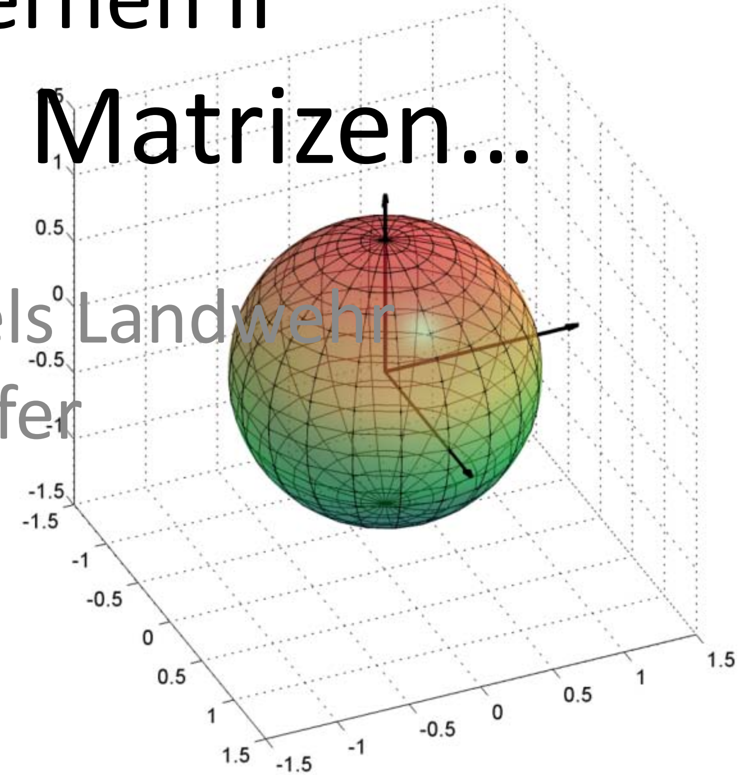
Maschinelles Lernen II
Grundlegendes zu Matrizen…

Christoph Sawade/Niels Landwehr
Tobias Scheffer



Matrix-Beispiele

Wiederholung: Daten

- Lineares Model: $f(\mathbf{x}, y) = \mathbf{w}^{\mathrm{T}}\Phi\left(\mathbf{x}, y\right)$

- Sei $\mathbf{X} = \begin{pmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{pmatrix}$ die Datenmatrix, $\mathbf{W} = \begin{pmatrix} | & & | \\ \mathbf{w}_1 & \cdots & \mathbf{w}_k \\ | & & | \end{pmatrix}$

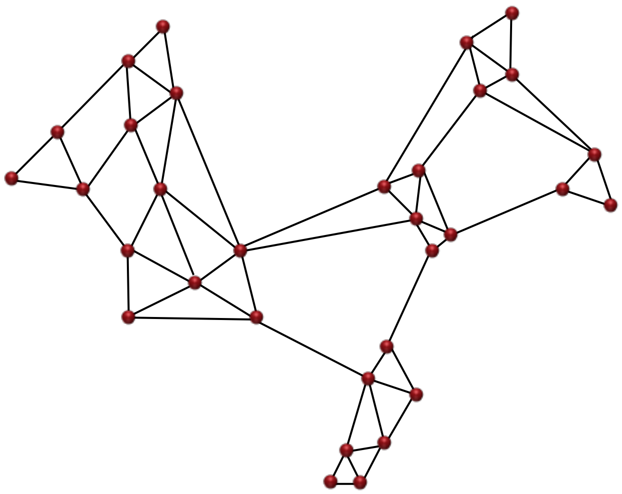
  die Matrix der Klassenmodelle, dann gilt

$$\mathbf{W}^T\mathbf{X} = \begin{pmatrix} f\left(\mathbf{x}_1, 1\right) & \cdots & f\left(\mathbf{x}_n, 1\right) \\ \vdots & \ddots & \vdots \\ f\left(\mathbf{x}_1, k\right) & \cdots & f\left(\mathbf{x}_n, k\right) \end{pmatrix}$$

Matrix-Beispiele

Wiederholung: Graphen

- Adjazenzmatrix



Kante zwischen
Knoten i und j?

