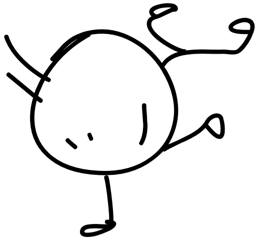


Merkmale und ihre Dimensionalität

Vorlesung 186.844

19.11.2015

TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

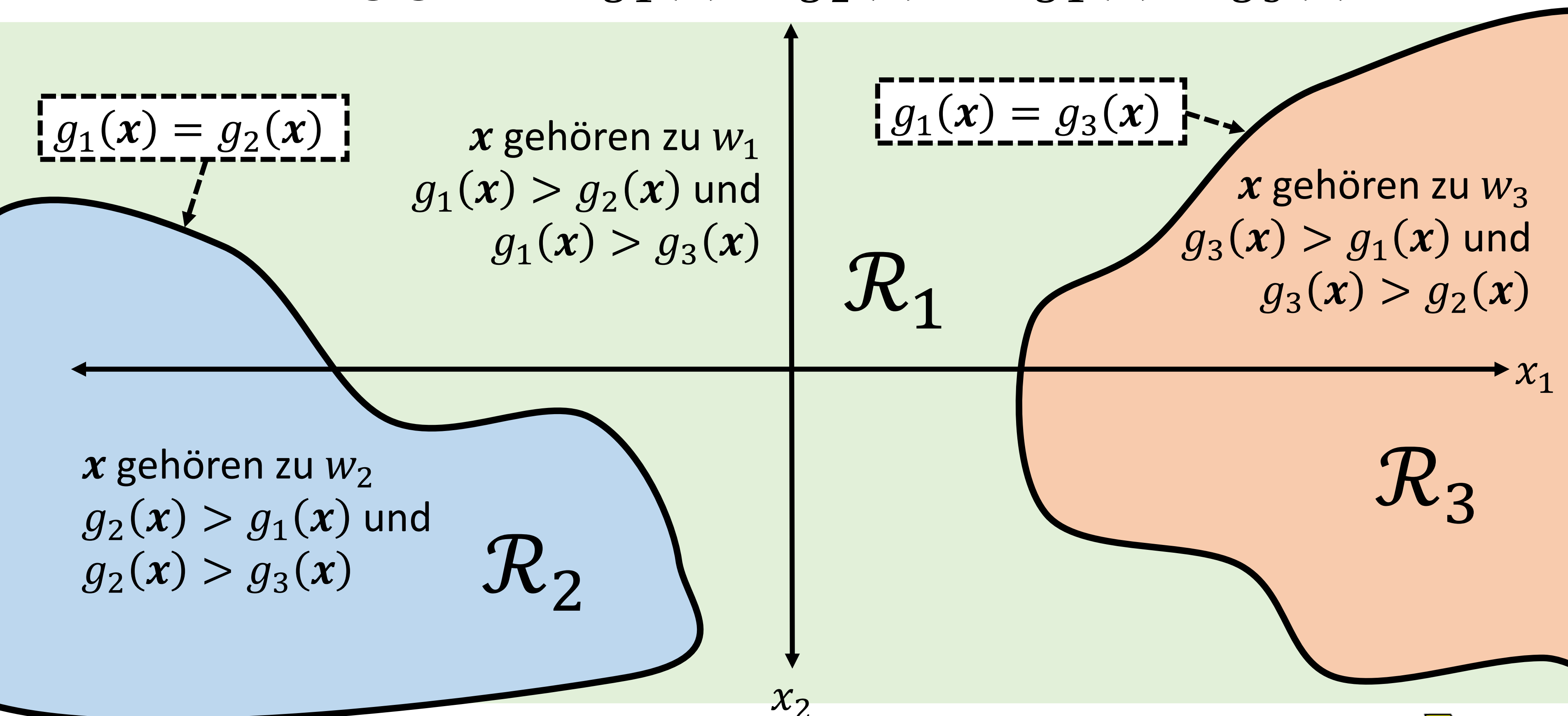
Überblick

I. WH: Diskriminantenfunktionen + Beispiele
II. Merkmale
III. Fluch der Dimensionalität
IV. Merkmalsselektion
V. Merkmalsreduktion
VI. Anwendungsbeispiele

I. Wiederholung: Diskriminantenfunktionen + Beispiele

Wofür?

- Entscheidungsregionen: $\mathcal{R}_1$, $\mathcal{R}_2$ und $\mathcal{R}_3$
- Entscheidungsgrenzen: $g_1(\boldsymbol{x}) = g_2(\boldsymbol{x})$ und $g_1(\boldsymbol{x}) = g_3(\boldsymbol{x})$



Diskriminatenfunktionen

… für multivariate Normalverteilungen unter Verwendung eines **Bayes-Klassifikators**

- durch Logarithmieren der Posteriors

$$g_j(\boldsymbol{x}) = \ln \frac{p(\boldsymbol{x}|\omega_j)P(\omega_j)}{p(\boldsymbol{x})}$$

- erhält man die folgenden (optimalen) Diskriminanten-funktionen

$$g_j(\boldsymbol{x}) = -\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu}_j)^T \Sigma_j^{-1}(\boldsymbol{x} - \boldsymbol{\mu}_j) - \frac{1}{2}\ln|\Sigma_j| + \ln P(\omega_j) - \frac{p}{2}\ln 2\pi - \ln p(\boldsymbol{x})$$

Vereinfachung

$$g_j(\boldsymbol{x}) = -\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_j)^T \Sigma_j^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_j) - \frac{1}{2}\ln|\Sigma_j| + \ln P(\omega_j) \cancel{- \frac{p}{2}\ln 2\pi - \ln p(\boldsymbol{x})}$$

- die letzten beiden Terme können weggelassen werden, weil sie unabhängig von den Klassen ω_j sind

- Diskriminantenfunktionen vereinfachen sich zu

$$g_j(\boldsymbol{x}) = -\frac{1}{2}\boxed{d_j^2(\boldsymbol{x})} + \left[-\frac{1}{2}\ln\boxed{|\boldsymbol{\Sigma}_j|} + \ln\boxed{P(\omega_j)}\right]$$

Mahalanobisdistanz

Determinante von $\boldsymbol{\Sigma}$

Prior

Beispiel: Angabe

Aufgabe: Ermittle die Diskriminantenfunktionen von Klasse A und B und ihre Entscheidungsgrenze…

1. für den allgemeinen Fall
2. für Spezialfall 1

Annahmen:

$p(\boldsymbol{x}|\omega_j)$ → multivariate Normalverteilungen

$P(w_A) = P(w_B) = 0,5$ → gleiche Priors

für Spezialfall 1: $\Sigma_A = \Sigma_B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Beispiel: Angabe

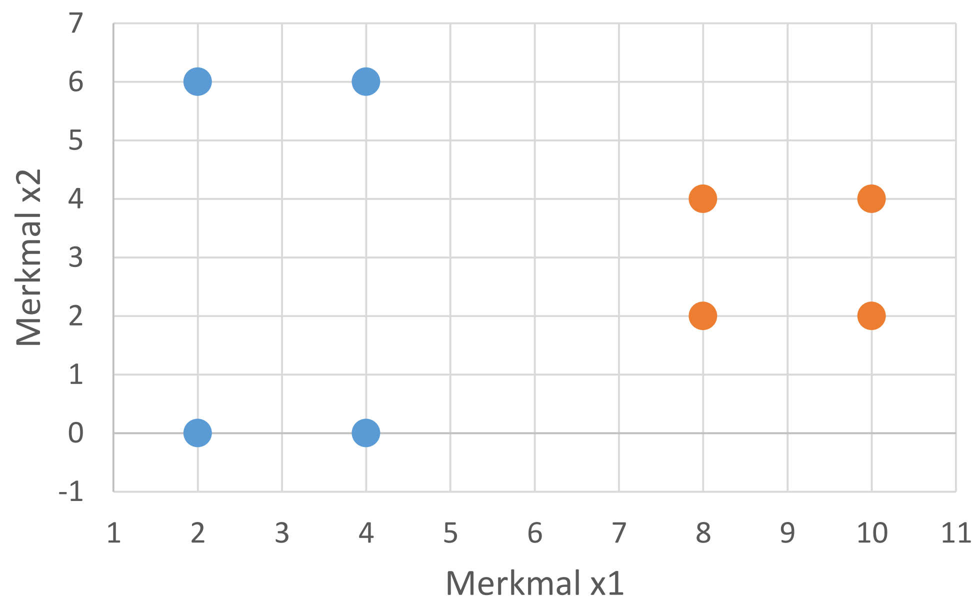
Trainingsdaten:

Klasse A

$$\begin{bmatrix}2\\6\end{bmatrix}, \begin{bmatrix}2\\0\end{bmatrix}, \begin{bmatrix}4\\6\end{bmatrix}, \begin{bmatrix}4\\0\end{bmatrix}$$

Klasse B

$$\begin{bmatrix}8\\2\end{bmatrix}, \begin{bmatrix}8\\4\end{bmatrix}, \begin{bmatrix}10\\2\end{bmatrix}, \begin{bmatrix}10\\4\end{bmatrix}$$



Beispiel: Lösung

1. Schätzung der Mittelwertvektoren

$$\hat{\boldsymbol{\mu}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}_i$$

$$\hat{\boldsymbol{\mu}}_A = \left(\begin{bmatrix}2\\6\end{bmatrix} + \begin{bmatrix}2\\0\end{bmatrix} + \begin{bmatrix}4\\6\end{bmatrix} + \begin{bmatrix}4\\0\end{bmatrix}\right) \cdot \frac{1}{4} = \begin{bmatrix}3\\3\end{bmatrix}$$

$$\hat{\boldsymbol{\mu}}_B = \left(\begin{bmatrix}8\\2\end{bmatrix} + \begin{bmatrix}8\\4\end{bmatrix} + \begin{bmatrix}10\\2\end{bmatrix} + \begin{bmatrix}10\\4\end{bmatrix}\right) \cdot \frac{1}{4} = \begin{bmatrix}9\\3\end{bmatrix}$$

Beispiel: Lösung

… für allgemeinen Fall

2. Schätzen der Kovarianzmatrizen

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{N-1}\sum_{i=1}^{N}(\boldsymbol{x}_i - \hat{\boldsymbol{\mu}})(\boldsymbol{x}_i - \hat{\boldsymbol{\mu}})^T$$

Beispiel: Lösung

… für allgemeinen Fall

2. Schätzen der Kovarianzmatrizen

$$\hat{\Sigma}_A = \frac{1}{4-1}\left\{\left(\begin{bmatrix}2\\6\end{bmatrix}-\begin{bmatrix}3\\3\end{bmatrix}\right)\left(\begin{bmatrix}2\\6\end{bmatrix}-\begin{bmatrix}3\\3\end{bmatrix}\right)^T+\left(\begin{bmatrix}2\\0\end{bmatrix}-\begin{bmatrix}3\\3\end{bmatrix}\right)\right.$$

$$\left(\begin{bmatrix}2\\0\end{bmatrix}-\begin{bmatrix}3\\3\end{bmatrix}\right)^T+\left(\begin{bmatrix}4\\6\end{bmatrix}-\begin{bmatrix}3\\3\end{bmatrix}\right)\left(\begin{bmatrix}4\\6\end{bmatrix}-\begin{bmatrix}3\\3\end{bmatrix}\right)^T+$$

$$\left.\left(\begin{bmatrix}4\\0\end{bmatrix}-\begin{bmatrix}3\\3\end{bmatrix}\right)\left(\begin{bmatrix}4\\0\end{bmatrix}-\begin{bmatrix}3\\3\end{bmatrix}\right)^T\right\}=$$

$$= \frac{1}{3}\left\{\begin{bmatrix}-1\\3\end{bmatrix}\begin{bmatrix}-1 & 3\end{bmatrix}+\begin{bmatrix}-1\\-3\end{bmatrix}\begin{bmatrix}-1 & -3\end{bmatrix}+\begin{bmatrix}1\\3\end{bmatrix}\begin{bmatrix}1 & 3\end{bmatrix}+\right.$$

$$\left.\begin{bmatrix}1\\-3\end{bmatrix}\begin{bmatrix}1 & -3\end{bmatrix}\right\}=\frac{1}{3}\left\{\begin{bmatrix}1 & -3\\-3 & 9\end{bmatrix}+\begin{bmatrix}1 & 3\\3 & 9\end{bmatrix}+\begin{bmatrix}1 & 3\\3 & 9\end{bmatrix}+\begin{bmatrix}1 & -3\\-3 & 9\end{bmatrix}\right\}=$$

$$=\frac{1}{3}\begin{bmatrix}4 & 0\\0 & 36\end{bmatrix}=\begin{bmatrix}4/3 & 0\\0 & 12\end{bmatrix}$$

Beispiel: Lösung

… für allgemeinen Fall

2. Schätzen der Kovarianzmatrizen

$$\hat{\Sigma}_B = \frac{1}{4-1}\left\{\left(\begin{bmatrix}8\\2\end{bmatrix}-\begin{bmatrix}9\\3\end{bmatrix}\right)\left(\begin{bmatrix}8\\2\end{bmatrix}-\begin{bmatrix}9\\3\end{bmatrix}\right)^T+\left(\begin{bmatrix}8\\4\end{bmatrix}-\begin{bmatrix}9\\3\end{bmatrix}\right)\left(\begin{bmatrix}8\\4\end{bmatrix}-\begin{bmatrix}9\\3\end{bmatrix}\right)^T+\right.$$

$$\left.\left(\begin{bmatrix}10\\2\end{bmatrix}-\begin{bmatrix}9\\3\end{bmatrix}\right)\left(\begin{bmatrix}10\\2\end{bmatrix}-\begin{bmatrix}9\\3\end{bmatrix}\right)^T+\left(\begin{bmatrix}10\\4\end{bmatrix}-\begin{bmatrix}9\\3\end{bmatrix}\right)\left(\begin{bmatrix}10\\4\end{bmatrix}-\begin{bmatrix}9\\3\end{bmatrix}\right)^T\right\} =$$

$$= \frac{1}{3}\left\{\begin{bmatrix}-1\\-1\end{bmatrix}\begin{bmatrix}-1 & -1\end{bmatrix}+\begin{bmatrix}-1\\1\end{bmatrix}\begin{bmatrix}-1 & 1\end{bmatrix}+\begin{bmatrix}1\\-1\end{bmatrix}\begin{bmatrix}1 & -1\end{bmatrix}+\begin{bmatrix}1\\1\end{bmatrix}\begin{bmatrix}1 & 1\end{bmatrix}\right\} =$$

$$= \frac{1}{3}\left\{\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}+\begin{bmatrix}1 & -1\\-1 & 1\end{bmatrix}+\begin{bmatrix}1 & -1\\-1 & 1\end{bmatrix}+\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}\right\} = \frac{1}{3}\begin{bmatrix}4 & 0\\0 & 4\end{bmatrix} =$$

$$= \begin{bmatrix}4/3 & 0\\0 & 4/3\end{bmatrix}$$

Beispiel: Lösung

… für allgemeinen Fall

3. Determinanten ermitteln

$$|\boldsymbol{\Sigma}| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

$$|\boldsymbol{\Sigma}|_A = \begin{vmatrix} {}^4\!/_3 & 0 \\ 0 & 12 \end{vmatrix} = {}^4\!/_3 \cdot 12 = 16$$

$$|\boldsymbol{\Sigma}|_B = \begin{vmatrix} {}^4\!/_3 & 0 \\ 0 & {}^4\!/_3 \end{vmatrix} = {}^4\!/_3 \cdot {}^4\!/_3 = {}^{16}\!/_9$$

