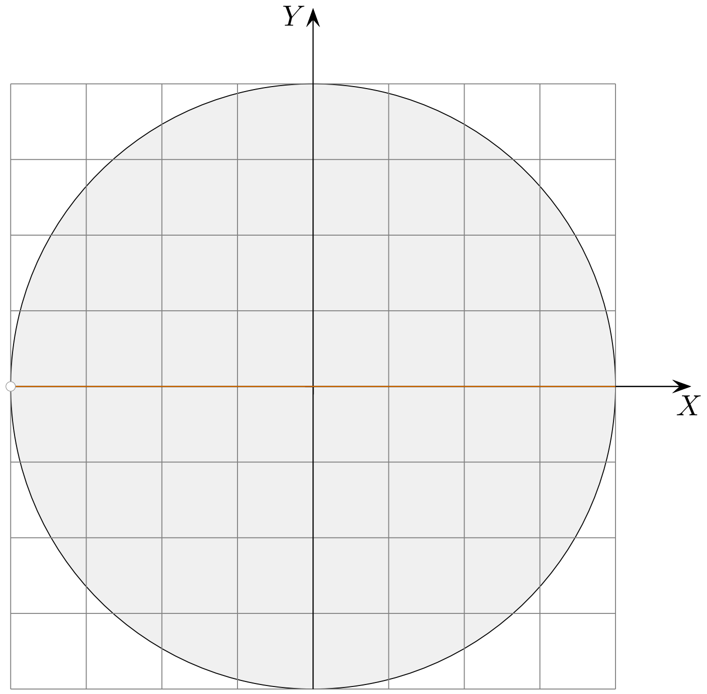
Die Abbildung (hier $\rho = 0{,}6$)

$$\begin{aligned} x^* &= x \\ y^* &= \rho x + \sqrt{1-\rho^2}\, y \end{aligned}$$

entpuppt sich als gescherte Stauchung,
wobei die x-Achse in die Regressionsgerade übergeht.
Diese halbiert die vertikalen Ellipsenlinien
(wie es auch zu erwarten ist).



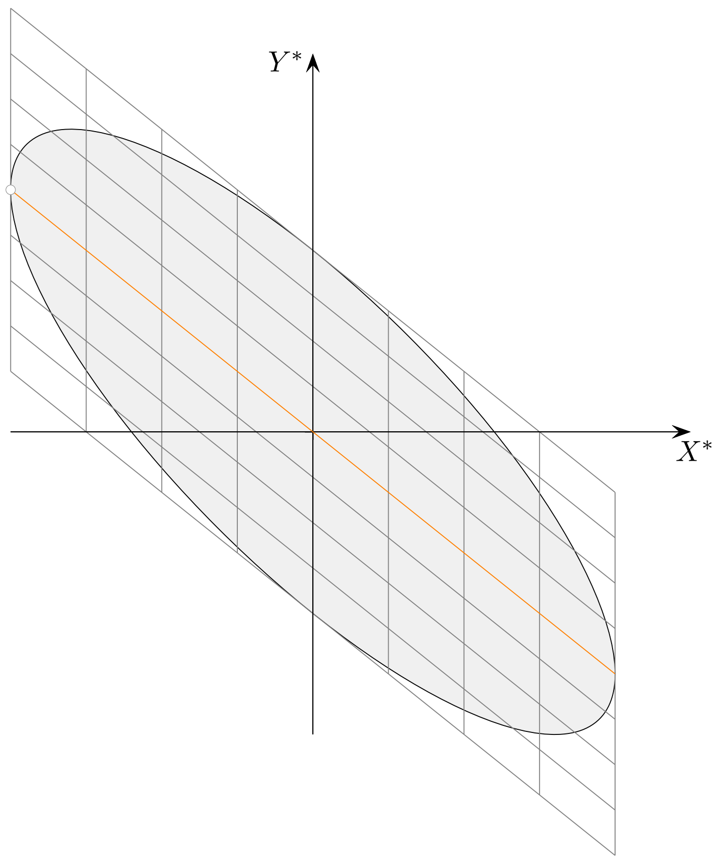
Geometrie der Abbildung



$$\rho = -0{,}8$$

$$x^* = x$$

$$y^* = \rho x + \sqrt{1-\rho^2}\, y$$

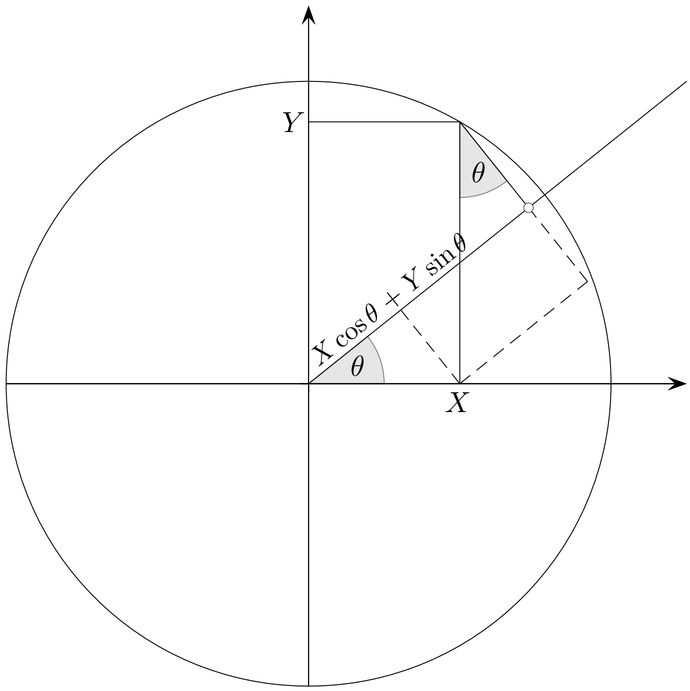


Geometrische Veranschaulichung der bivariaten Normalverteilung

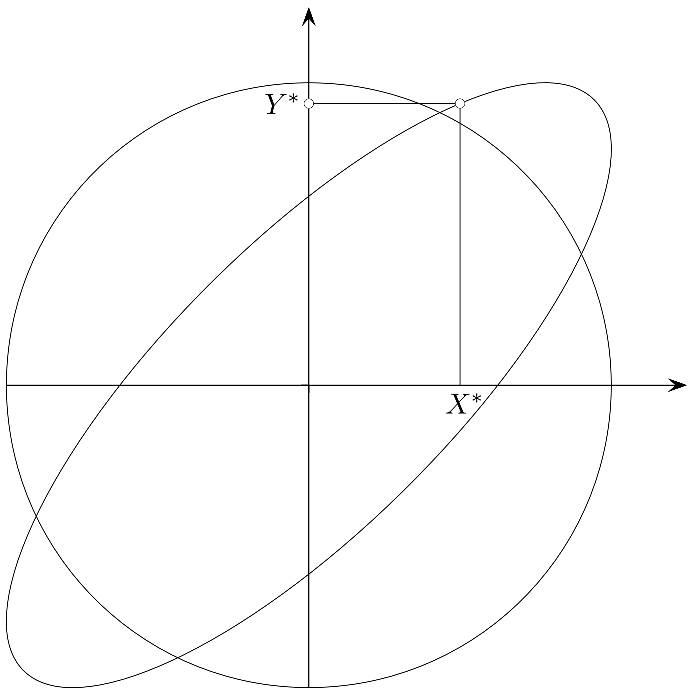
X und Y sind unabhängig und $\mathcal{N}(0,1)$-verteilt.
Wir veranschaulichen uns:

$$X^* = X$$
$$Y^* = \rho X + \sqrt{1-\rho^2}\,Y$$

Zu $-1 \leq \rho \leq 1$ gibt es ein θ mit $\cos\theta = \rho$.