$$G = \begin{pmatrix} & \vdots & \\ \cdots & g_{i,j} & \cdots \\ & \vdots & \end{pmatrix}$$

- Inzidenzmatrix, Laplace-Matrix,…

Matrix-Beispiele

Wiederholung: Gleichungssysteme

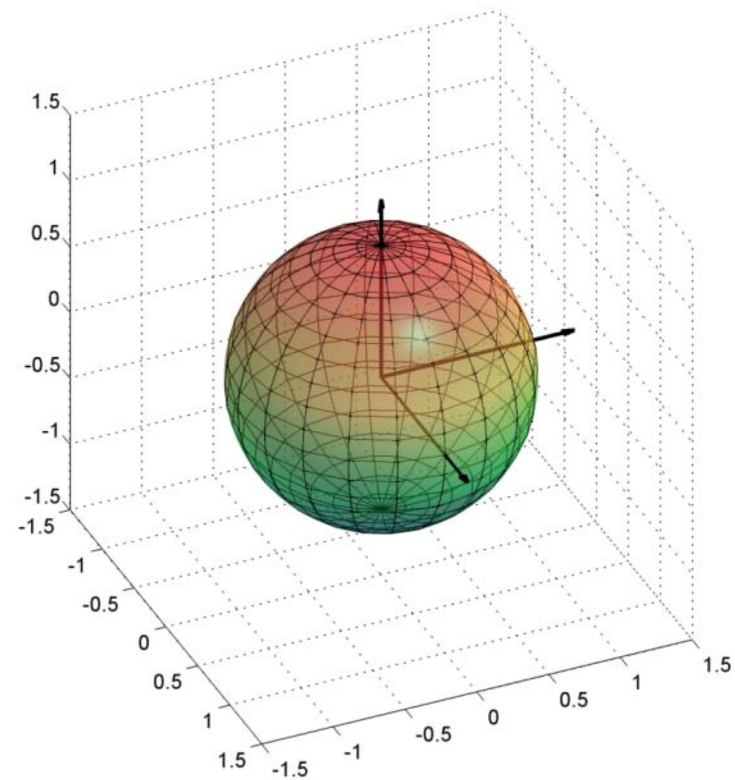
- Elementarmatrizen $\mathbf{R} \in \mathbb{R}^{n \times n}$: Grundoperationen des Gauß'sches Eliminierungsverfahren
  – Addition von Zeilen/Spalten
  – Multiplikation einer Zeile/Spalte mit Skalar
  – Vertauschen von Zeilen/Spalten

- $\mathbf{R}$ entsteht durch die Anwendung der gewünschten Operation auf die Einheitsmatrix

Matrix-Beispiele

Wiederholung: Lineare Abbildung

- Matrixmultiplikationen sind lineare Abbildungen: $\mathbf{A}(c\mathbf{x}+\mathbf{y}) = c\mathbf{A}\mathbf{x}+\mathbf{A}\mathbf{y}$