Beispiel: Lösung

… für allgemeinen Fall

4. Inverse berechnen

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$\boldsymbol{\Sigma}^{-1}{}_{A} = \frac{1}{16}\begin{bmatrix} 12 & 0 \\ 0 & {}^{4}/_{3} \end{bmatrix} = \begin{bmatrix} {}^{3}/_{4} & 0 \\ 0 & {}^{1}/_{12} \end{bmatrix}$$

$$\boldsymbol{\Sigma}^{-1}{}_{B} = \frac{9}{16}\begin{bmatrix} {}^{4}/_{3} & 0 \\ 0 & {}^{4}/_{3} \end{bmatrix} = \begin{bmatrix} {}^{3}/_{4} & 0 \\ 0 & {}^{3}/_{4} \end{bmatrix}$$

Beispiel: Lösung

… für allgemeinen Fall

5. Diskriminantenfunktionen bestimmen

$$g_j(\boldsymbol{x}) = -\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{\mu}_j\right)^T \Sigma_j^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}_j\right) - \frac{1}{2}\ln\left|\Sigma_j\right| + \cancel{\ln P(\omega_j)}$$

$$g_A(x) = -\frac{1}{2}\left\{\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 3 \\ 3 \end{pmatrix}\right)^T \begin{pmatrix} 3/4 & 0 \\ 0 & 1/12 \end{pmatrix}\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 3 \\ 3 \end{pmatrix}\right)\right\} - \frac{1}{2}\ln(16)$$

$$\begin{pmatrix} (x_1-3) & (x_2-3) \end{pmatrix} \begin{pmatrix} 3/4 & 0 \\ 0 & 1/12 \end{pmatrix} \begin{pmatrix} (x_1-3) \\ (x_2-3) \end{pmatrix} =$$

$$\begin{pmatrix} \left(\frac{3x_1}{4} - \frac{9}{4}\right) & \left(\frac{x_2}{12} - \frac{1}{4}\right) \end{pmatrix} \begin{pmatrix} (x_1-3) \\ (x_2-3) \end{pmatrix} =$$

$$\left(\frac{3x_1}{4} - \frac{9}{4}\right)(x_1-3) + \left(\frac{x_2}{12} - \frac{1}{4}\right)(x_2-3) =$$

$$3/4\, x_1^2 - 9/4\, x_1 - 9/4\, x_1 + 27/4 + 1/12\, x_2^2 - 1/4\, x_2 - 3/12\, x_2 + 3/4 =$$

$$3/4\, x_1^2 - 18/4\, x_1 + 1/12\, x_2^2 - 2/4\, x_2 + 30/4 = d_A^2$$

Beispiel: Lösung

… für allgemeinen Fall

5. Diskriminantenfunktionen bestimmen

$$g_A(\mathbf{x}) = -\tfrac{1}{2}\, d_A^2 - \tfrac{1}{2}\ln(16) = -\tfrac{3}{8}x_1^2 + \tfrac{18}{8}x_1 - \tfrac{1}{24}x_2^2 + \tfrac{2}{8}x_2 - \tfrac{30}{8} - \tfrac{1}{2}\ln(16) = \boxed{-\tfrac{3}{8}x_1^2 + \tfrac{9}{4}x_1 - \tfrac{1}{24}x_2^2 + \tfrac{1}{4}x_2 - \tfrac{15}{4} - \tfrac{1}{2}\ln(16)}$$

$$g_B(\mathbf{x}) = -\tfrac{1}{2}\underbrace{\left\{\left(\begin{pmatrix}x_1\\x_2\end{pmatrix}-\begin{pmatrix}9\\3\end{pmatrix}\right)^T \begin{pmatrix}3/4 & 0\\ 0 & 3/4\end{pmatrix}\left(\begin{pmatrix}x_1\\x_2\end{pmatrix}-\begin{pmatrix}9\\3\end{pmatrix}\right)\right\}} - \tfrac{1}{2}\ln(16/9)$$

$$\big((x_1-9)\,(x_2-3)\big)\begin{pmatrix}3/4 & 0\\ 0 & 3/4\end{pmatrix}\begin{pmatrix}(x_1-9)\\(x_2-3)\end{pmatrix} =$$

$$\left(\left(\frac{3x_1}{4}-\frac{27}{4}\right)\left(\frac{3x_2}{4}-\frac{9}{4}\right)\right)\begin{pmatrix}(x_1-9)\\(x_2-3)\end{pmatrix} =$$

$$\left(\frac{3x_1}{4}-\frac{27}{4}\right)(x_1-9) + \left(\frac{3x_2}{4}-\frac{9}{4}\right)(x_2-3) =$$

$$\tfrac{3}{4}x_1^2 - \tfrac{27}{4}x_1 - \tfrac{27}{4}x_1 + \tfrac{243}{4} + \tfrac{3}{4}x_2^2 - \tfrac{9}{4}x_2 - \tfrac{9}{4}x_2 + \tfrac{27}{4} =$$

$$\tfrac{3}{4}x_1^2 - \tfrac{54}{4}x_1 + \tfrac{3}{4}x_2^2 - \tfrac{18}{4}x_2 + \tfrac{270}{4} = d_B^2$$

Beispiel: Lösung

… für allgemeinen Fall

5. Diskriminantenfunktionen bestimmen

$$g_B(\mathbf{x}) = -\tfrac{1}{2}\, d_B^2 - \tfrac{1}{2}\ln(16/9) = -\tfrac{3}{8}x_1^2 + \tfrac{54}{8}x_1 - \tfrac{3}{8}x_2^2 + \tfrac{18}{8}x_2 + \tfrac{270}{8} - \tfrac{1}{2}\ln(16/9) = \boxed{-\tfrac{3}{8}x_1^2 + \tfrac{27}{4}x_1 - \tfrac{3}{8}x_2^2 + \tfrac{9}{4}x_2 - \tfrac{135}{4} - \tfrac{1}{2}\ln(16/9)}$$

Beispiel: Lösung

… für allgemeinen Fall

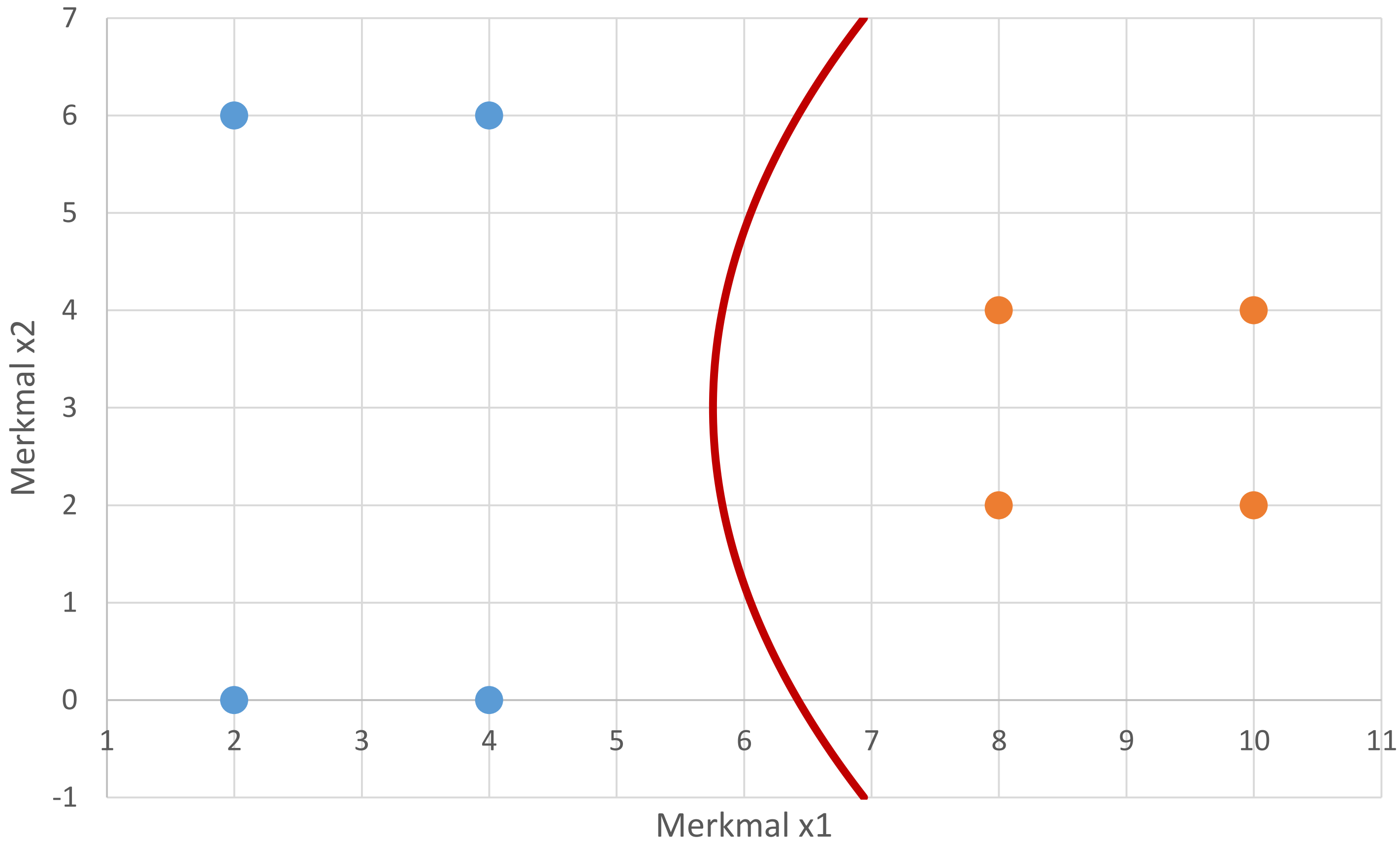
6. Entscheidungsgrenze berechnen

$$g_A(\mathbf{x}) = g_B(\mathbf{x}) \qquad 0 = g_A(\mathbf{x}) - g_B(\mathbf{x})$$

$$
\begin{aligned}
0 = {} & -\tfrac{3}{8} x_1^2 + \tfrac{9}{4} x_1 - \tfrac{1}{24} x_2^2 + \tfrac{1}{4} x_2 - \tfrac{15}{4} - \tfrac{1}{2} \ln(16) + \tfrac{3}{8} x_1^2 - \tfrac{27}{4} x_1 + \tfrac{3}{8} x_2^2 - \\
& \tfrac{9}{4} x_2 + \tfrac{135}{4} + \tfrac{1}{2} \ln(16/9) = -\tfrac{18}{4} x_1 + \tfrac{8}{24} x_2^2 - \tfrac{8}{4} x_2 + \tfrac{120}{4} - \\
& \tfrac{1}{2} \ln(16) + \tfrac{1}{2} \ln(16/9) = -\tfrac{9}{2} x_1 + \tfrac{1}{3} x_2^2 - 2 x_2 - \tfrac{1}{2} \ln(16) + \\
& \tfrac{1}{2} \ln(16/9) + 30 = \boxed{-\tfrac{9}{2} x_1 + \tfrac{1}{3} x_2^2 - 2 x_2 + 28{,}9014}
\end{aligned}
$$

Beispiel: Ergebnis

... für allgemeinen Fall



Beispiel: Lösung

… für Spezialfall 1

2. Diskriminantenfunktionen ermitteln

$$g_j(\boldsymbol{x}) = -\frac{1}{2\sigma^2}\left(-2\boldsymbol{\mu}_j^T\boldsymbol{x} + \boldsymbol{\mu}_j^T\boldsymbol{\mu}_j\right) + \cancel{\ln P(\omega_j)}$$

$$g_A(\boldsymbol{x}) = -\tfrac{1}{2}\left(-2\,(3\;\;3)\begin{pmatrix}x_1\\x_2\end{pmatrix} + (3\;\;3)\begin{pmatrix}3\\3\end{pmatrix}\right) =$$

$$-\tfrac{1}{2}\left((-6\;\;-6)\begin{pmatrix}x_1\\x_2\end{pmatrix} + 18\right) =$$

$$-\tfrac{1}{2}\left(-6x_1 - 6x_2 + 18\right) =$$

$$\boxed{3x_1 + 3x_2 - 9}$$

Beispiel: Lösung

… für Spezialfall 1

2. Diskriminantenfunktionen ermitteln

$$g_B(\mathbf{x}) = -\frac{1}{2}\left(-2\,(9\;\;3)\begin{pmatrix}x_1\\x_2\end{pmatrix} + (9\;\;3)\begin{pmatrix}9\\3\end{pmatrix}\right) =$$

$$-\frac{1}{2}\left((-18\;\;-6)\begin{pmatrix}x_1\\x_2\end{pmatrix} + 90\right. =$$