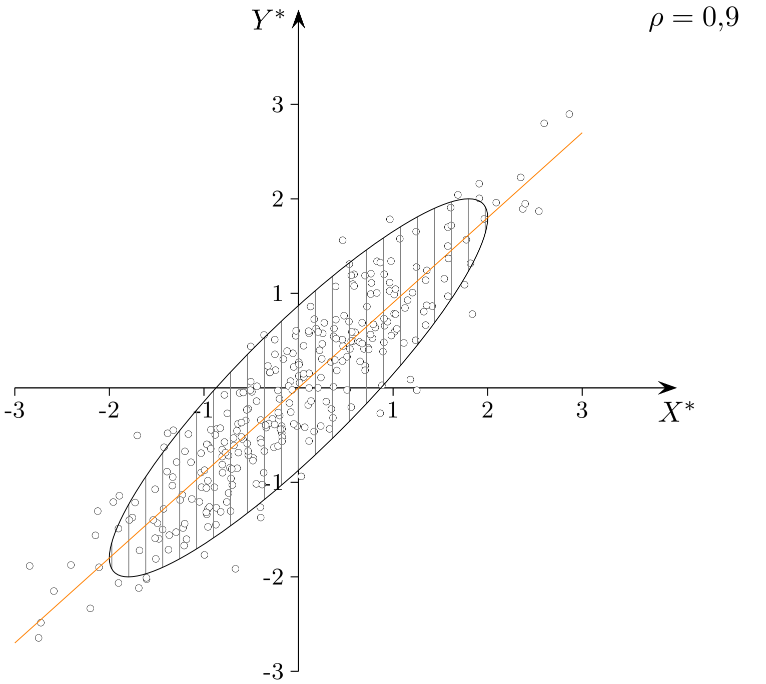
Tragen wir $Y^* = X\cos\theta + Y\sin\theta$ auf der y-Achse ab, so entsteht folgende Grafik:

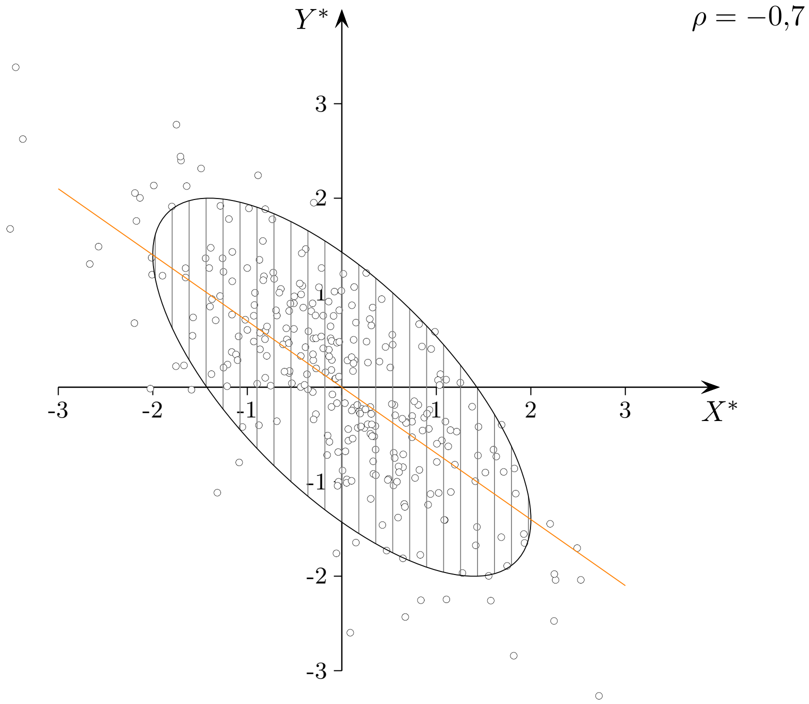


Simulation

$$X^* = X$$
$$Y^* = \rho X + \sqrt{1-\rho^2}\, Y$$

X und Y sind unabhängig und $\mathcal{N}(0,1)$-verteilt.
Die Bilder des Kreises mit dem Radius 2 wurden gezeichnet.

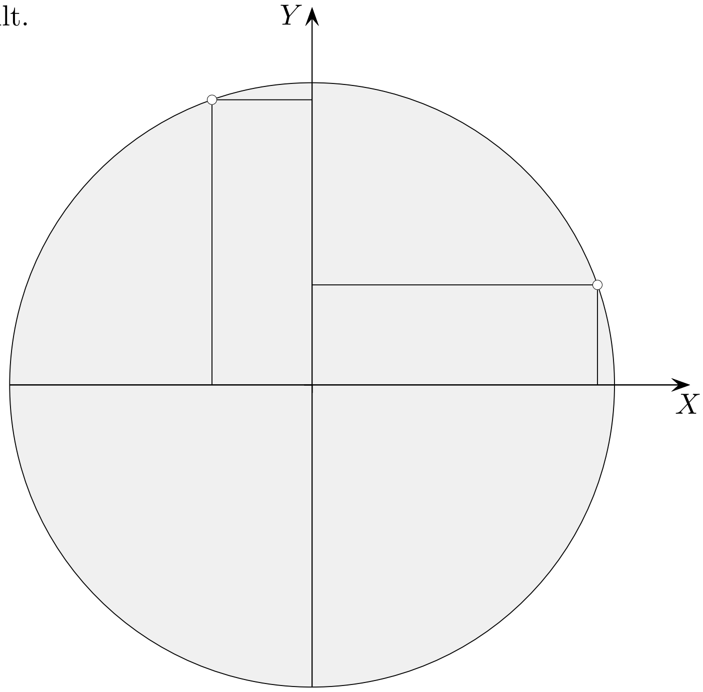




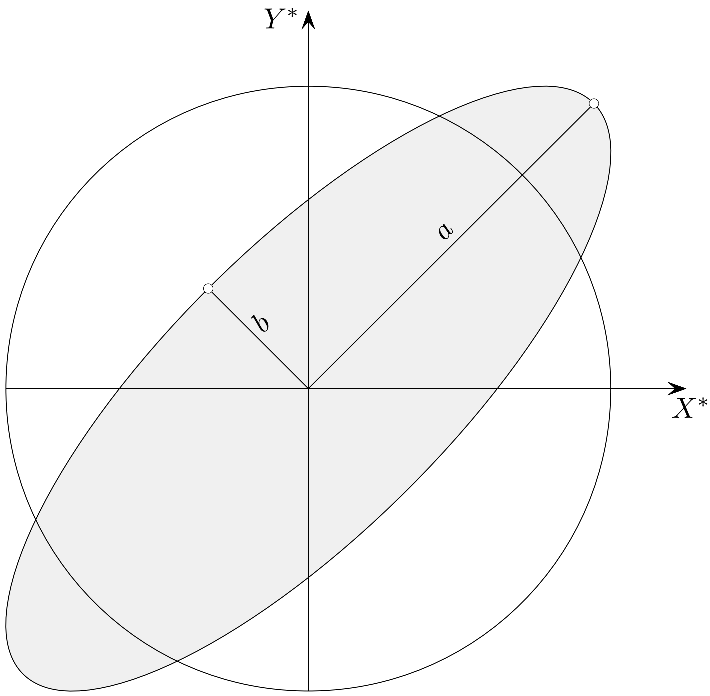
Hauptachsen

X und Y sind unabhängig und $\mathcal{N}(0,1)$-verteilt.