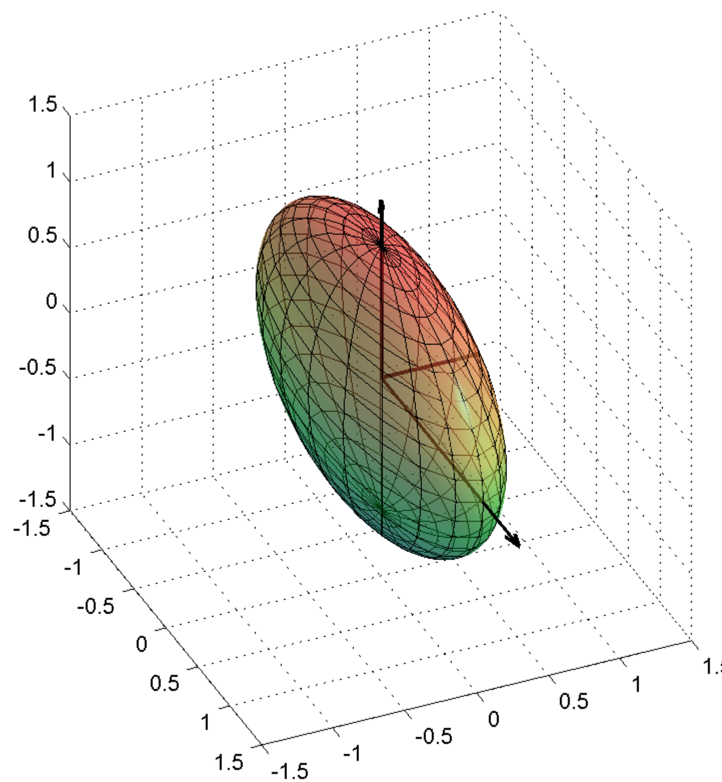


Matrix-Beispiele

Wiederholung: Lineare Abbildung

- Skalierung

$$\mathbf{A}_{skal}\left(\mathbf{c}\right)=\begin{pmatrix} c_1 & 0 & 0 \\ 0 & c_2 & \\ 0 & 0 & \end{pmatrix}$$



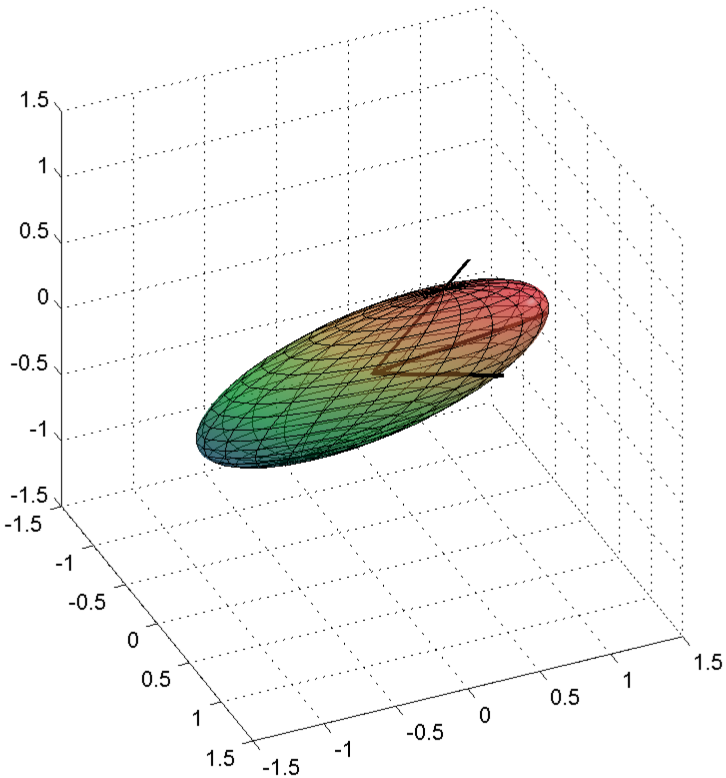
Matrix-Beispiele

Wiederholung: Lineare Abbildung

- Scherung

$$\mathbf{A}_{sch}(c) = \begin{pmatrix} c & 1-c & 1-c \\ 1-c & c & \\ 1-c & 1-c & \end{pmatrix}$$