$$-\frac{1}{2}\left(-18x_1 - 6x_2 + 90\right) =$$

$$\boxed{9x_1 + 3x_2 - 45}$$

Beispiel: Lösung

… für Spezialfall 1

3. Entscheidungsgrenze berechnen

$$g_A(x) = g_B(x) \qquad 0 = g_A(x) - g_B(x)$$

$$3x_1 + 3x_2 - 9 = 9x_1 + 3x_2 - 45$$

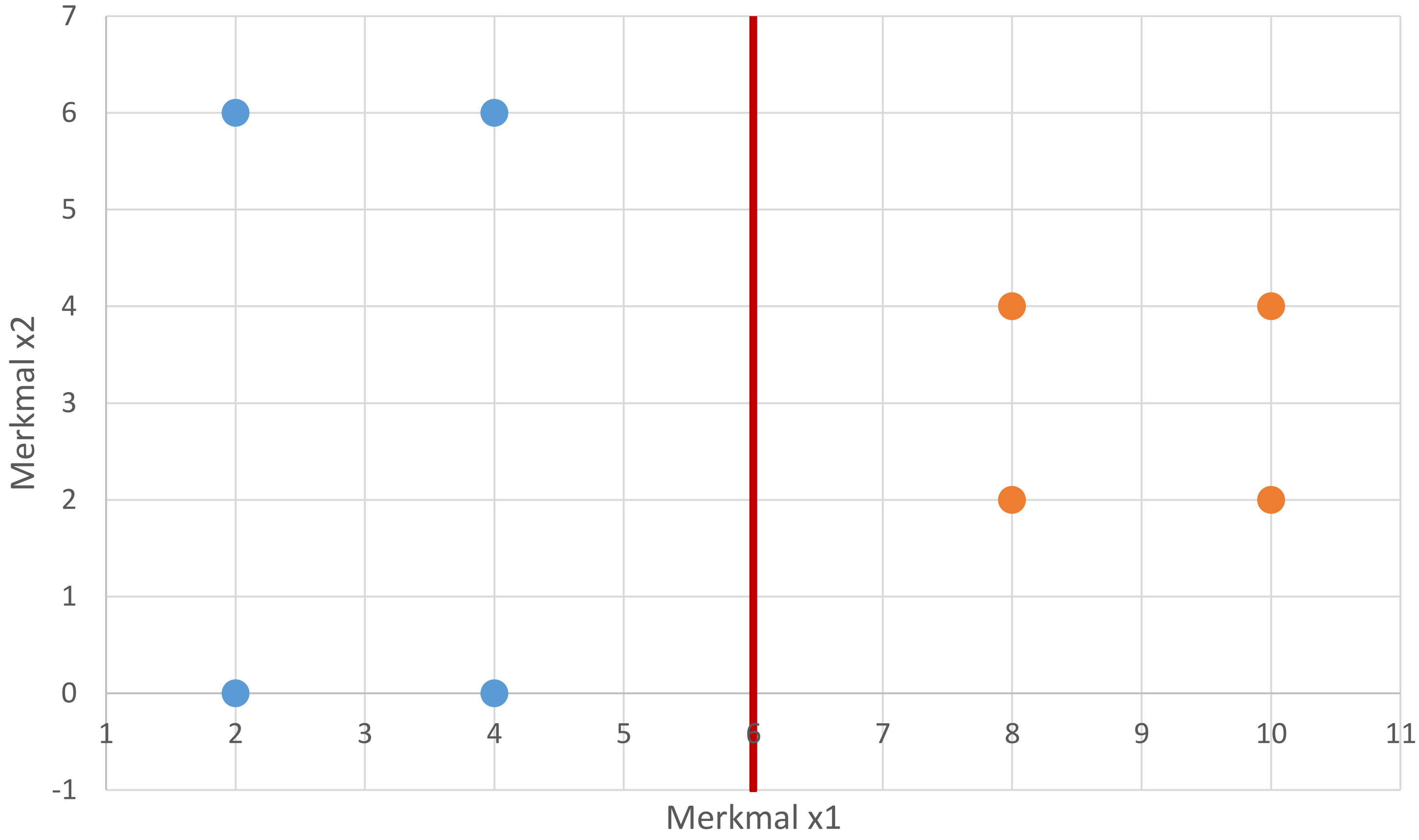
$$-6x_1 - 9 = -45$$

$$-6x_1 = -36$$

$$\boxed{x_1 = 6}$$

Beispiel: Ergebnis

... für Spezialfall 1



II. Merkmale

Merkmale

… für Muster in Bildern

- Beschreibung des „Aussehens“ (Appearance)
  - durch Template, z.B. Template im RGB-Farbraum

Eingabe

Extraktion

Template

n

m

Merkmalsvektor

$$\boldsymbol{x} \in \mathbb{R}^{n \cdot m \cdot f}$$

n … Zeilen
m … Spalten
f … Farbkanäle

Merkmale

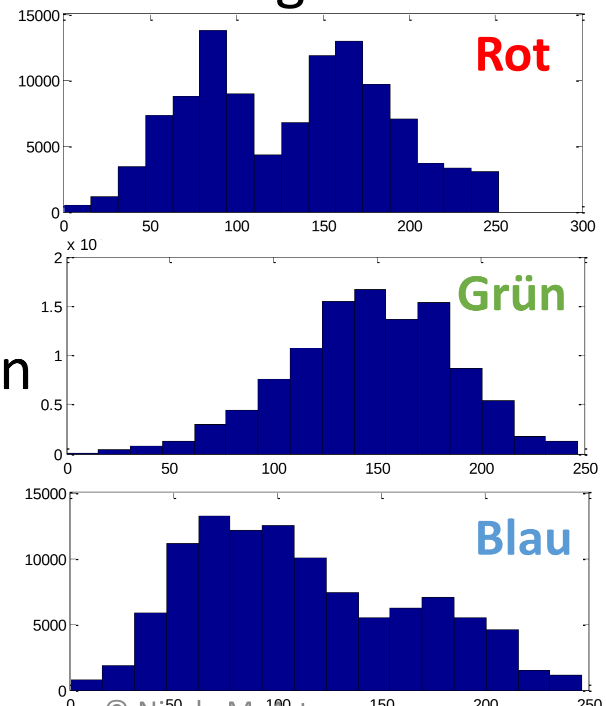
... für Muster in Bildern

- Beschreibung des „Aussehens“ (Appearance)
  - durch Histogramm, z.B. 3D-Histogramm mit 16 Bins

Eingabe

Extraktion

Histogramme



Merkmalsvektor

$$\boldsymbol{x} \in \mathbb{R}^{b \cdot f}$$

b ... Bins

f ... Farbkanäle

Merkmale

… für Muster in Bildern

- Beschreibung geometrischer Eigenschaften durch:

Position $\boldsymbol{x} \in \mathbb{R}^2$

Bounding Box $\boldsymbol{x} \in \mathbb{R}^4$

Konvexe Hülle $\boldsymbol{x} \in \mathbb{R}^{2n}$ n…Points

Kontur …

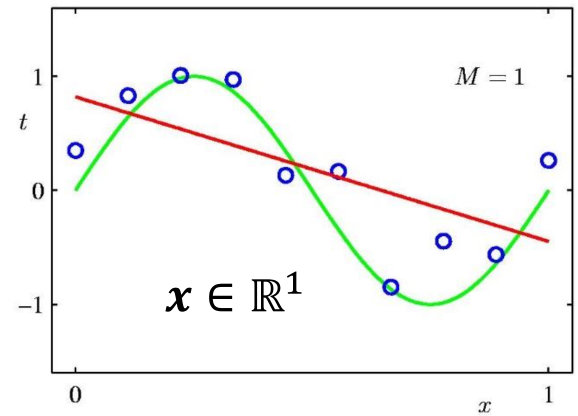
Silhouette …

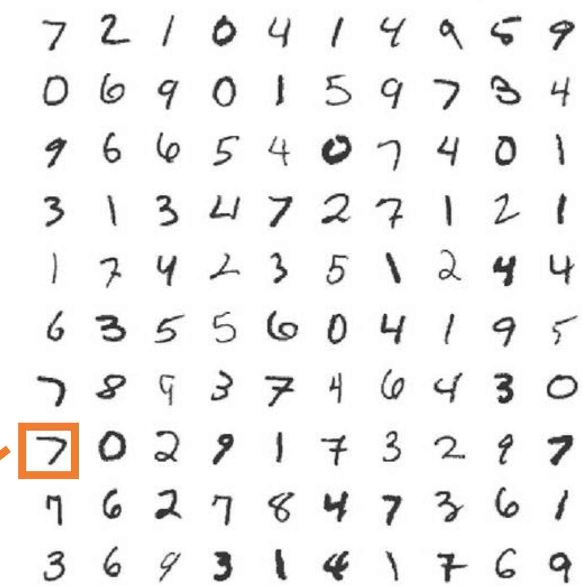
Skeleton …

Dimensionalität von

… Merkmalsvektoren $\boldsymbol{x} \in \mathbb{R}^n$

- viele Beispiele dieser Vorlesung verwenden Merkmale/Merkmals-vektoren im $\mathbb{R}^1$ und $\mathbb{R}^2$
- in praktischen Anwendungen der Mustererkennung sind Merkmals-vektoren oft hochdimensional
  - z.B.: optische Buchstabenerkennung (OCR = optical character recognition)





$\boxed{\boldsymbol{x} \in \mathbb{R}^{n \cdot m}}$

[Quelle: Bishop, 2006]

Bemerkungen zu Merkmalen

- „Gute“ Merkmale sind diskriminierend und einfach zu berechnen

- Wenn mehrere Merkmale kombiniert werden, sollte jedes Merkmal unterschiedliche Information enthalten
  - im Idealfall sind Merkmale nicht stark korreliert

- Viele Merkmale enthalten nicht zwingend mehr Information als wenige Merkmale
  - sonst würde man z.B. bei Mustern in Bildern einfach alle Pixelwerte verwenden die das Muster beschreiben

Bemerkungen zu Merkmalen

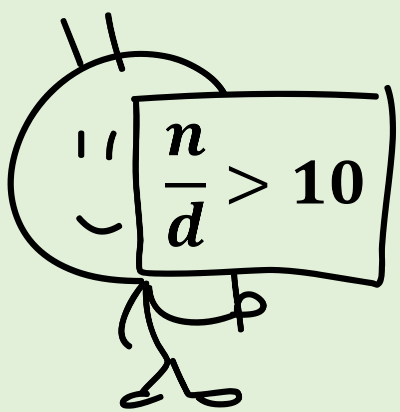
- Generell neigt man dazu die Anzahl der Merkmale zu erhöhen, wenn man mit dem Klassifikationsergebnis nicht zufrieden ist.

- Mehr Merkmale → eine höhere Dimensionalität ist gerechtfertigt, wenn die Ergebnisse des Klassifikators verbessert werden können.

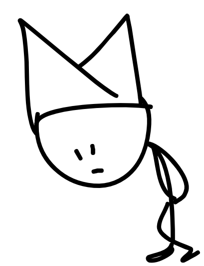
- Leider ist es nicht immer der Fall, dass mehr Merkmale zu besseren Ergebnissen führen müssen.

Faustregel

Sei n die Größe des Trainingsdatensatzes pro Klasse und d die Anzahl der Merkmale. Dann sollte das Verhältnis zwischen Trainingsdaten und Merkmale wie folgt aussehen:

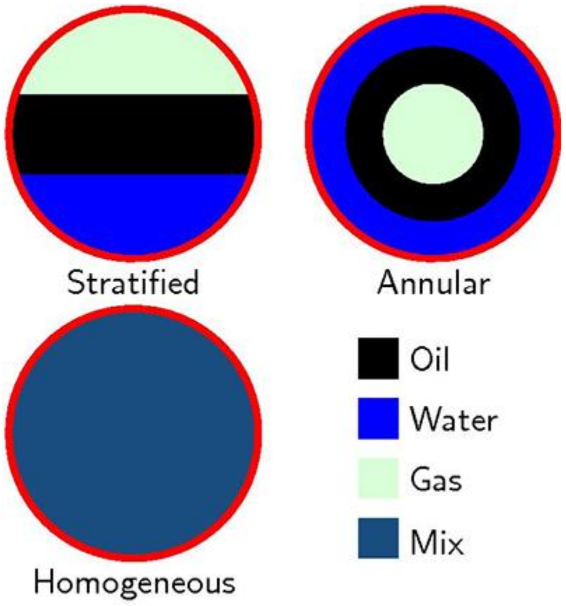


III. Fluch der Dimensionalität



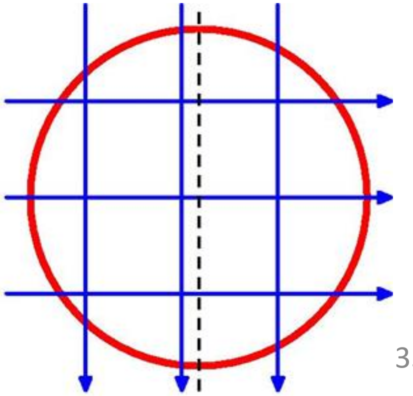
Beispiel

[Quelle: Bishop, 2006]



- synthetischer Datensatz
  - Messungen der Verhältnisse von Öl, Wasser und Gas mit Hilfe von Gammastrahlen
  - Merkmalsvektor $\boldsymbol{x} \in \mathbb{R}^{12}$:
    - 6 Messungen mit je zwei Gammastrahlen unterschiedlicher Energie
  - drei Klassen: „Stratified“, „Annular“ und „Homogeneous“

6 Messungen

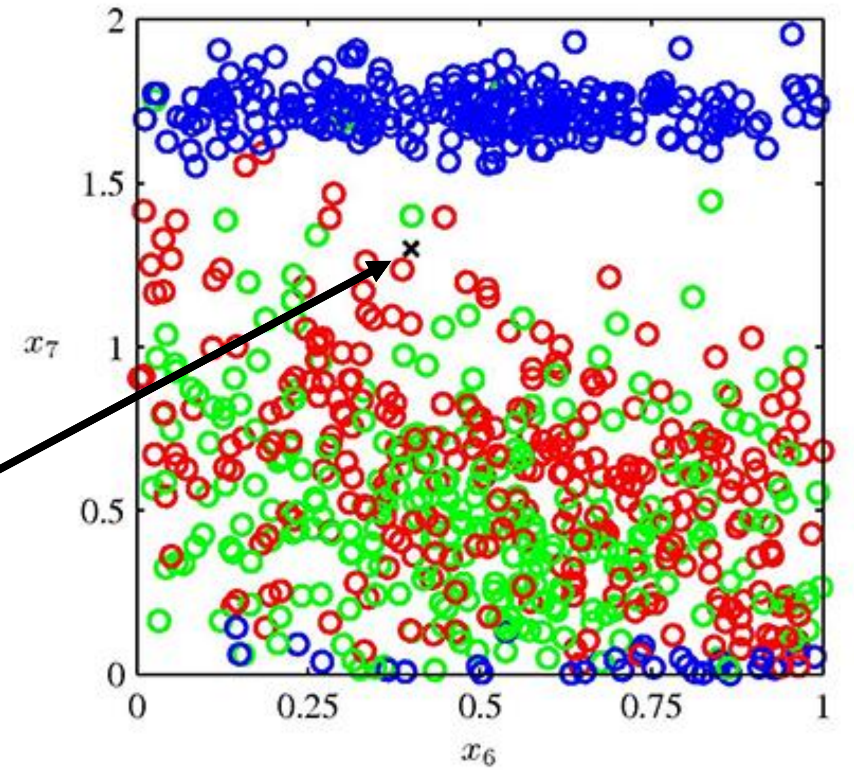


Beispiel

- Trainingsdaten: 100 Datenpunkte $\in \mathbb{R}^2$
  - Merkmale x_6 und x_7 der drei Klassen „Homogeneous“, „Annular“ und „Stratified“