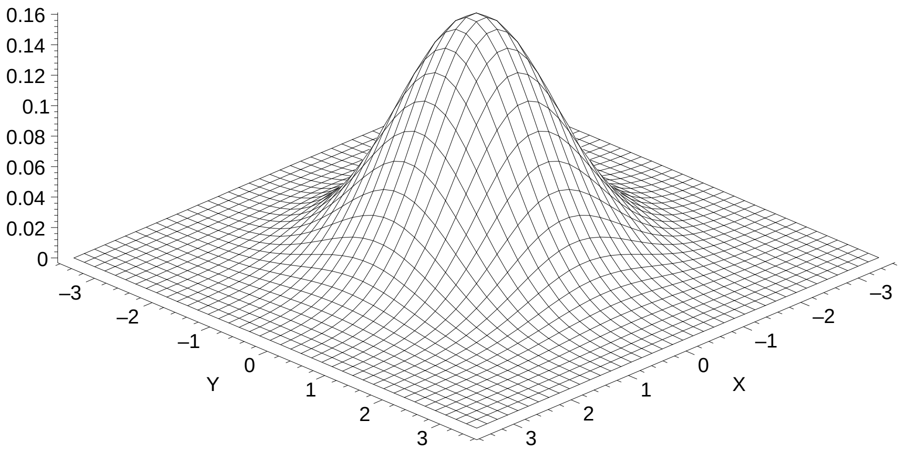
$$X^* = X$$
$$Y^* = \rho X + \sqrt{1-\rho^2}\,Y$$
$$\cos\theta = \rho$$



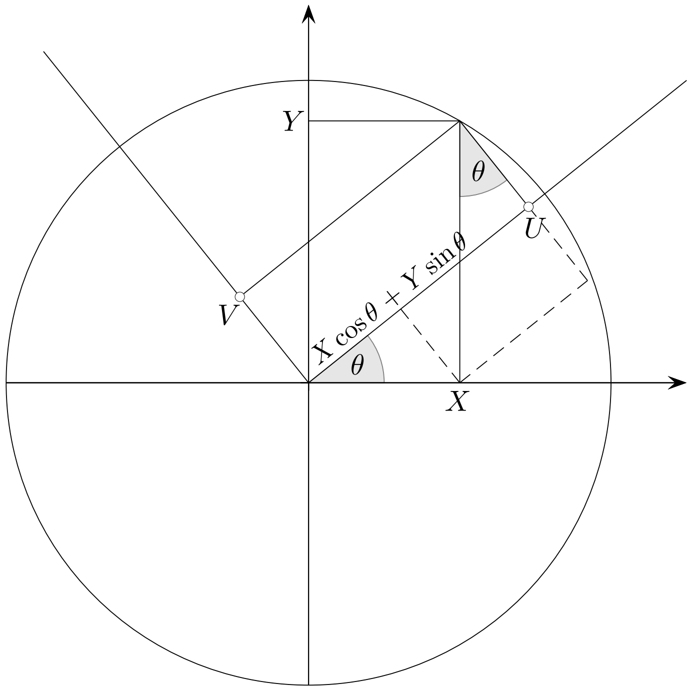
Bei der Abbildung $(X,Y) \longrightarrow (X^*,Y^*)$ des Einheitskreises liegen die Bilder der Punkte $(\cos\frac{\theta}{2}|\sin\frac{\theta}{2})$ und $(\cos\frac{\theta}{2}+\frac{\pi}{2}|\sin\frac{\theta}{2}+\frac{\pi}{2})$ auf den Winkelhalbierenden.
Nach einer längeren Rechnung (Pythagoras, trigonometrische Identitäten, Wurzel-Regeln) erhalten wir das überraschend einfache Ergebnis $a = \sqrt{1+\rho}$, $b = \sqrt{1-\rho}$.



Linearkombination von X und Y



Seien X und Y unabhängig und $\mathcal{N}(0,1)$-verteilt.
Die gemeinsame Dichte $f(x,y) = \dfrac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}$ ist das Produkt der Dichten $f_X(x)$ und $f_Y(y)$.



Mit X und Y (beide $\mathcal{N}(0,1)$-verteilt) ist für $\alpha^2 + \beta^2 = 1$ auch $\alpha X + \beta Y$ $\mathcal{N}(0,1)$-verteilt.

Hierzu betrachten wir für (X,Y) die Koordinaten (U,V) in einem um θ gedrehten Koordinatensystem. Für die Zufallsvariable U gilt $U = X\cos\theta + Y\sin\theta$, siehe Grafik.

Die Dichte von U kann mit $f_U(u) = \int\limits_{-\infty}^{-\infty} f(u,v)\,dv$ (Randdichte) ermittelt werden.

Da die gemeinsame Dichte drehsymmetrisch ist, stimmen $f_U(u)$ und $f_X(u)$ überein.

Summe von normalverteilten Zufallsvariablen

Seien X und Y unabhängig und $\mathcal{N}(\lambda, \sigma^2)$- bzw. $\mathcal{N}(\mu, \tau^2)$-verteilt.
Dann ist $X + Y$ $\mathcal{N}(\lambda + \mu, \tau^2 + \tau^2)$-verteilt, d. h. Erwartungswerte und Varianzen addieren sich.

Die Zufallsvariaben $U = \frac{X - \lambda}{\sigma}$ und $V = \frac{Y - \mu}{\tau}$ sind standardnormalverteilt (siehe Voraussetzung).

Dies trifft auch für $W = \alpha U + \beta V$ mit $\alpha^2 = \frac{\sigma^2}{\sigma^2 + \tau^2}$ und $\beta^2 = \frac{\tau^2}{\sigma^2 + \tau^2}$ zu, da $\alpha^2 + \beta^2 = 1$ gilt.

Eingesetzt und umgestellt erhalten wir die Behauptung:

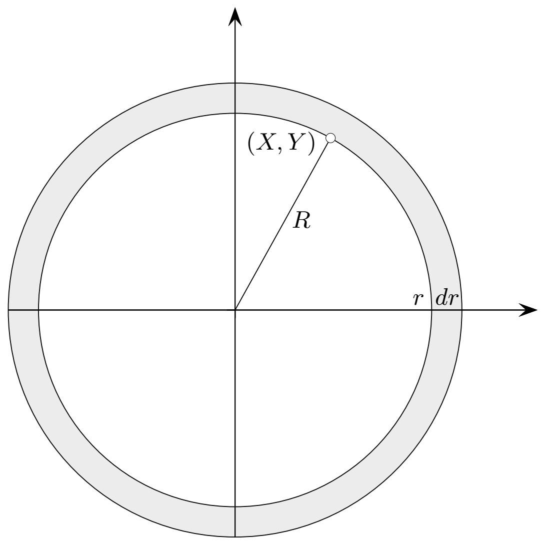
$$X + Y = (\lambda + \mu) + \sqrt{\sigma^2 + \tau^2}\, W$$

Rayleigh-Verteilung

Seien X und Y unabhängig und $\mathcal{N}(0,1)$-verteilt.

Die gemeinsame Dichte lautet: $f(x,y) = \dfrac{1}{2\pi}\, e^{-\frac{1}{2}(x^2+y^2)}$.

Wir untersuchen die Verteilung von $R = \sqrt{X^2+Y^2}$.
Mit ihr können die Wahrscheinlichkeiten von konzentrischen Kreisflächen leicht ermittelt werden.