- $\begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$ wird nur skalier



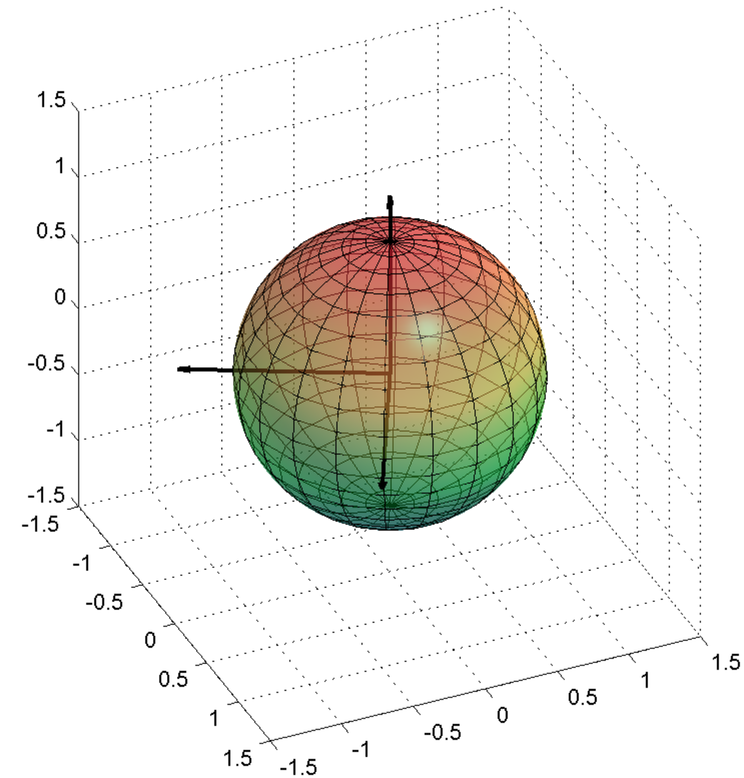
Matrix-Beispiele

Wiederholung: Lineare Abbildung

- Rotation

$$\mathbf{A}_{rot}^{z}(\theta)=\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos & \\ 0 & 0 & \end{pmatrix}$$

- z-Achse ist invariant

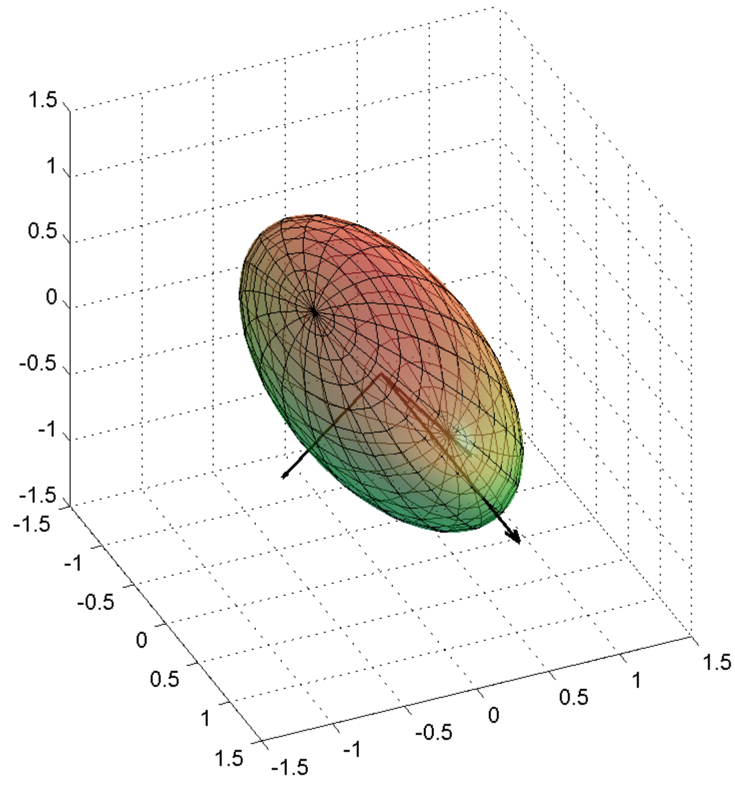


Matrix-Beispiele

Wiederholung: Lineare Abbildung

- Kombinationen

$$\mathbf{A}_{rot}^{x}(\theta)\cdot\mathbf{A}_{skal}(\mathbf{c})$$



- x-Achse wird nur ska

Eigenwertproblem

- Eigenvektor $\mathbf{u_i} \in \mathbb{C}^m$ einer Abbildung $\mathbf{A} \in \mathbb{R}^{m \times m}$ wird durch $\mathbf{A}$ nur skaliert.

- Eigenwert $\lambda_i \in \mathbb{C}$ einer Abbildung $\mathbf{A}$ ist der zu $\mathbf{u_i}$ gehörende Skalierungsfaktor.

- Eigenwertproblem: Bestimme $\mathbf{u_i}$, s.d.
$$\mathbf{A}\mathbf{u_i} = \lambda_i \mathbf{u_i}$$