- Ziel: Klassifikation einer neuen Beobachtung $\boldsymbol{x} = (x_6, x_7)$

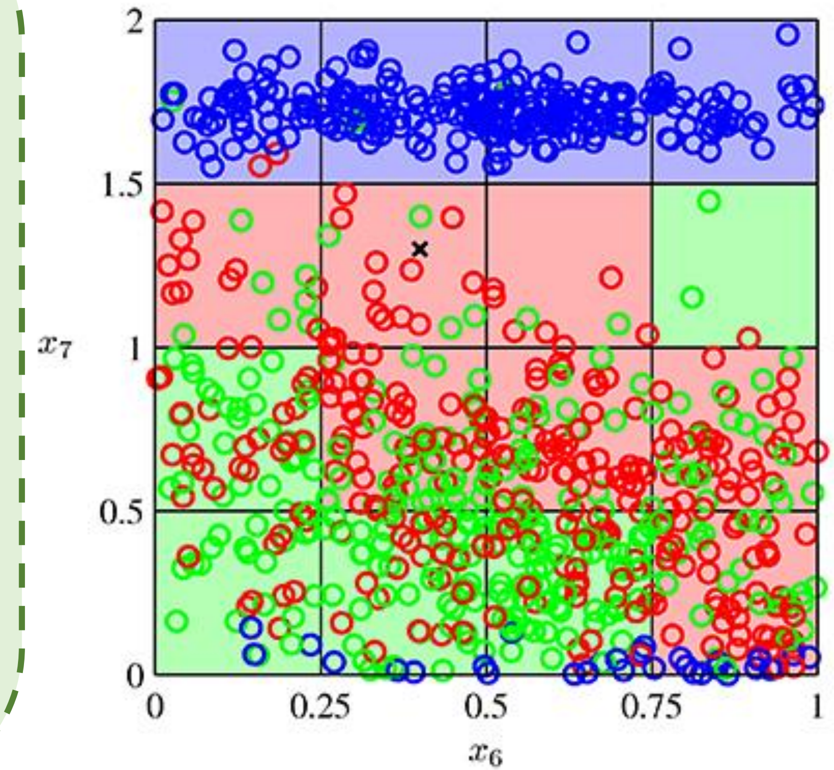
[Quelle: Bishop, 2006]



Beispiel

Einfaches Verfahren:
- unterteile den Merkmalsraum des Trainingsdatensatzes in gleich große Zellen (Nachbarschaften)
- finde heraus in welche Zelle die Beobachtung $\boldsymbol{x}$ fällt
- weise $\boldsymbol{x}$ der Klasse mit den meisten Datenpunkte in dieser Zelle zu

[Quelle: Bishop, 2006]



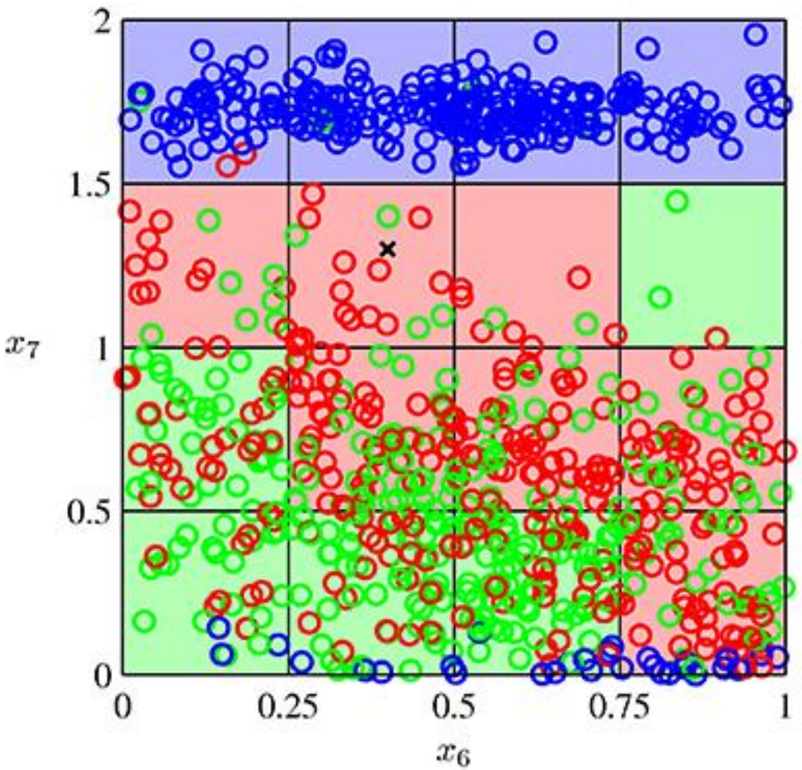
Diese naive Lösung hat einige Probleme ….

Beispiel

Eines der **gravierendsten** Probleme wird klar, wenn man die **Dimensionalität des Merkmalsraums** erhöht:

- $\mathbb{R}^2$ → 16 Zellen, 4^2 (siehe Abbildung)
- $\mathbb{R}^3$ → 64 Zellen, 4^3
- …
- $\mathbb{R}^{12}$ → 16.777.216 Zellen, 4^{12}

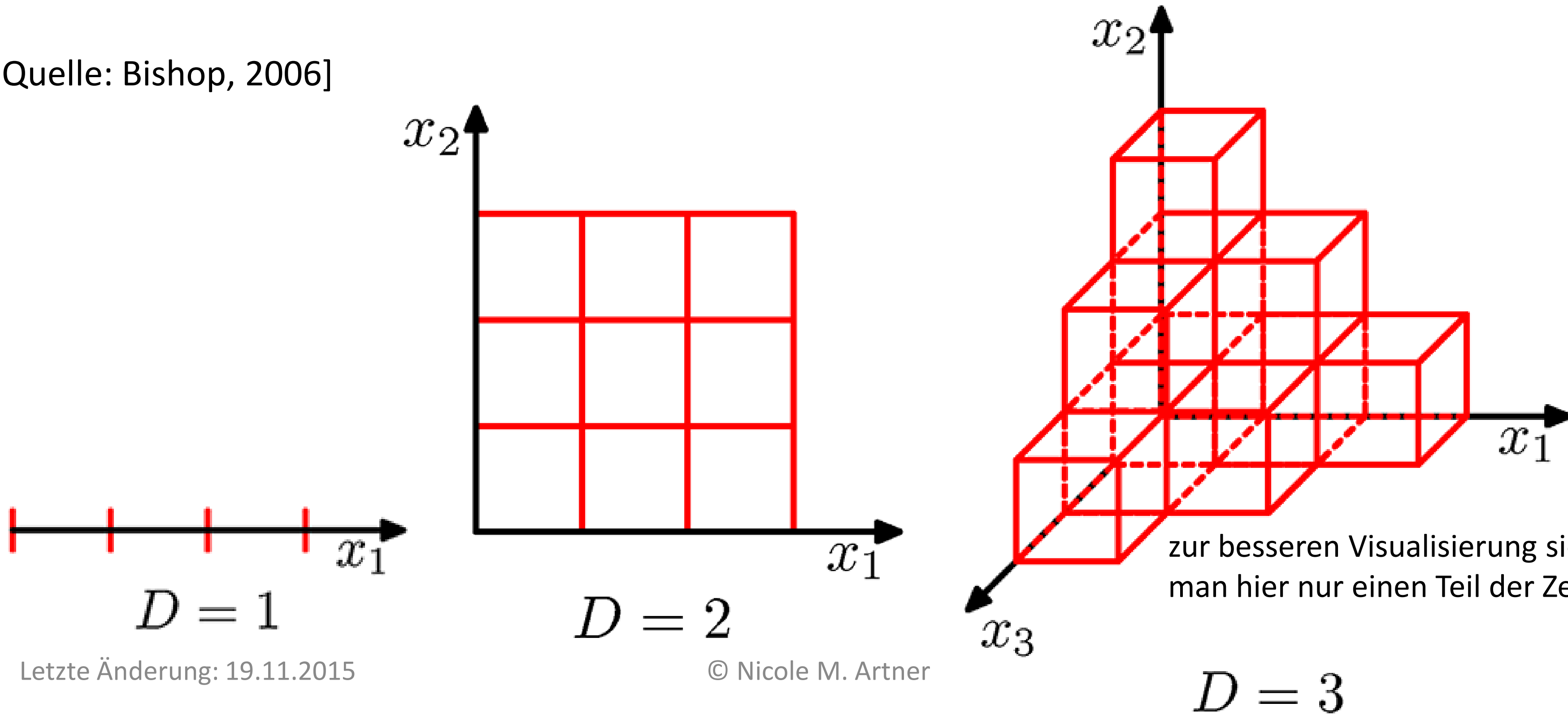
[Quelle: Bishop, 2006]



Beispiel

- Anzahl der Zellen wächst **exponentiell** mit den Dimensionen des Merkmalsraums
- es ist schwierig sicherzustellen, dass Zellen in hoch-dimensionalen Räumen **nicht leer** sind (keine Trainingsdaten vorhanden)

[Quelle: Bishop, 2006]



„Curse of dimensionality“

… zu Deutsch: Fluch der Dimensionalität

Der Fluch der Dimensionalität ist das exponentielle Wachstum der Möglichkeiten als eine Funktion der Dimensionalität. Wenn die Dimensionalität wächst,

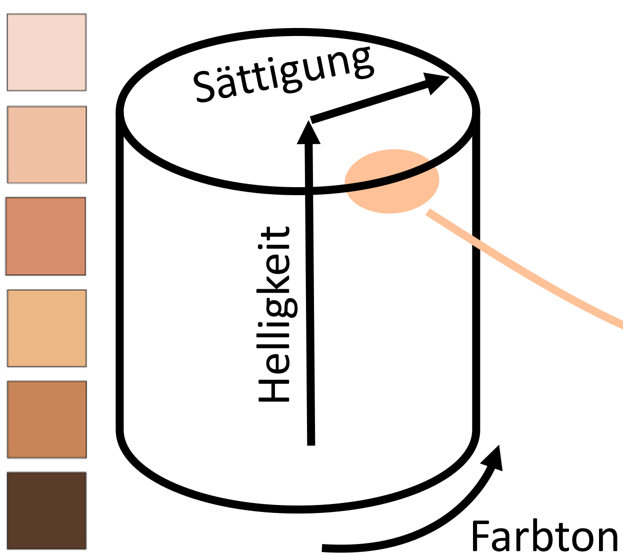
- dann ist eine **exponentiell wachsende Menge an Daten** notwendig, um den Merkmalsraum abzudecken und Training zu ermöglichen;
- dann wächst auch der **Rechenaufwand** und die **Komplexität**;

Ins Deutsche übersetzt, aus „Dictionary of Computer Vision and Image Processing“, Fisher et al., 2005

VI. Merkmalsselektion

Grundregel

Wähle Merkmale, sodass im Merkmalsraum
- der Abstand zwischen unterschiedlichen Klassen groß ist
- und die Varianz innerhalb der Klassen klein



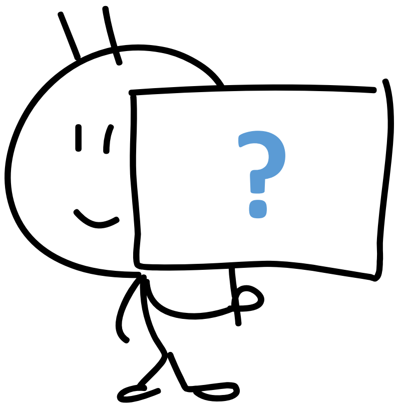
- z.B. Klassifikation von Pixeln in „Haut“ oder „nicht Haut“
- Mögliche Merkmale:
  - Rot, Grün und Blau (RGB-Farbraum) ODER
  - Farbton, Sättigung und Helligkeit (HSV-Farbraum)
- Wahl:
  - Farbton + Sättigung um unabhängig vom Hautton (Helligkeit) zu sein → Varianz entlang dieser Achse ist am größten

Vorgehensweise

Ziel: Auswahl einer Teilmenge aus allen möglichen Merkmalen

Wie kann man entscheiden welche Merkmale ausgewählt werden?