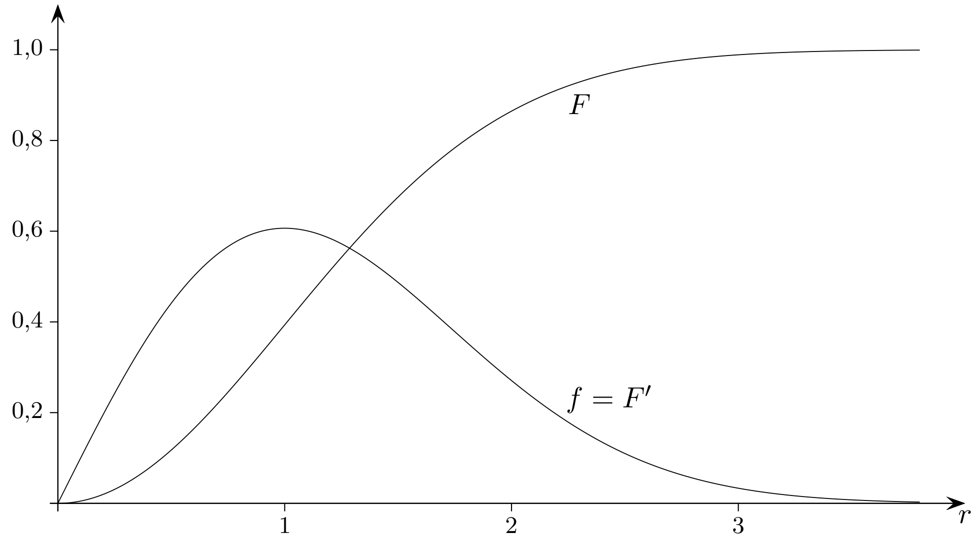
Das infinitesimale Volumenelement beträgt $V_{\circledcirc} = 2\pi r dr \cdot \dfrac{1}{2\pi}\, e^{-\frac{1}{2}r^2} = r\, e^{-\frac{1}{2}r^2} dr$.

Die Dichte von R lautet somit $f(r) = r\, e^{-\frac{1}{2}r^2}$, $r \geq 0$.

$F(r) = 1 - e^{-\frac{1}{2}r^2}$ ist dann die Verteilungsfunktion.

$F(2) = 0{,}865$

$2{,}45 = F^{-1}(0{,}95)$



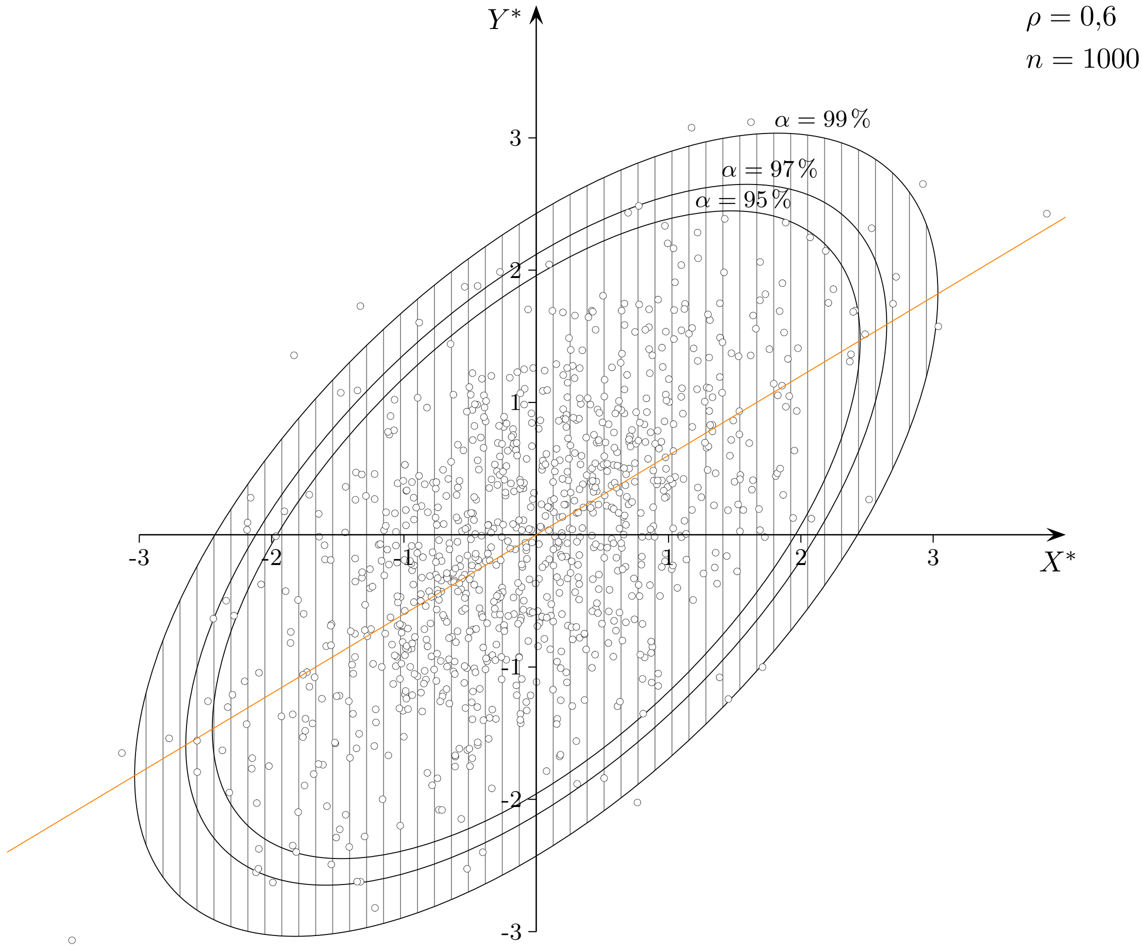
Simulation

$$X^* = X$$
$$Y^* = \rho X + \sqrt{1-\rho^2}\,Y$$

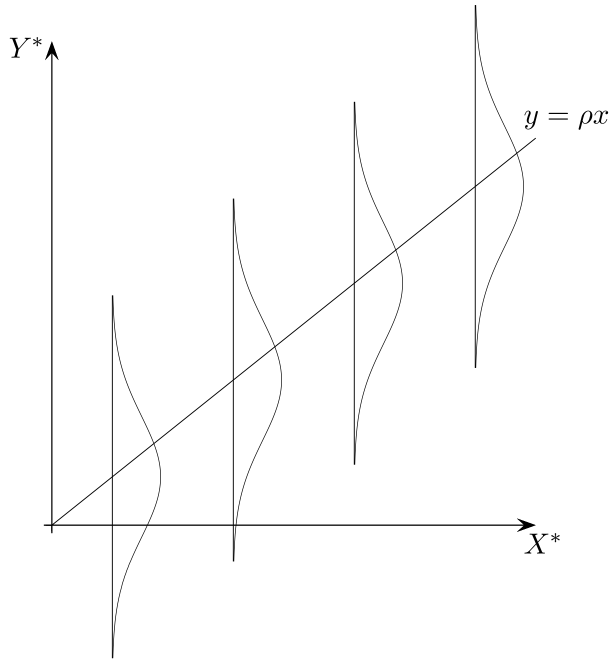
X und Y sind unabhängig und $\mathcal{N}(0,1)$-verteilt.
Die Bilder der Kreise mit den Wahrscheinlichkeiten α wurden gezeichnet.

Hauptachsen (siehe S. 4)

$$a = F^{-1}(\alpha)\sqrt{1+\rho}$$
$$b = F^{-1}(\alpha)\sqrt{1-\rho}$$



Dichte der bivariaten Normalverteilung



$$X^* = X$$
$$Y^* = \rho X + \sqrt{1-\rho^2}\, Y$$

Die Dichte

$$g(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)}$$

kann auch auf einfache Weise hergeleitet werden.