Eigenwertproblem

$$\boxed{\mathbf{A}\mathbf{u_i} = \lambda_i \mathbf{u_i}}$$

- $\mathrm{eig}(\mathbf{A}^k) = \{\lambda_1^k, \ldots, \lambda_m^k\}, k = ?$
- $\det(\mathbf{A}) = \prod_{i=1}^m \lambda_i, \mathrm{trace}(\mathbf{A}) = \sum_{i=1}^m \lambda_i$
- $\det(\mathbf{A} - \lambda_i \mathbf{I}) = 0$
- Wenn $\mathbf{A}$ symmetrisch, dann $\forall i : \lambda_i \in \mathbb{R}$
- Wenn $\mathbf{A}$ positiv (semi-) definit, dann $\forall i : \lambda_i \geq 0$
- Eigenwertzerlegung:

$$\mathbf{A} = \begin{pmatrix} | & & | \\ \mathbf{u}_1 & \cdots & \mathbf{u}_m \\ | & & | \end{pmatrix} \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & 0 \\ 0 & \cdots & \lambda_m \end{pmatrix} \begin{pmatrix} - & \mathbf{u}_1 & - \\ & \vdots & \\ - & \mathbf{u}_m & - \end{pmatrix}$$

- Analogon für $\mathbf{B} \in \mathbb{R}^{m_1 \times m_2}$: Singulärwerte

Eigenwertproblem

symmetrische Matrizen

$$\boxed{\mathbf{A}\mathbf{u_i} = \lambda_i \mathbf{u_i}}$$

- Es gilt:
  - Wenn $\mathbf{u_i}$ und $\mathbf{u_j}$ Lösungen sind, dann sind auch $c\mathbf{u_i} + \mathbf{u_j}$ Lösungen f.a. $c \in \mathbb{R}$ (Linearität)
  - $\mathbf{u_i} = \mathbf{0}$ ist triviale Lösung

- Für symmetrische Matrizen, können Eigenvektoren orthonormal gewählt werden:
  - $\mathbf{u_i}^T \mathbf{u_j} = 0$
  - $\mathbf{u_i}^T \mathbf{u_i} = 1$

Bestimmung der Eigenwerte

Power Iteration

- Iterative Annäherung an den größten Eigenvektor λ_1:
  - Initialisiere b_0 zufällig
  - Wiederhole bis ... $b_{k+1} = \frac{Ab_k}{\| Ab_k \|} \to \lambda_1$

- Konvergenzgeschwindigkeit:

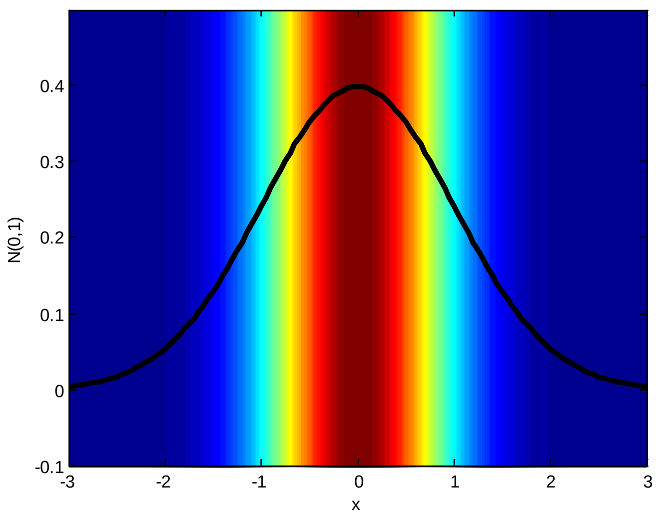
$$\left|\lambda_1 - b_k\right| = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$$

Positive Definitheit

- $\mathbf{A} \in \mathbb{R}^{m \times m}$ ist positiv semi-definit:

$$\forall \mathbf{z} \in \mathbb{R}^m : \mathbf{z}^T \mathbf{A} \mathbf{z} \geq 0$$

- $\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$



$\mu = 0$

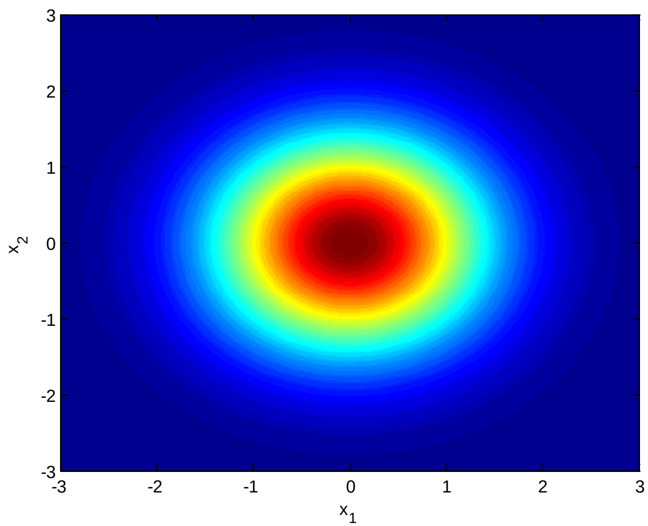
$\sigma^2 = 1$

Positive Definitheit

- $\mathbf{A} \in \mathbb{R}^{m \times m}$ ist positiv semi-definit:

$$\forall \mathbf{z} \in \mathbb{R}^m : \mathbf{z}^T \mathbf{A} \mathbf{z} \geq 0$$

- $\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{x}-\boldsymbol{\mu})^T(\mathbf{x}-\boldsymbol{\mu})\right)$



$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\sigma^2 = 1$$

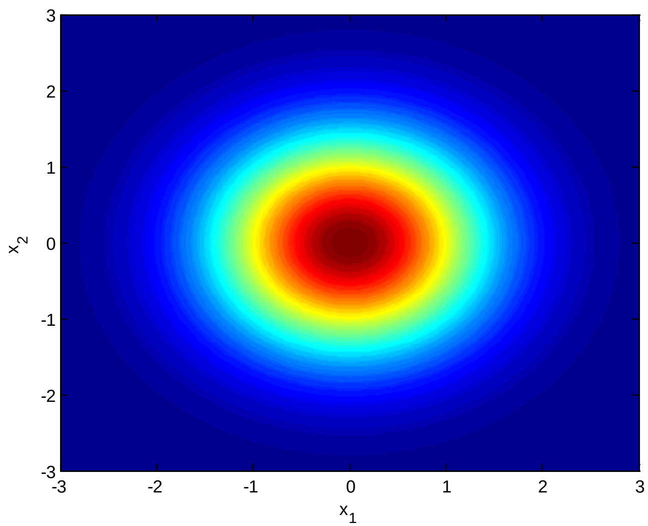
Positive Definitheit

- $\mathbf{A} \in \mathbb{R}^{m \times m}$ ist positiv semi-definit:

$$\forall \mathbf{z} \in \mathbb{R}^m : \mathbf{z}^T \mathbf{A} \mathbf{z} \geq 0$$

- $$\frac{1}{(2\pi)^{m/2}|\mathbf{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2} \overbrace{(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})}^{\text{quadrierte Mahalanobis-Distanz}}\right)$$

Kovarianz-Matrix



$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

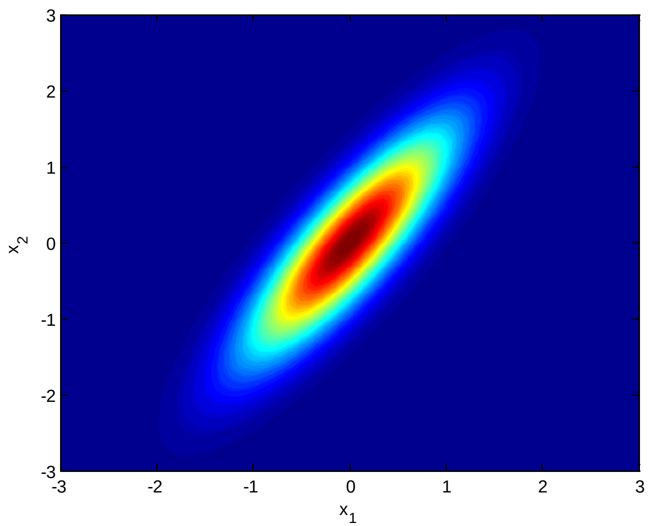
Positive Definitheit

- $\mathbf{A} \in \mathbb{R}^{m \times m}$ ist positiv semi-definit:

$$\forall \mathbf{z} \in \mathbb{R}^m : \mathbf{z}^T \mathbf{A} \mathbf{z} \geq 0$$

- $\frac{1}{(2\pi)^{m/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}\overbrace{(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}^{\text{quadrierte Mahalanobis-Distanz}}\right)$

Kovarianz-Matrix



$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 0.5 & 0.6 \\ 0.6 & 1 \end{pmatrix}$$