1. individuell für jedes Merkmal x
2. für Merkmalsvektoren $\boldsymbol{x}$ bestehend aus unterschiedlichen Kombinationen von Merkmalen

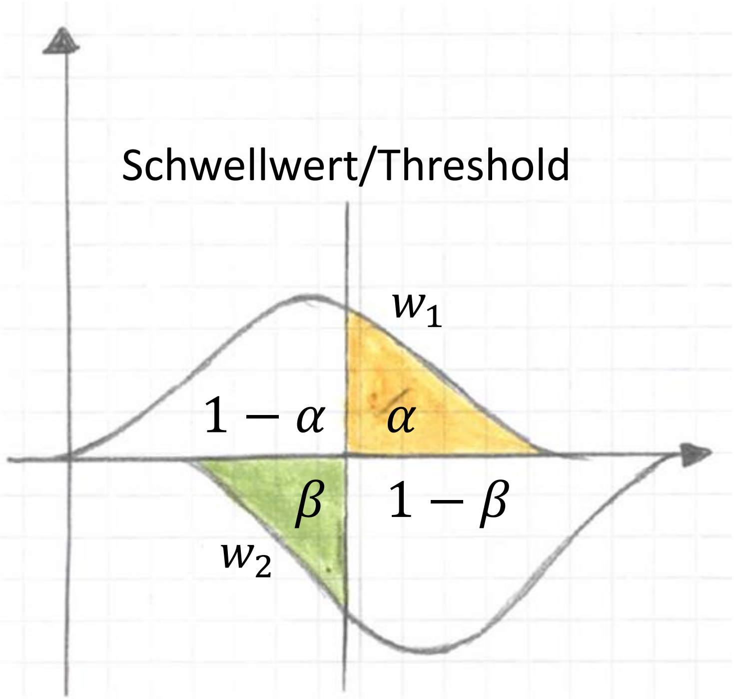


Entscheidungskriterien

- Trennbarkeit der Klassen → unabhängig vom Klassifikator
- Performance des Klassifikators (Fehlerrate) → abhängig von einem bestimmten Klassifikator

Entscheidungskriterium: ROC

ROC = Receiver Operating Characteristic

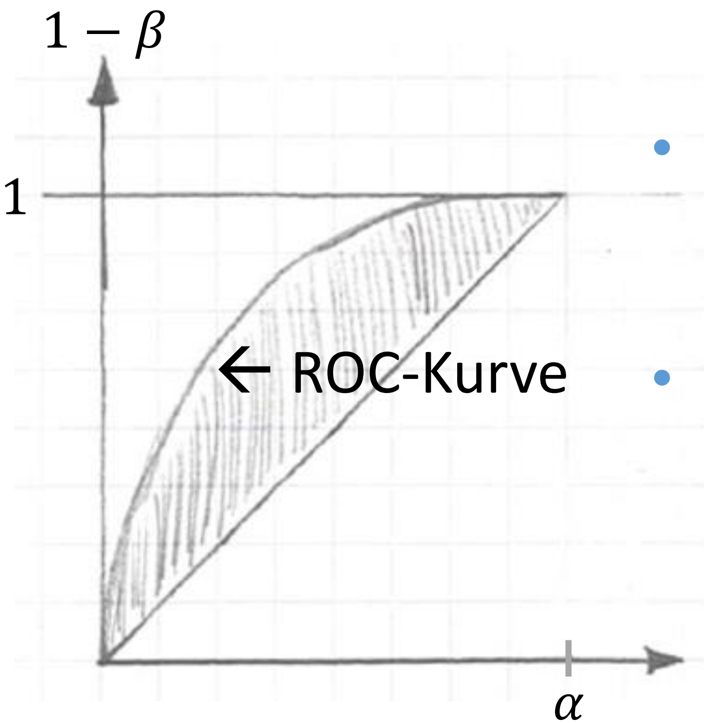


- Quantifizierung der Überlappung von Klassen (in einem Merkmal)
- α … Fehlerwahrscheinlichkeit Klasse w_1
  (gelbe Fläche unter der Kurve)
- β … Fehlerwahrscheinlichkeit Klasse w_2
  (grüne Fläche unter der Kurve)

Achtung: Nicht zu verwechseln mit der **bedingten Fehlerwahrscheinlichkeit** (posterior der nicht gewählten Klasse).

Entscheidungskriterium: ROC

- wenn man den Schwellwert über „alle“ möglichen Positionen schiebt, erhält man unterschiedliche Werte für α und β



- vollständiger Überlapp zwischen Klassen → $\alpha = 1 - \beta$ → gerade Linie in Abbildung
- umso weiter Klassen voneinander entfernt → umso größer ist die Fläche zwischen Kurve und gerader Linie

Entscheidungskriterium: FDR

FDR = Fisher‘s discriminant ratio

FDR für ein Merkmal ist wie folgt definiert:

für zwei Klassen

$$FDR = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}$$

für M Klassen

$$FDR_1 = \sum_{i}^{M} \sum_{j \neq i}^{M} \frac{\left(\mu_i - \mu_j\right)^2}{\sigma_i^2 + \sigma_j^2}$$

M ...Anzahl der Klassen

- je größer FDR eines Merkmals umso besser
  - Idealfall: Mittelwerte der Klassen sind weit entfernt und Varianz innerhalb der Klassen ist klein (siehe Folie 40)

Auswahlprozess

… für individuelle Merkmale

1. **Entscheidungskriterium** k wird für alle Merkmale m ermittelt (z.B. FDR)

2. **Merkmale** m werden absteigend/aufsteigend (je nach Kriterium) nach k **sortiert**

3. die **ersten $\boldsymbol{l}$** (Länge des Vektors) Merkmale werden **ausgewählt**

Auswahlprozess

… für Merkmalsvektoren

- die „einfachste“ Lösung wäre einfach alle möglichen Permutationen (Länge des Vektors l, Merkmale m) zu testen
- mit Hilfe der Entscheidungskriterien entscheidet man sich dann für die „beste“ Teilmenge an Merkmalen

Die vollständige Suche (exhaustiv search) nach der besten Teilmenge ist praktisch nur für eine kleine Anzahl an Merkmalen möglich. Problem: kombinatorische Explosion!

- Lösung: Prozess durch ein iteratives Verfahren approximieren
    - es kann nicht garantiert werden, dass die „ideale“ Teilmenge gefunden wird

Sequential Backward Selection

… sequentielle Rückwärtsauswahl (frei übersetzt)

Beispiel: $m = 4$, Merkmale x_1, x_2, x_3, x_4, l soll 2 sein

1. Berechne **Entscheidungskriterium** für $[x_1, x_2, x_3, x_4]^T$
2. **Reduziere die Länge** des Vektors um ein Merkmal und ermittle Entscheidungskriterium für $[x_1, x_2, x_3]^T$, $[x_1, x_2, x_4]^T$, $[x_1, x_3, x_4]^T$ und $[x_2, x_3, x_4]^T$
3. Wähle die **beste Kombination**, z.B.: $[x_1, x_2, x_3]^T$
4. **Reduziere wieder** die Länge des Vektors um eins und ermittle Entscheidungskriterium für $[x_1, x_2\ ]^T$, $[x_1, x_3]^T$ und $[x_2, x_3]^T$
5. Wähle die **beste Kombination**, das Ergebnis der Auswahl, z.B.: $[x_1, x_2\ ]^T$

Sequential Backward Selection

… sequentielle Rückwärtsauswahl (frei übersetzt)

Allgemeiner formuliert:

Beginnend mit einem Merkmalsvektor der Länge m entfernt man jeweils ein Merkmal von der „besten" Kombination (abhängig vom Entscheidungskriterium) bis man einen Merkmalsvektor der Länge l erhält.

Anzahl der durchsuchten Kombinationen:

$$1 + \frac{1}{2}\big((m+1)m - l(l+1)\big)$$

Im vorherigen Beispiel:

$$1 + \frac{1}{2}\big((4+1)4 - 2(2+1)\big) = 8$$

Sequential Forward Selection

… sequentielle Vorwärtsauswahl (frei übersetzt)

Beispiel: $m = 4$, Merkmale x_1, x_2, x_3, x_4, l soll 2 sein

1. Berechne **Entscheidungskriterium für jedes Merkmal**: x_1, x_2, x_3 und x_4
2. Wähle **das „beste" Merkmal** (basierend auf Entscheidungskriterium), z.B.: x_1
3. **Bilde alle möglichen Vektoren** $l = 2$ die Merkmal x_1 enthalten: $[x_1, x_2]^T$, $[x_1, x_3]^T$, $[x_1, x_4]^T$
4. Ermittle **Entscheidungskriterium** für alle Vektoren und wähle den „besten", z.B.: $[x_1, x_2]^T$

…wenn $l = 3$, dann müsste der Prozess weitergehen

Sequential Forward Selection

… sequentielle Vorwärtsauswahl (frei übersetzt)

Allgemeiner formuliert:

Beginnend mit den einzelnen Merkmalen m bildet man Merkmalsvektoren wachsender Länge basierend auf der bisherigen „besten“ Kombination (Merkmal) bis man einen Merkmalsvektor der Länge l erhält.

Anzahl der durchsuchten Kombinationen:

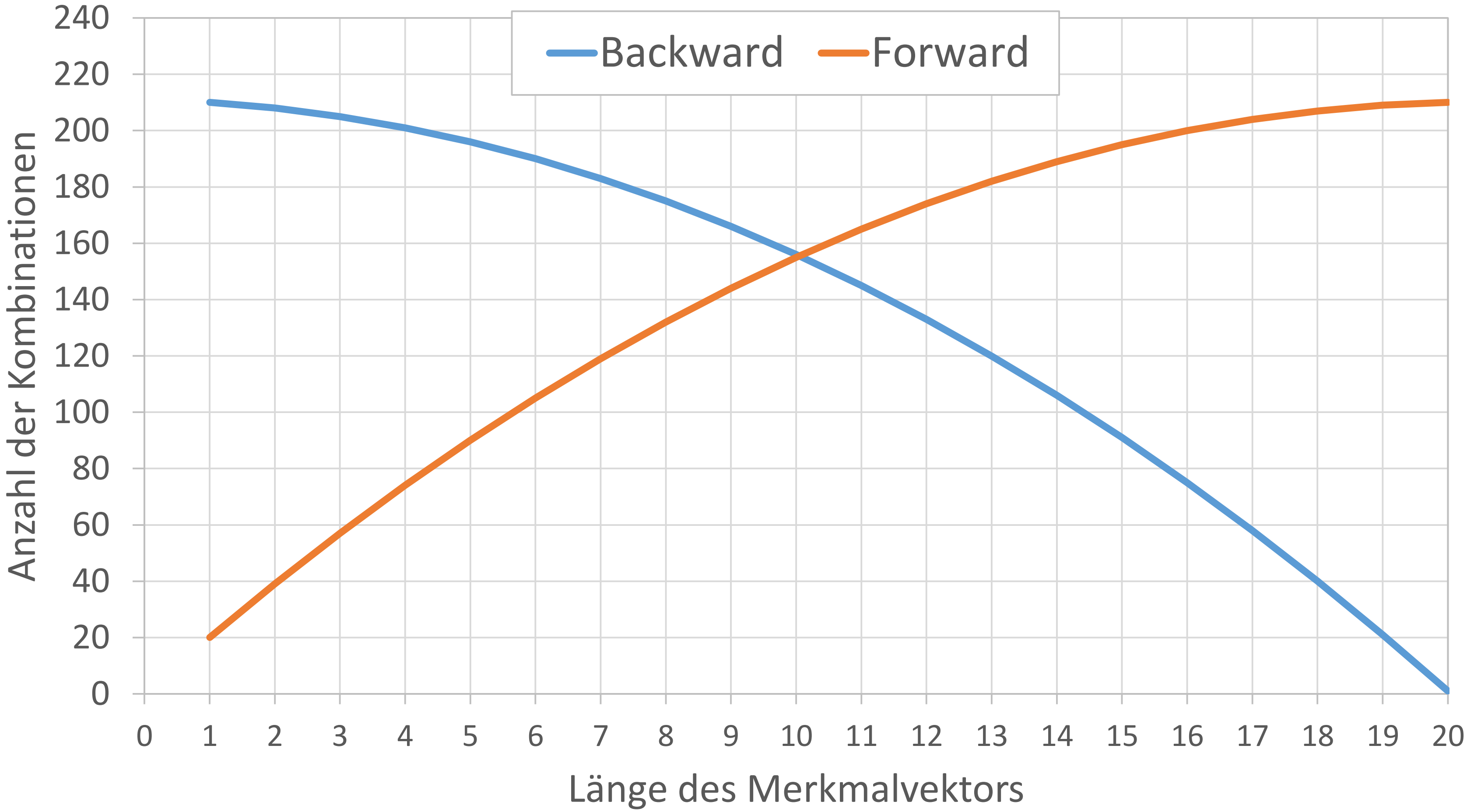
$$lm - \frac{l(l-1)}{2}$$

Im vorherigen Beispiel:

$$8 - \frac{2(2-1)}{2} = 7$$

Sequential Forward vs. Sequential Backward

- für $m = 20$



V. Merkmalsreduktion

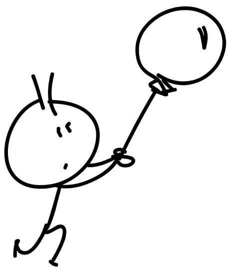
Grundlagen

Man spricht von einer **Merkmalsreduktion**, wenn man die Dimensionalität und die damit verbundenen Probleme reduziert in dem man Merkmale geschickt kombiniert.

Mögliche Verfahren:

- LDA (=Linear Discriminant Analysis): maximiert die Separierbarkeit der Klassen, basiert auf FDR (siehe Folie 44)
- PCA (=Principal Component Analysis): erzeugt unkorrelierte Merkmale
- ICA (=Independant Component Analysis): erzeugt unabhängige Merkmale

Grundidee



... der Merkmalsreduktion (PCA und ICA)

- die vorhandenen Merkmale zu transformieren
- relevante Information für Klassifikation in weniger Merkmalen zu konzentrieren
- redundante Information zu finden und zu entfernen
- reduzieren der Dimensionalität des Merkmalraums

Lineare Transformationen

… sind besonders attraktiv zur Merkmalsreduktion, weil sie einfach zu berechnen sind

Transformierter Merkmalsvektor

$$\boxed{\boldsymbol{x}} = \boxed{\boldsymbol{U}}\boxed{\boldsymbol{y}} \quad \xleftrightarrow[\text{Invertierbar}]{\hspace{3em}} \quad \boldsymbol{y} = \boxed{\boldsymbol{U}^{-1}}\boldsymbol{x}$$

Lineare Transformation

Merkmalsvektor

Inverse, lineare Transformation

Lineare Transformationen

Ziel: Finde lineare Transformation $\boldsymbol{U}$, welche den Merkmalsvektor $\boldsymbol{x} \in R^d$ reduziert zu $\boldsymbol{y} \in R^{d'}$, wobei $d' < d$

- lineare Transformationen können hoch-dimensionale Merkmalsvektoren in Merkmalsräume mit niedriger Dimensionalität projizieren

- Vorteile
  - reduzierte Komplexität der Schätzung (z.B.: pdf) und Klassifikation (z.B.: Entscheidungsgrenze)
  - ermöglicht die Visualisierung von multivariaten Merkmalsvektoren in zwei oder drei Dimensionen

WH: Lineare Algebra

Ein p-dimensionaler Vektor $\boldsymbol{x}$ kann ohne Fehler durch die Summe von p linear unabhängigen Vektoren dargestellt werden:

$$\boldsymbol{x} = \sum_{i=1}^{p} y_i \boldsymbol{u}_i = \boldsymbol{U}\boldsymbol{y}$$

$\boldsymbol{u}_i$ ... Basisvektoren, Spalten von der Basis $\boldsymbol{U}$

y_i ... Koeffizienten bezüglich der Basis $\boldsymbol{U}$

Beispiel

$$\begin{pmatrix} 5 \\ 2 \\ -1 \end{pmatrix} = 5\boldsymbol{u}_1 + 2\boldsymbol{u}_2 - \boldsymbol{u}_3$$

$$\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{u}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Hauptkomponentenanalyse

… im Englischen „Principal Component Analysis“ (PCA)