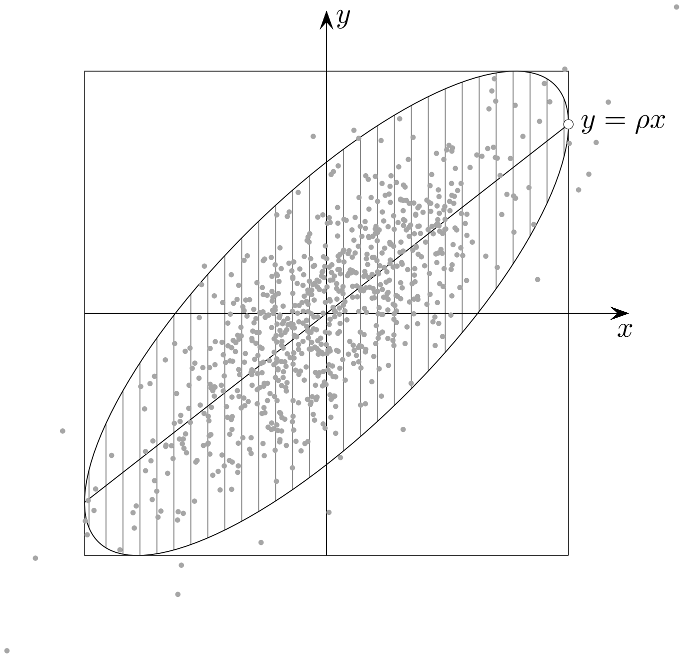
Hierzu ist zu berücksichtigen, dass X^* standardnormalverteilt mit der Dichte $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ ist und dass die bedingte Verteilung (x ist gegeben) den Erwartungswert $\mu = \rho x$ und die Standardabweichung $\sigma = \sqrt{1-\rho^2}$ hat.

Die Dichte der bedingten Verteilung lautet somit $\psi_x(y) = \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}(y-\rho x)^2}$.

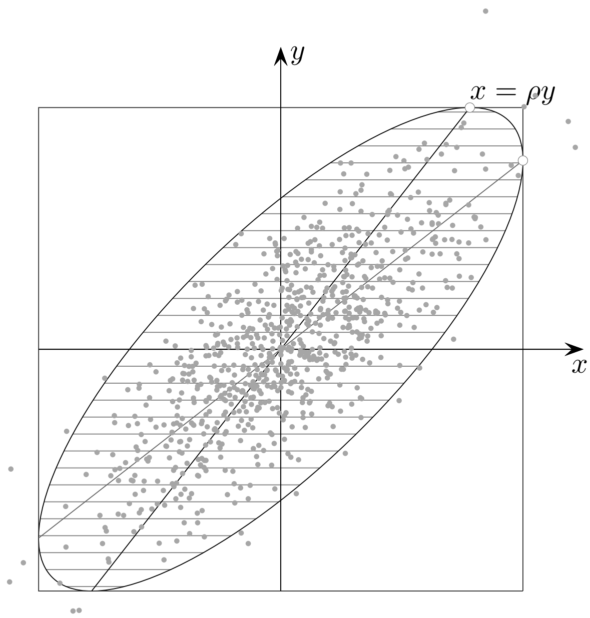
Die Dichte der bivariaten Normalverteilung ergibt sich nun offensichtlich als Produkt $\varphi(x)\psi_x(y)$. Dieses umgeformt stimmt mit $g(x,y)$ überein.

Bedingte Wahrscheinlichkeiten ($X = x$ liegt vor) werden - wie gewohnt - mit der Normalverteilung $\mathcal{N}(\rho x, 1-\rho^2)$ berechnet. Um die Wahrscheinlichkeit eines Ereignisses E von zwei korrelierenden Zufallsvariablen U und V zu ermitteln, kann E durch die beiden Standardisierungen U^* und V^* und weiter mit Hilfe der obigen Abbildungsgleichungen durch die $\mathcal{N}(0,1)$-verteilten Zufallsvariablen X und Y ausgedrückt werden, für die Unabhängigkeit und Rotationssymmetrie der gemeinsamen Dichte vorliegt.
Das Integrieren der bivariaten Dichte wird so vermieden.

Beste Prognosen

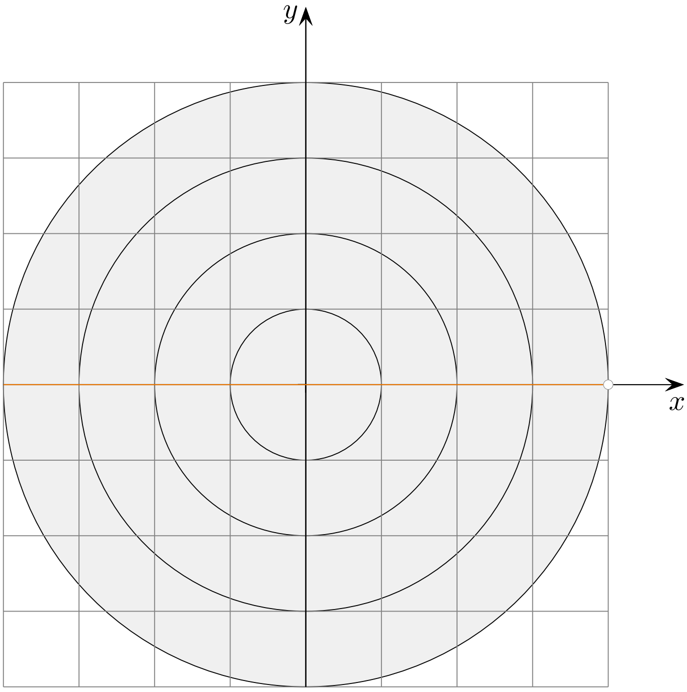


Falls x gegeben ist, so ist der Wert der Regressionsgeraden $y = \rho x$ die beste Prognose,
da dieser y-Wert die vertikalen Ellipsenschnitte halbiert.



Falls y gegeben ist, so ist der Wert $y = \rho x$ die beste Prognose,
da dieser y-Wert die horizontalen Ellipsenschnitte halbiert.
Die bedingten Wahrscheinlichkeiten ($Y = y$ liegt vor) werden mit
der Normalverteilung $\mathcal{N}(\rho y, 1 - \rho^2)$ berechnet.

Dichte der bivariaten Normalverteilung



Die Abbildung (hier $\rho = 0{,}6$)

$$x^* = x$$
$$y^* = \rho x + \sqrt{1-\rho^2}\, y$$

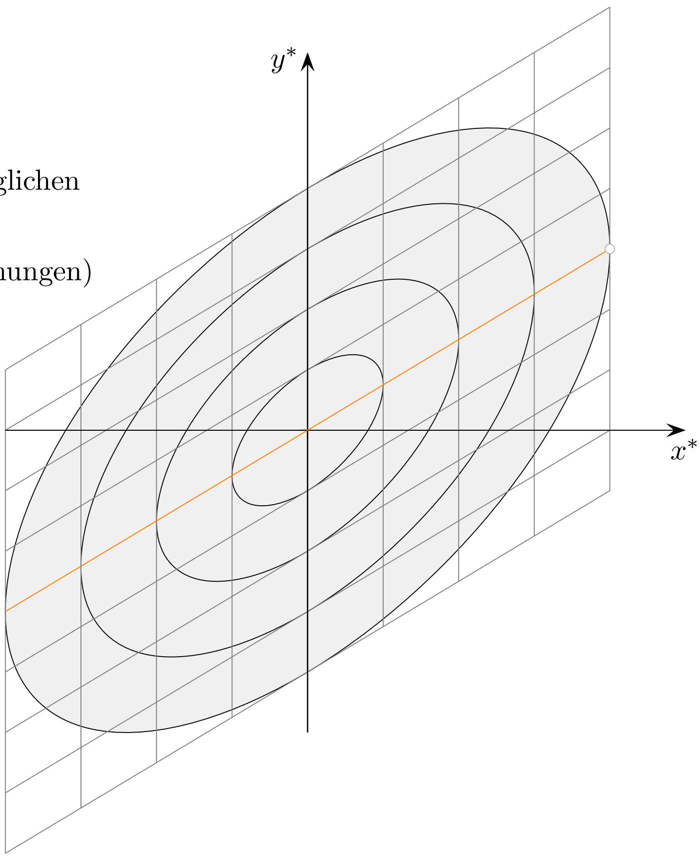
setzt sich aus einer Stauchung und einer Scherung zusammen.

Die Umkehrabbildung

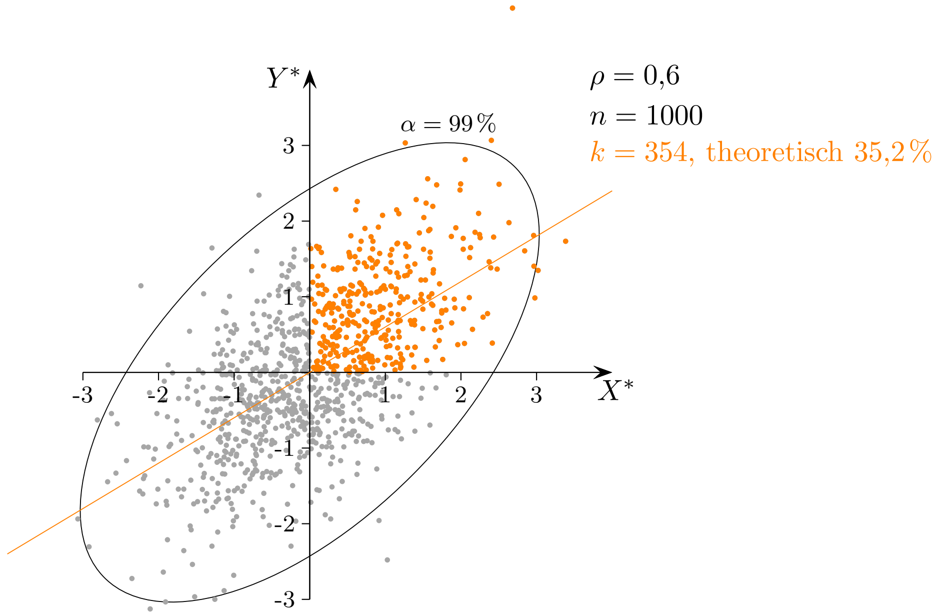
$$x = x^*$$
$$y = \frac{y^* - \rho^*}{\sqrt{1-\rho^2}}$$

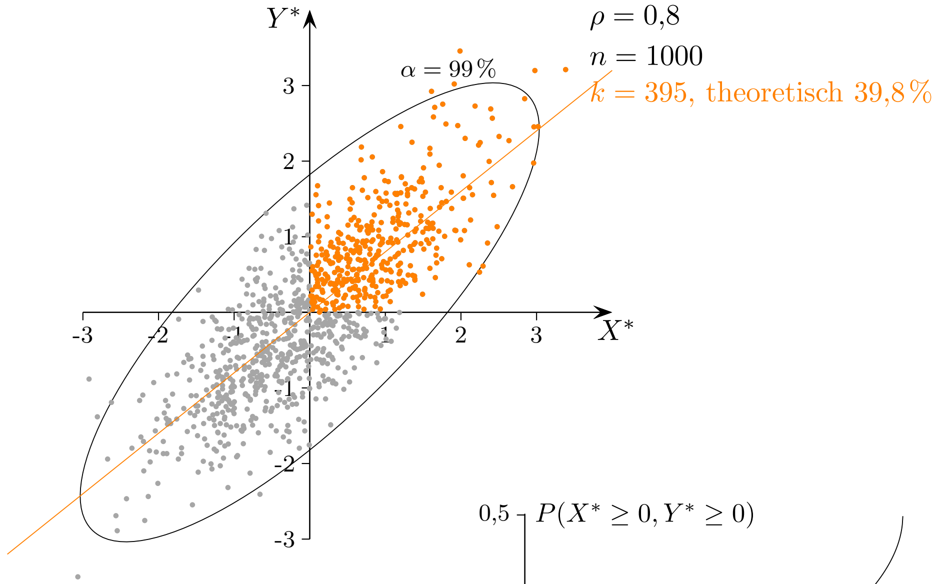
wird in $f(x,y) = \dfrac{1}{2\pi}\, e^{-\frac{1}{2}(x^2+y^2)}$ eingesetzt
und der Stauchfaktor $\sqrt{1-\rho^2}$ mit $\dfrac{1}{\sqrt{1-\rho^2}}$ ausgeglichen
(das Volumen muss 1 bleiben).
Die führt direkt (leichte Umformungen, Umbenennungen)
zur bivariaten Dichte

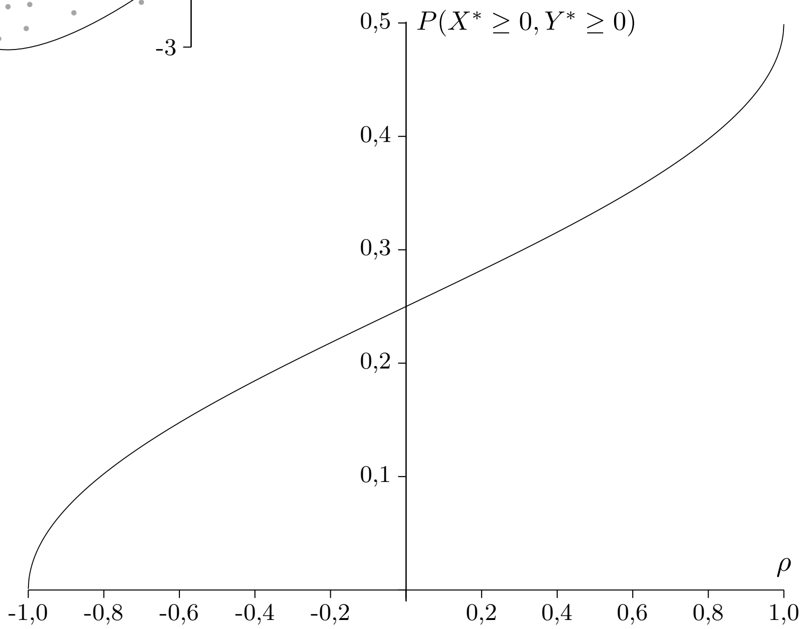
$$g(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho x y + y^2)}.$$



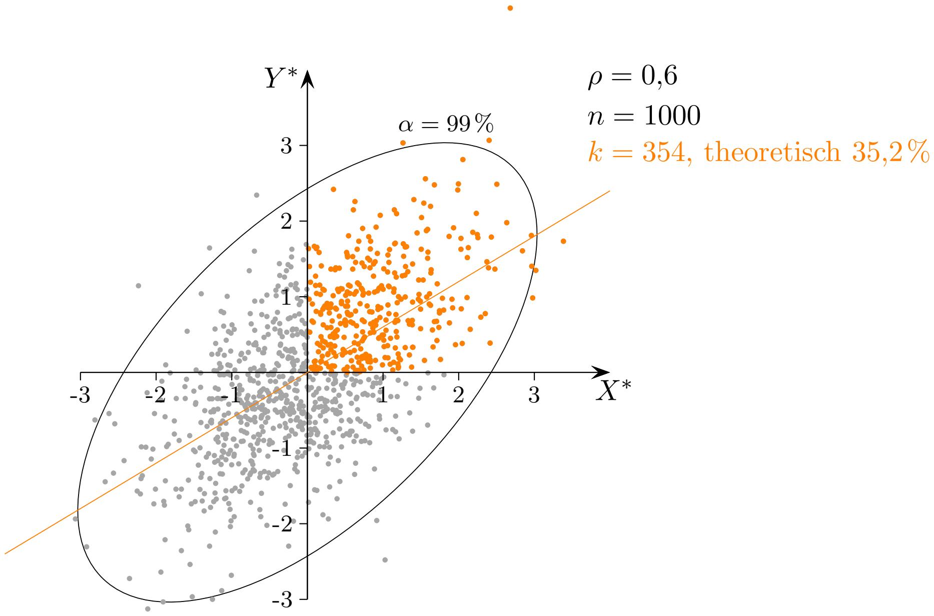
Zusammenhang von ρ und $P(X^* \geq 0, Y^* \geq 0)$