- ist eine Möglichkeit wie man eine reduzierte Basis $\boldsymbol{U}$ finden kann mit $m \ll p$, wobei der Rekonstruktionsfehler ε^2 minimiert wird

$$\boldsymbol{x} = \boxed{\sum_{i=1}^{p} y_i \boldsymbol{u}_i}$$

fehlerfreie Rekonstruktion

$$\hat{\boldsymbol{x}}(m) = \sum_{i=1}^{m} y_i \boldsymbol{u}_i$$

$$\boldsymbol{x} = \hat{\boldsymbol{x}}(m) + \boxed{\varepsilon^2}$$

Rekonstruktionsfehler

Hauptkomponentenanalyse

Die Hauptkomponentenanalyse (PCA) findet eine **reduzierte Basis** $\boldsymbol{U}$ die den Rekonstruktionsfehler ε^2 minimiert, weil sie als Basisvektoren die **Eigenvektoren** $\boldsymbol{u}_i$ mit den m **größten Eigenwerten** λ_i der Kovarianzmatrix $\widehat{\boldsymbol{\Sigma}}$ des ursprünglichen Merkmalsvektors $\boldsymbol{x}$ verwendet.

$$\widehat{\boldsymbol{\Sigma}}\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i$$

Eigenvektor & Eigenwert

Wenn man eine Matrix A mit einem Vektor $\boldsymbol{v}$ multipliziert und das Ergebnis ist ein Vielfaches $\lambda\boldsymbol{v}$ des Originalvektors $\boldsymbol{v}$, dann ist $\boldsymbol{v}$ ein Eigenvektor und der Faktor λ ein Eigenwert.

$$A\boldsymbol{v} = \lambda\boldsymbol{v}$$

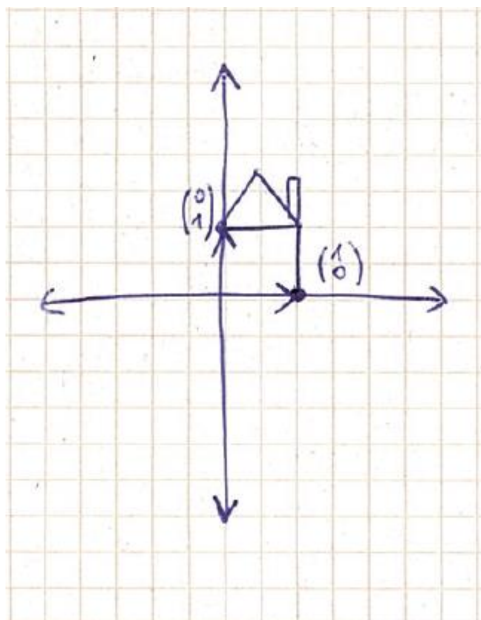
Geometrisch interpretiert: Die Matrix macht den Vektor zu einem Vektor der parallel ist.

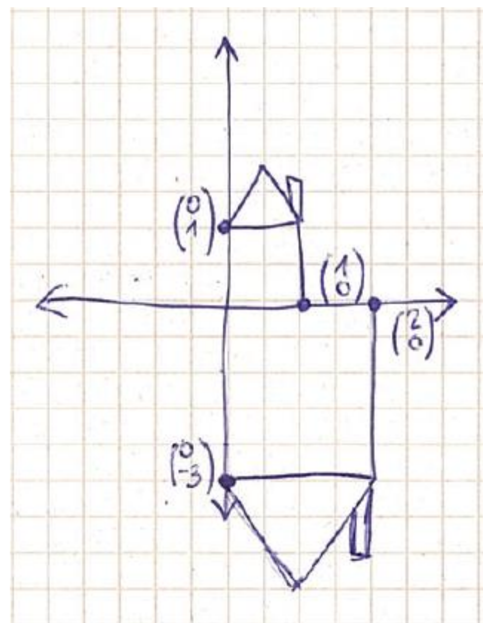
Regel: Eigenvektoren dürfen keine Nullvektoren sein!

Beispiel

Eigenvektoren und Eigenwerte von: $A = \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}$

Fragen: Welcher Vektor wird durch diese Matrix zu einem Vielfachen von sich selbst?



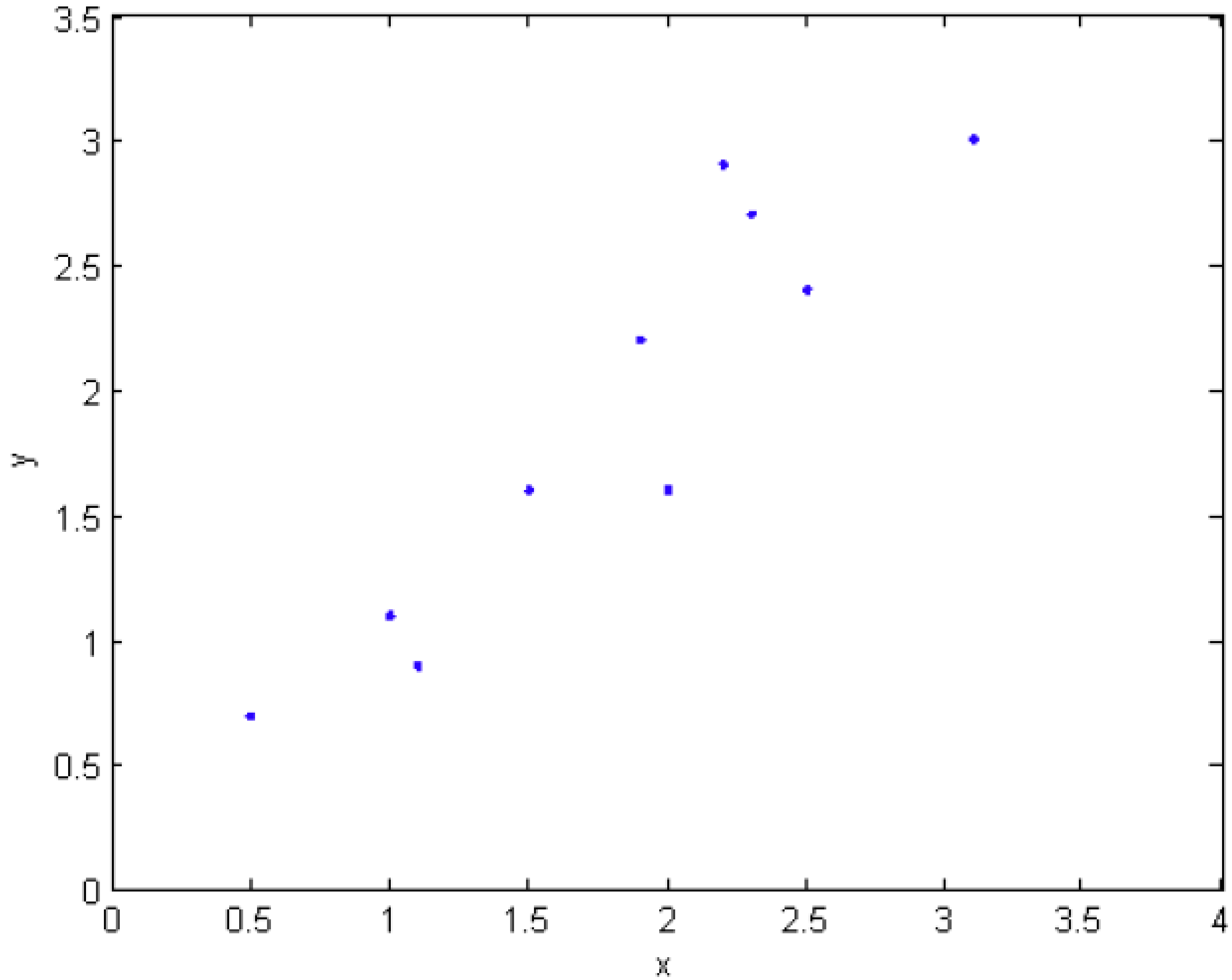


Eigenvektoren: $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ Eigenwerte: 2 und -3

zurück zur Merkmalsreduktion ....

Beispiel

$x_{1,k}$	$x_{2,k}$
2.5	2.4
0.5	0.7
2.2	2.9
1.9	2.2
3.1	3.0
2.3	2.7
2	1.6
1	1.1
1.5	1.6
1.1	0.9



Beispiel

- Mittelwert Vektor:

$$\hat{\boldsymbol{\mu}} = \begin{pmatrix} 1{,}81 \\ 1{,}91 \end{pmatrix}$$

- Differenz Vektoren:

$$\Phi_k = \boldsymbol{x}_k - \hat{\boldsymbol{\mu}}$$

- Kovarianzmatrix:

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{9} \sum_{k=1}^{10} \Phi_k \Phi_k^T$$

$\Phi_{1,k}$	$\Phi_{2,k}$
.69	.49
-1.31	-1.21
.39	.99
.09	.29
1.29	1.09
.49	.79
.19	-.31
-.81	-.81
-.31	-.31
-.71	-1.01

Beispiel: Eigenwerte

- Ergebnis für Kovarianzmatrix: $\hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 0{,}6166 & 0{,}6154 \\ 0{,}6154 & 0{,}7166 \end{pmatrix}$

- Berechnung der zwei Eigenwerte λ_1 und λ_2:

$$\boxed{det}(\hat{\boldsymbol{\Sigma}} - \lambda \boxed{E}) = 0$$

Determinante

$$E = \begin{pmatrix} 1 & 0 & \\ 0 & 1 & \\ & & \ddots \end{pmatrix}$$

Beispiel: Eigenwerte

$$\boxed{det\left(\widehat{\boldsymbol{\Sigma}} - \lambda E\right)} = 0$$

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

$$\begin{vmatrix} 0{,}6166 - \lambda & 0{,}6154 \\ 0{,}6154 & 0{,}7166 - \lambda \end{vmatrix} = \boxed{\lambda^2 - 1{,}3332\lambda + 0{,}0632}$$

Polynom 2. Grades

(charakteristisches Polynom der Matrix)

Die Nullstellen des Polynoms sind die Eigenwerte λ_1 und λ_2.
Ergebnisse:

$$\lambda_1 = 1{,}2840, \lambda_2 = 0{,}0491$$

Beispiel: Eigenvektoren

- zur Bestimmung der Eigenvektoren löst man folgendes Gleichungssystem für jeden Eigenwert λ_i

$$\widehat{\boldsymbol{\Sigma}}\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i \Rightarrow \left(\widehat{\boldsymbol{\Sigma}} - \lambda_i E\right)\boldsymbol{u}_i = \vec{0}$$

- z.B.: für λ_1

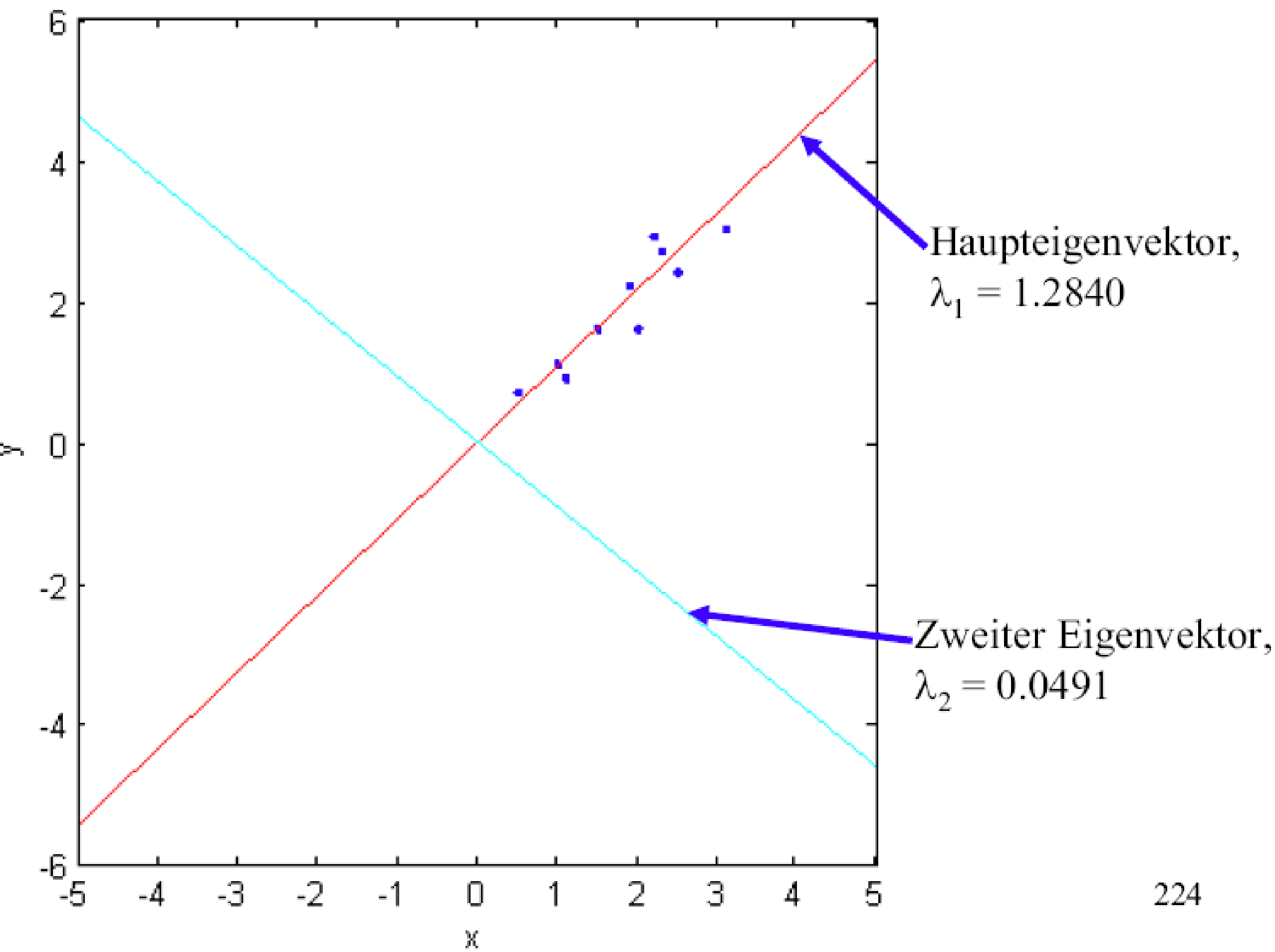
$$\begin{pmatrix} 0,6166 - 1,2840 & 0,6154 \\ 0,6154 & 0,7166 - 1,2840 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

- Ergebnisse: Normalisierte Eigenvektoren

Länge ~1

$$\boldsymbol{u}_1 = \begin{pmatrix} \boxed{\begin{matrix} 0,6779 \\ 0,7352 \end{matrix}} \end{pmatrix} \qquad \boldsymbol{u}_2 = \begin{pmatrix} -0,7352 \\ 0,6779 \end{pmatrix}$$

Beispiel: Eigenvektoren



Der Haupteigen-vektor liegt entlang der Richtung der größten Varianz der Trainingsdaten.