Zusammenhang von ρ und $P(X^* \geq 0, Y^* \geq 0)$



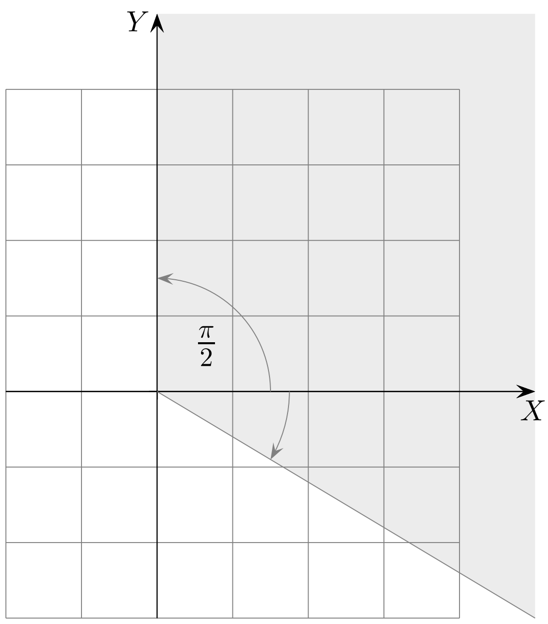
Das Ereignis $X^* \geq 0, Y^* \geq 0$ bedeutet, dass die unstandardisierten Zufallsvariablen nicht unter dem jeweiligen Erwartungswert liegen.

Aufgrund der Abbildungsgleichungen

$$X^* = X$$
$$Y^* = \rho X + \sqrt{1-\rho^2}\,Y$$

ist das gleichbedeutend mit

$$X \geq 0$$
$$Y \geq \frac{-\rho}{\sqrt{1-\rho^2}}\,X.$$



Die Gerade, die die Fläche nach unten begrenzt, besitzt die Steigung $m = \frac{-\rho}{\sqrt{1-\rho^2}}$.

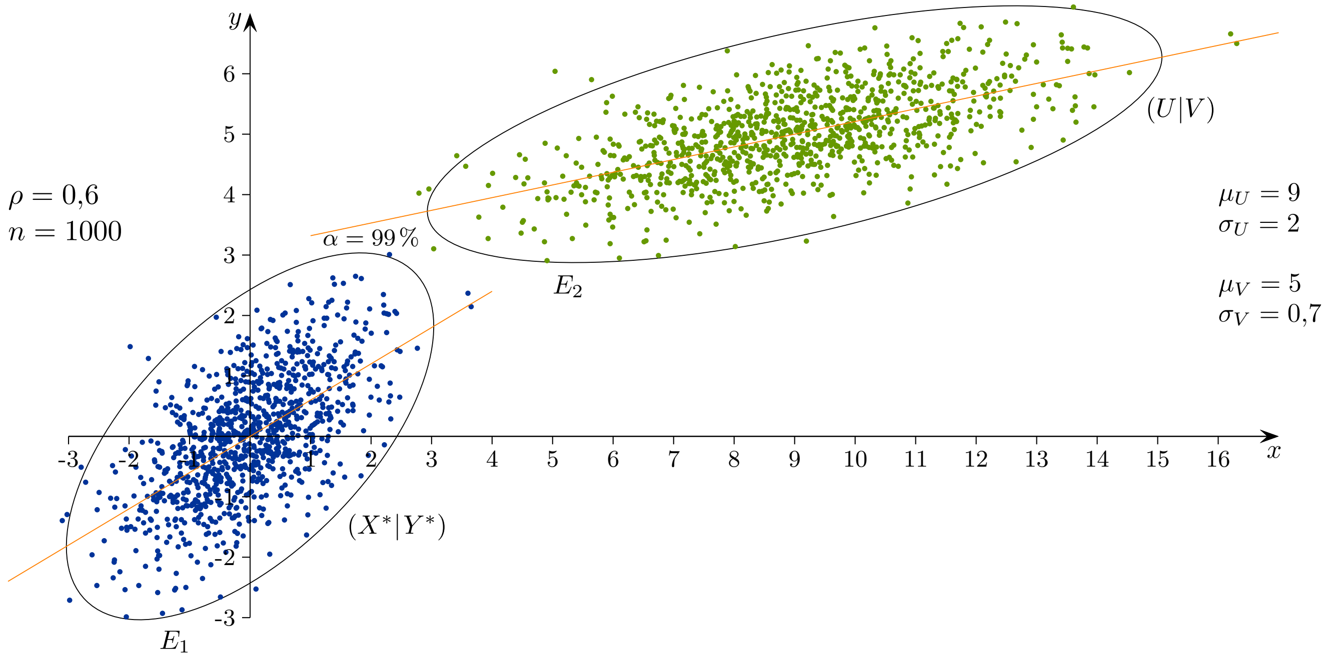
Hiermit kann der gesamte Winkel und der Anteil am Ganzen (2π) ermittelt werden.

Wir erhalten:

$$P(X^* \geq 0, Y^* \geq 0) = \frac{-\arctan\frac{-\rho}{\sqrt{1-\rho^2}} + \frac{\pi}{2}}{2\pi} = \frac{\arctan\frac{\rho}{\sqrt{1-\rho^2}} + \frac{\pi}{2}}{2\pi}$$

Umgekehrt kann durch Auszählen ρ geschätzt werden.

Unstandardisierte Daten



Die Gleichung der Ellipse E_2 lautet

$$\frac{\left(\frac{x-\mu_U}{\sigma_U}+\frac{y-\mu_V}{\sigma_V}\right)^2}{2(1+\rho)}+\frac{\left(-\frac{x-\mu_U}{\sigma_U}+\frac{y-\mu_V}{\sigma_V}\right)^2}{2(1-\rho)}=r^2, \qquad r=F^{-1}(\alpha)=\sqrt{-2\ln(1-\alpha)}$$