Koeffizienten

- Man kann jetzt die Vektoren $\boldsymbol{x_k}$ als eine Summe (linear Kombination) der neuen Basisvektoren schreiben:

$$\boldsymbol{x} = \sum_{i=1}^{p} y_i \boldsymbol{u}_i = \boldsymbol{U}\boldsymbol{y}$$

- Die Koeffizienten y_i werden folgendermaßen berechnet:

$$y_{\boldsymbol{i}} = \boldsymbol{u}_i^T \boldsymbol{x} \qquad\qquad \boldsymbol{y} = \boldsymbol{U}^T \boldsymbol{x}$$

Beispiel: Koeffizienten

- Wenn man die berechneten Eigenvektoren als Basisvektoren annimmt, dann hat man

$$\boldsymbol{U} = \begin{pmatrix} \boxed{\begin{matrix} 0{,}6779 \\ 0{,}7352 \end{matrix}} & \boxed{\begin{matrix} -0{,}7352 \\ 0{,}6779 \end{matrix}} \end{pmatrix}$$

$$\boldsymbol{u}_1 \qquad \boldsymbol{u}_2$$

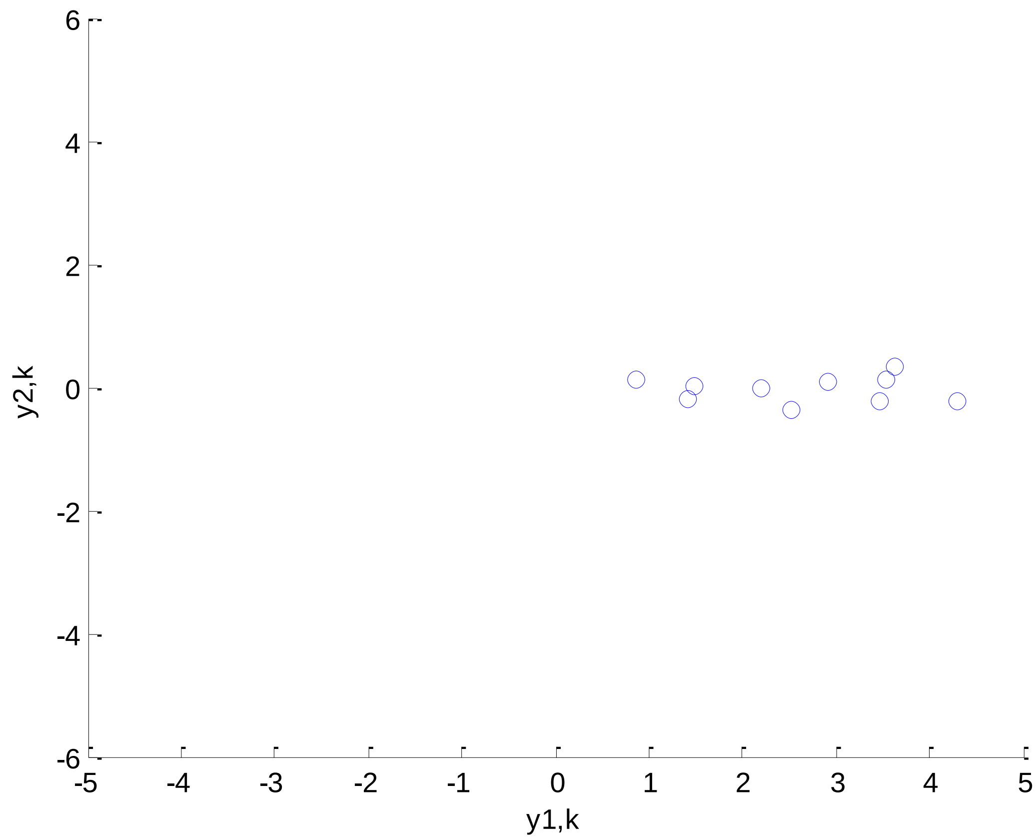
- das ergibt für $\boldsymbol{x}_1$

$$\boldsymbol{x}_1 = \begin{pmatrix} 2{,}5 \\ 2{,}4 \end{pmatrix} \Rightarrow \boldsymbol{y}_1 = \begin{pmatrix} \boxed{0{,}6779 \quad 0{,}7352} \\ \boxed{-0{,}7352 \quad 0{,}6779} \end{pmatrix} \begin{pmatrix} 2{,}5 \\ 2{,}4 \end{pmatrix} = \begin{pmatrix} 3{,}4592 \\ -0{,}2110 \end{pmatrix}$$

Beispiel: Koeffizienten

Koeffizienten für $\boldsymbol{x}_k$

$y_{1,k}$	$y_{2,k}$
3,4592	-0,2110
0,8536	0,1069
3,6235	0,3485
2,9055	0,0945
4,3071	0,1394
3,5442	-0,2454
2,5321	-0,3858
1,4866	0,0105
2,1932	-0,0182
1,4074	-0,1986



Beispiel: Reduktion

- Reduzieren der Basis → in unserem Beispiel nehmen wir nur den Eigenvektor des größten Eigenwertes, d.h.:

$$\boldsymbol{U} = \begin{pmatrix} 0{,}6779 \\ 0{,}7352 \end{pmatrix}$$

- Für $\boldsymbol{x}_1$

$$\boldsymbol{x}_1 = \begin{pmatrix} 2{,}5 \\ 2{,}4 \end{pmatrix} \Rightarrow \boldsymbol{y}_1 = \begin{pmatrix} 0{,}6779 & 0{,}7352 \end{pmatrix} \begin{pmatrix} 2{,}5 \\ 2{,}4 \end{pmatrix} = 3{,}4592$$

Beispiel: Rekonstruktion

- Invertieren der Transformation

$$\boldsymbol{y} = \boldsymbol{U}^T \boldsymbol{x} \longrightarrow \boldsymbol{x} = (\boldsymbol{U}^T)^{-1} \boldsymbol{y}$$

- Falls die Spalten der Matrix $\boldsymbol{U}$ orthonormale Vektoren sind, dann kann man folgendermaßen vereinfachen

$$\boldsymbol{U}^{-1} = \boldsymbol{U}^T \longrightarrow \boldsymbol{x} = \boldsymbol{U}\boldsymbol{y}$$

Beispiel: Rekonstruktion

Koeffizienten aus der reduzierten Basis → $\boldsymbol{U} = \begin{pmatrix} 0{,}6779 \\ 0{,}7352 \end{pmatrix}$

$y_{1,k}$
3,4592
0,8536
3,6235
2,9055
4,3071
3,5442
2,5321
1,4866
2,1932
1,4074

Rekonstruktion der Daten mit $\boldsymbol{x} = \boldsymbol{U}\boldsymbol{y}$

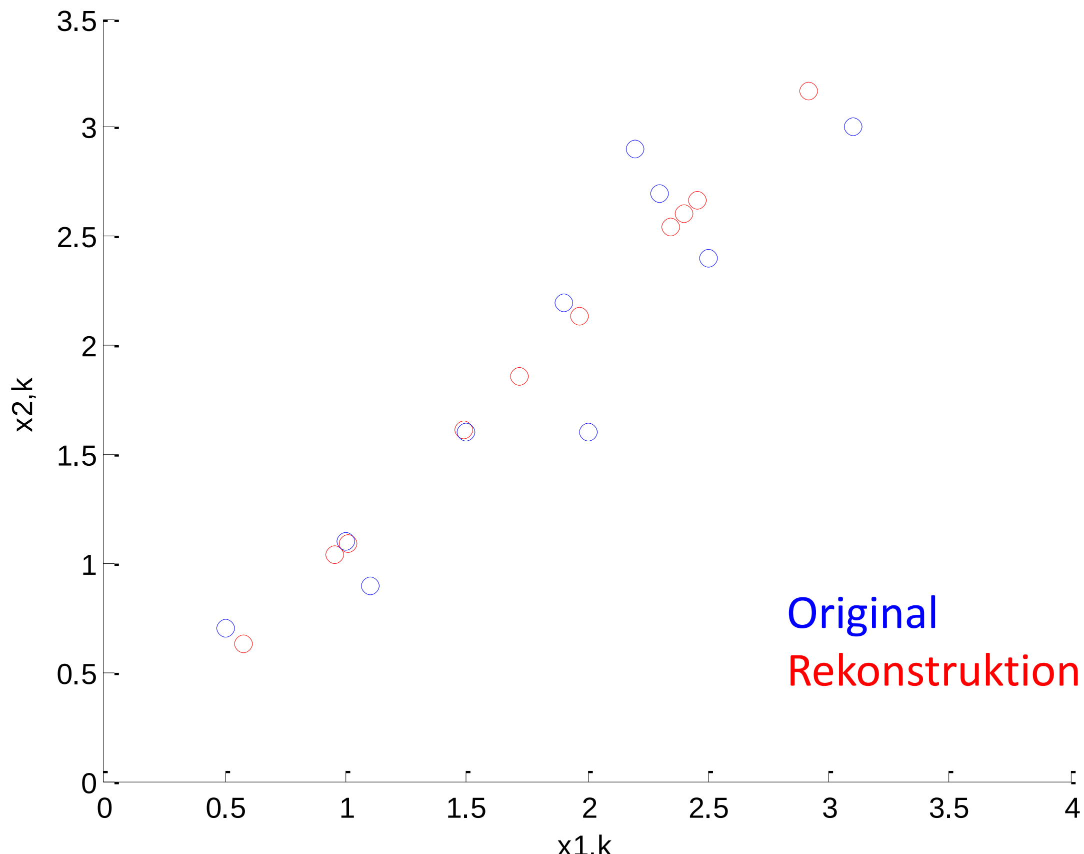
$$\boldsymbol{x}_1^{neu} = \boldsymbol{U}y_1 = \begin{pmatrix} 0{,}6779 \\ 0{,}7352 \end{pmatrix} \cdot 3{,}4592 = \begin{pmatrix} 2{,}3450 \\ 2{,}5432 \end{pmatrix}$$

zum Vergleich: $\boldsymbol{x}_1 = \begin{pmatrix} 2{,}5 \\ 2{,}4 \end{pmatrix}$

Beispiel: Ergebnis

- Ergebnis der Rekonstruktion

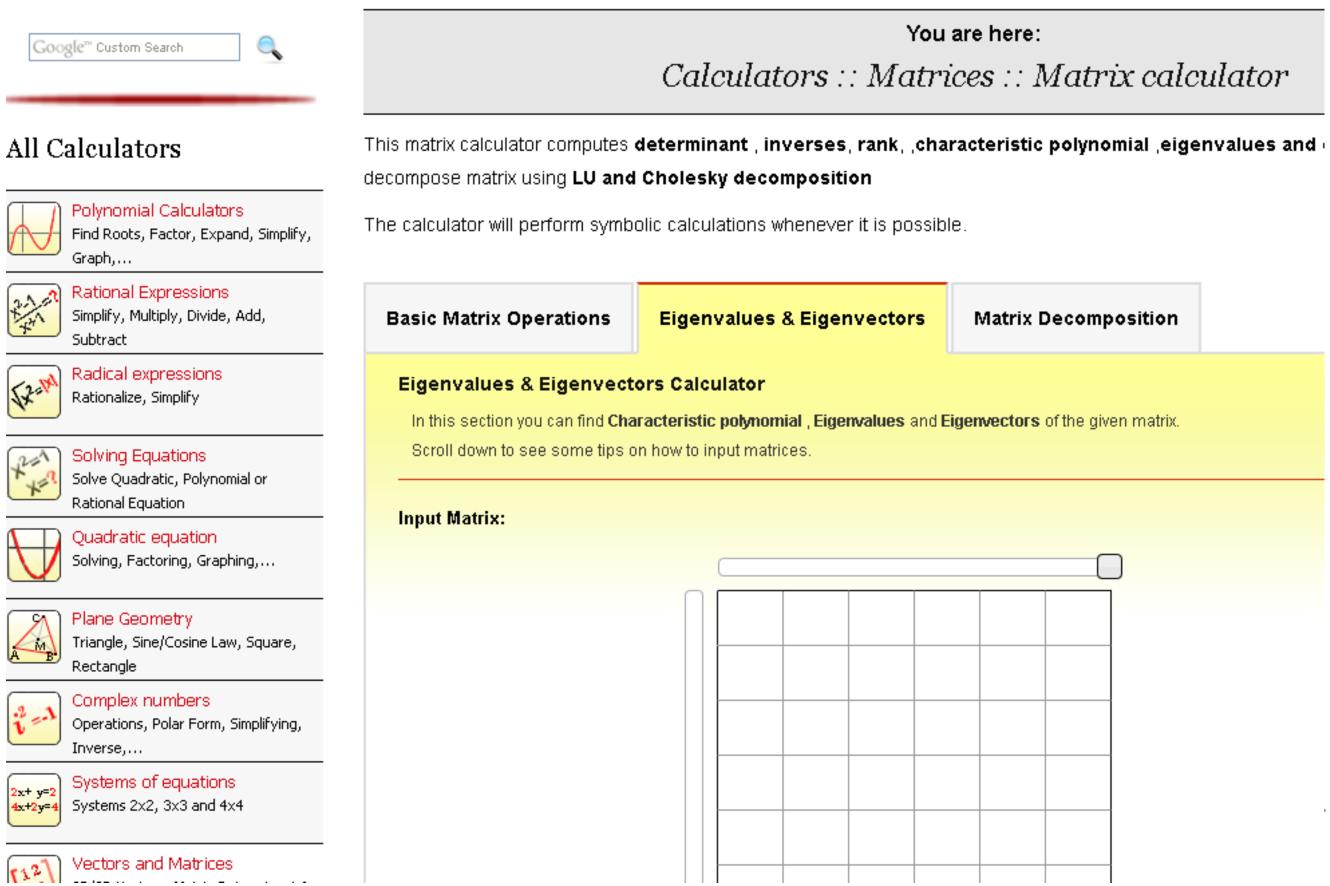
$x_{1,k}^{neu}$	$x_{2,k}^{neu}$
2.3450	2.5432
0.5786	0.6276
2.4563	2.6640
1.9696	2.1361
2.9198	3.1666
2.4026	2.6057
1.7165	1.8616
1.0078	1.0930
1.4867	1.6124
0.9541	1.0347