Durch Standardisierung geht E_2 in E_1 über

$$\frac{x+y}{2(1+\rho)}+\frac{-x+y}{2(1-\rho)}=r^2,$$

vereinfacht ergibt das

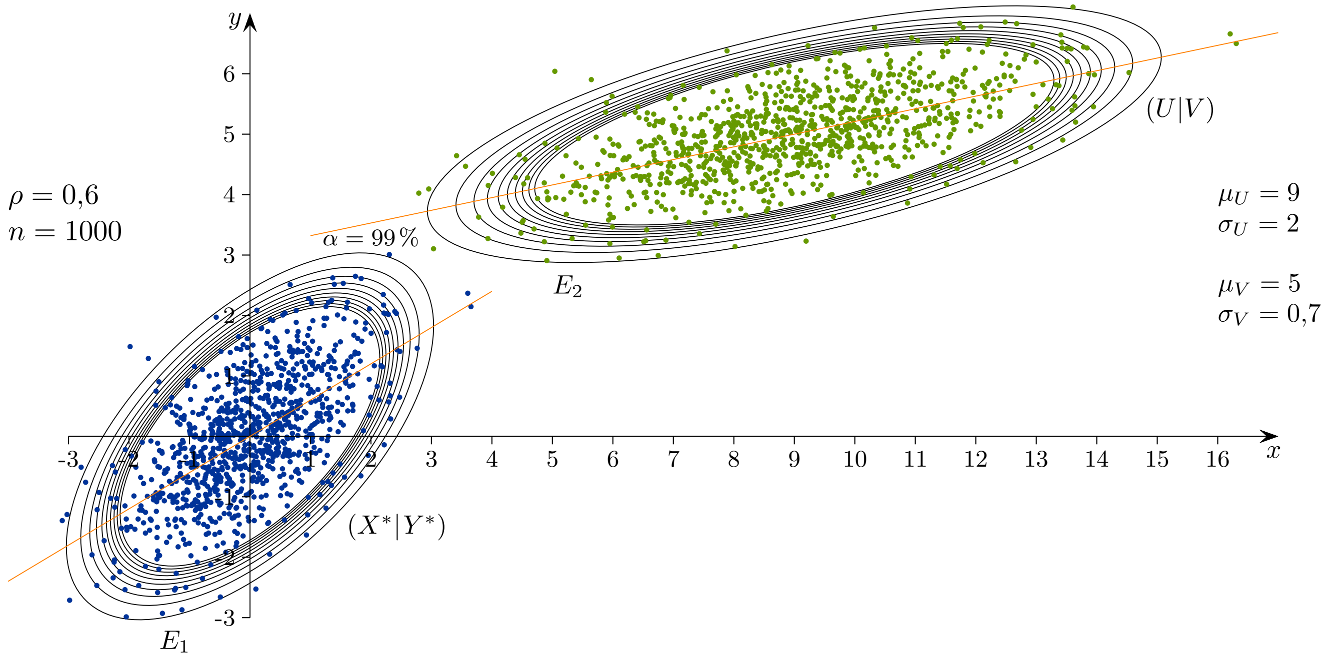
$$x^2-2\rho xy+y^2=r^2\cdot(1-\rho^2).$$

Diese Ellipse ist, wie wir gesehen haben, das Bild des Kreises $x^2+y^2=r^2$
unter der Abbildung

$$\begin{aligned} x^* &= x \\ y^* &= \rho x+\sqrt{1-\rho^2}\,y. \end{aligned}$$

Die Gleichung für die in den Ursprung verschobene Ellipse E_2 stimmt mit $(x,y)\,\Sigma^{-1}\begin{pmatrix}x\\y\end{pmatrix}=r^2$ überein, siehe Seite 2 unten. Hauptachsen von E_2: $a=\sqrt{r^2/e_1}$, $b=\sqrt{r^2/e_2}$, Eigenwerte e_i von Σ^{-1}

Unstandardisierte Daten



Je Ellipse verringert sich der Anteil von 99 % um 1 %.

Auf die Matrizenschreibweise der Dichte mit Hilfe der Kovarianzmatrix Σ soll noch eingegangen werden. Für standardisierte Daten gilt:

$$g(x,y) = \frac{1}{2\pi\sqrt{|\,\Sigma\,|}}\, e^{-\frac{1}{2}(x|y)\;\Sigma^{-1}\binom{x}{y}}$$

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \qquad |\,\Sigma\,| = 1 - \rho^2$$

Mit der Abbildung
$$\begin{aligned} x' &= \mu_X + \sigma_X x \\ y' &= \mu_Y + \sigma_X y \end{aligned} \qquad \begin{vmatrix} x'_x & x'_y \\ y'_x & y'_y \end{vmatrix} = \sigma_X \sigma_Y$$

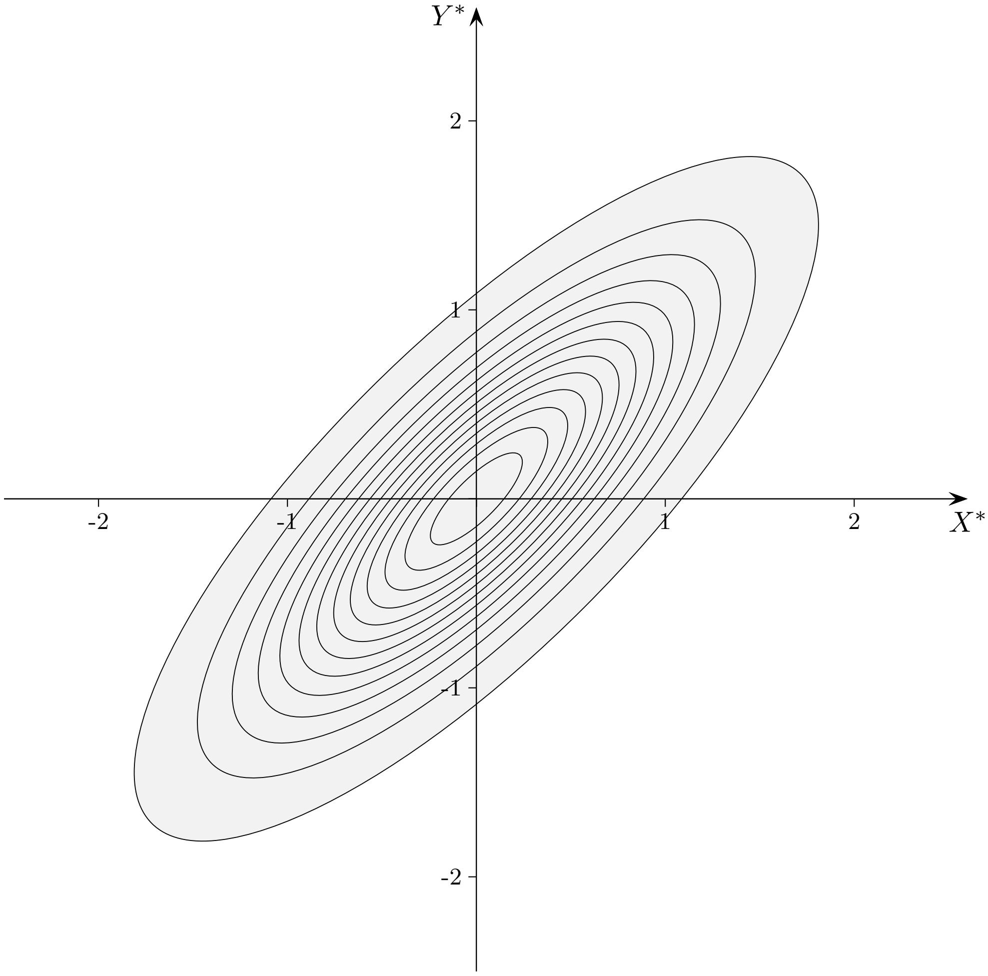
geht die Dichte über in

$$g(x,y) = \frac{1}{2\pi\sqrt{|\,\Sigma\,|}}\, e^{-\frac{1}{2}(x-\mu_X\,|\,y-\mu_Y)\;\Sigma^{-1}\binom{x-\mu_X}{y-\mu_Y}}$$

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}, \qquad |\,\Sigma\,| = \sigma_X^2\sigma_Y^2(1-\rho^2), \qquad \sqrt{|\,\Sigma\,|} = \sigma_X\sigma_Y\sqrt{1-\rho^2}$$

Höhenlinien

$\rho = 0{,}8$



Je Ellipse verringert sich die Höhe von 0,25 um 0,02. Die maximale Höhe beträgt 0,265.