Online „Rechner“

… für Hauptkomponentenanalyse

http://www.mathportal.org/calculators/matrices-calculators/matrix-calculator.php



VI. Anwendungsbeispiele

Anwendung

Gesichtserkennung

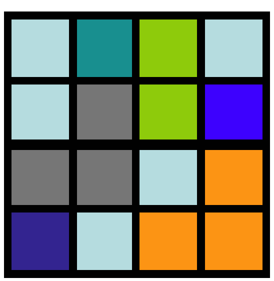
- benutzerfreundliches Biometrie-Verfahren
- selbst Menschen sind nicht 100% fehlerfrei
- Anwendungsgebiet für PCA

Gesichtserkennung …

- **Eigenfaces** → Anwendung der Hauptkomponenten-analyse (PCA)

1. Schritt:

Transformation der 2D Bilder in 1D Merkmalsvektoren



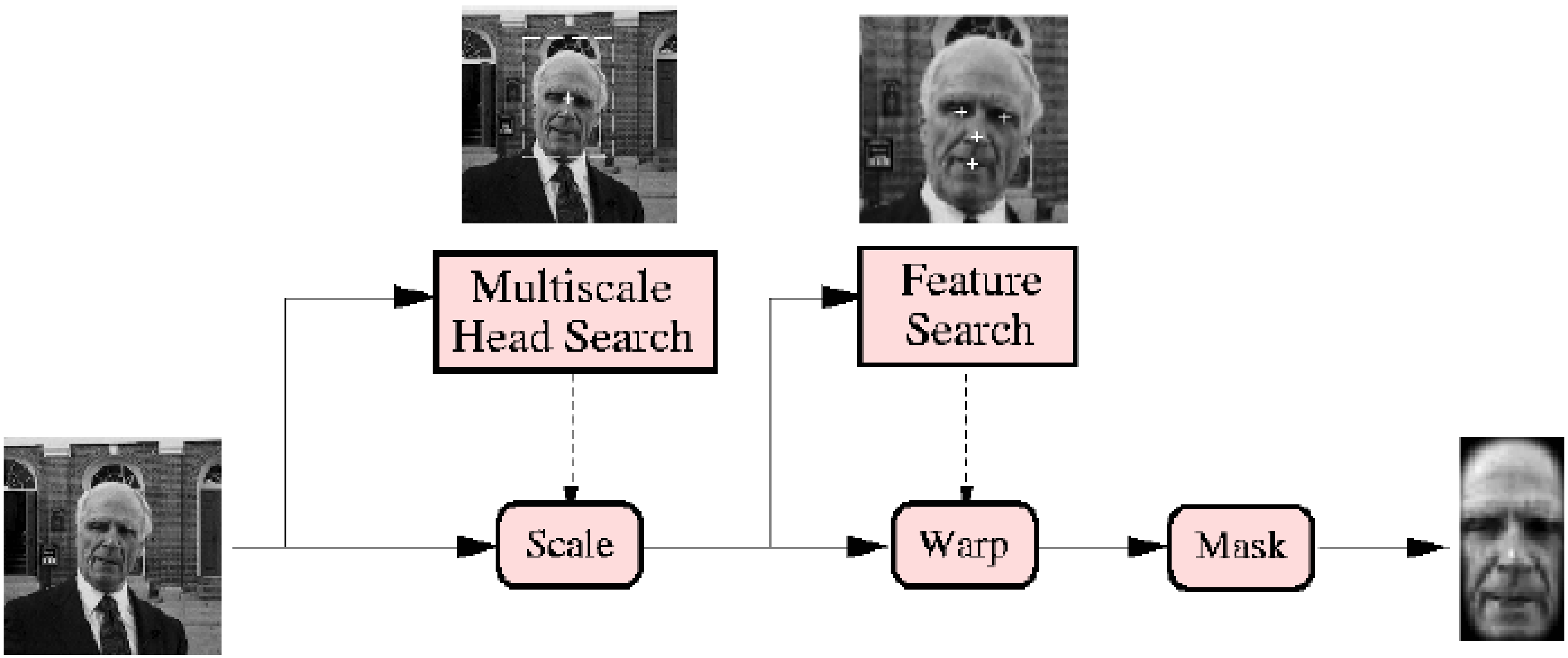
Bild

$$\mathbb{R}^{N\times N} \rightarrow \mathbb{R}^{N^2\times 1}$$

$$f(x,y) \rightarrow f(i) = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{N^2} \end{pmatrix}$$

Gesichtserkennung

- Trainingsdaten: Standardisierte Bilder von Gesichtern



[Quelle: Moghaddam & Pentland, 1995]

Gesichtserkennung

- Repräsentation der Gesichter aus den Trainingsdaten durch eine reduzierte Basis von Eigenvektoren → die **Eigenfaces**

$$\boldsymbol{x}_k = \sum_{i=1}^{m} \omega_{k,i} \boldsymbol{u}_i$$

Trainingsdaten

...

Eigenfaces

...

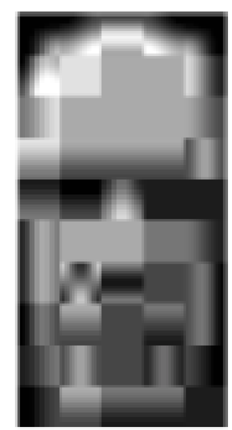
Gesichtserkennung

… zusätzlicher Aspekt → **Datenkompression**

Rekonstruktion eines Gesichts (aus dem Trainingsdaten-satz) als gewichtete Summe der Eigenvektoren

Eigenface Rekonstruktion (85 bytes)

VS.

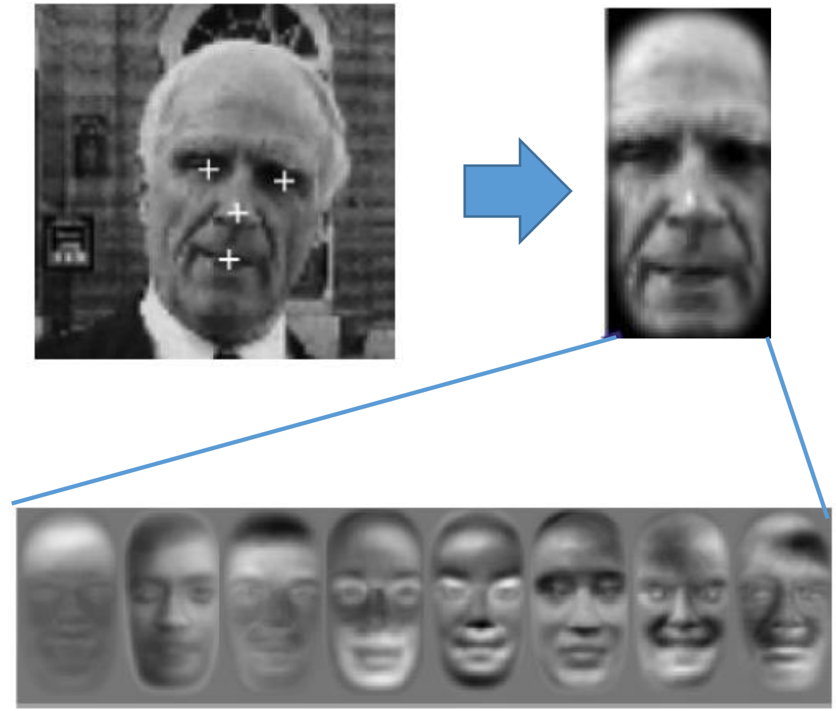


JPEG Kompression (530 bytes)

Gesichtserkennung

Ablauf der Erkennung:

1. Vorverarbeitung der Eingabe
2. Rekonstruktion aus bestehenden Eigenfaces in „Datenbank“
3. Vergleich der verwendeten Koeffizienten mit den Koeffizienten von gespeicherten Gesichtern (Datenbank)
4. Identifikation des Gesichts

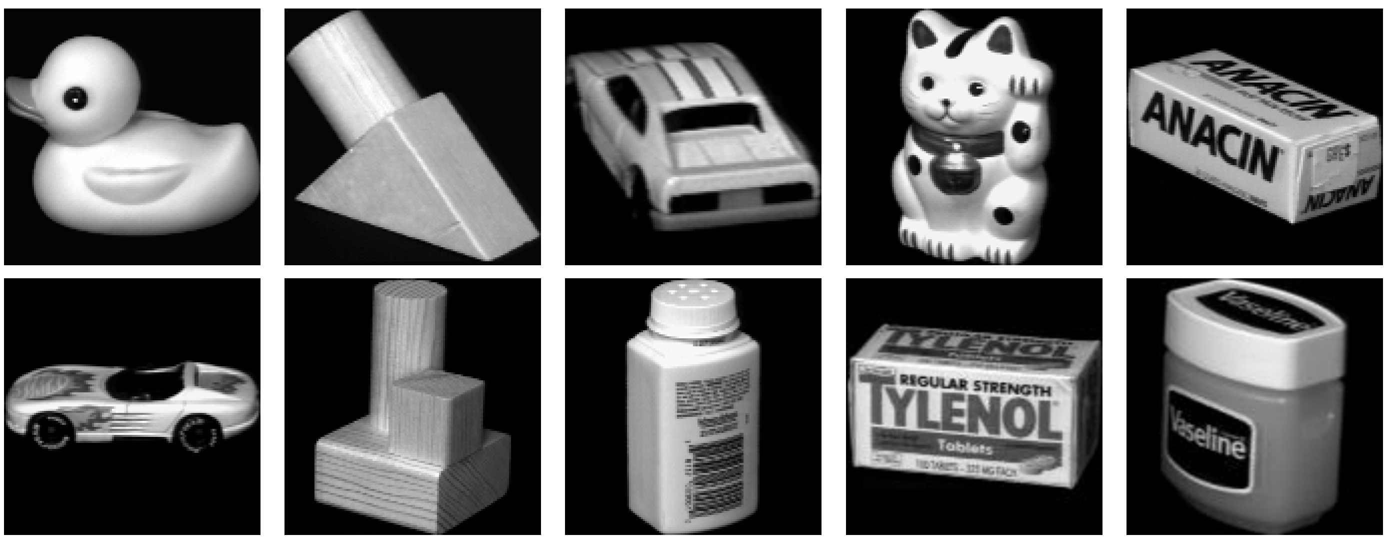


Koeffizienten aus Rekonstruktion

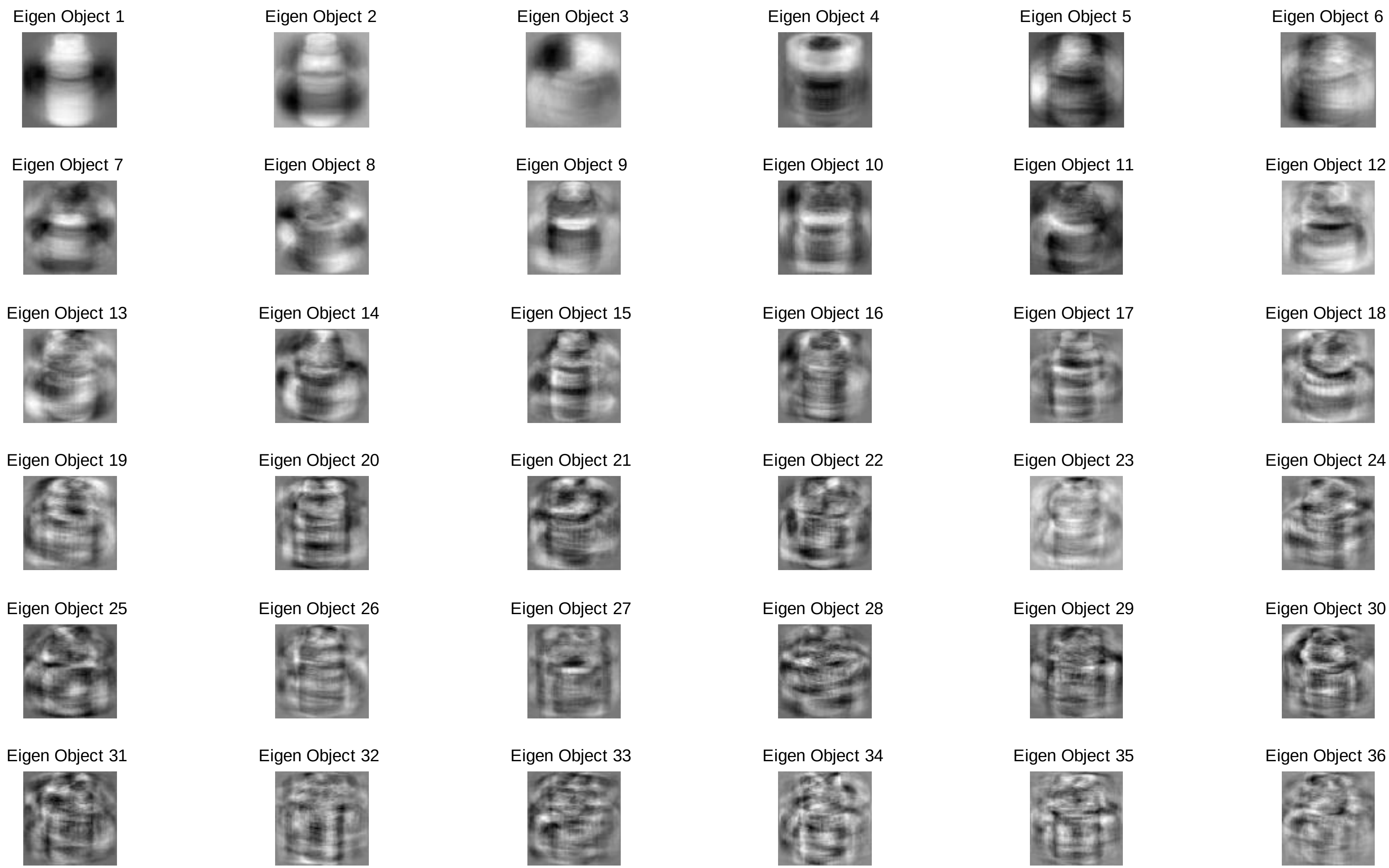
Vergleich mit Koeffizienten in einer Datenbank von bekannten Gesichtern

Anwendung

- nicht auf Gesichter beschränkt
- kann allgemein für „Objekte“ verwendet werden: Eigenobjects
- **Beispiel: 10 verschiedene Objekte**



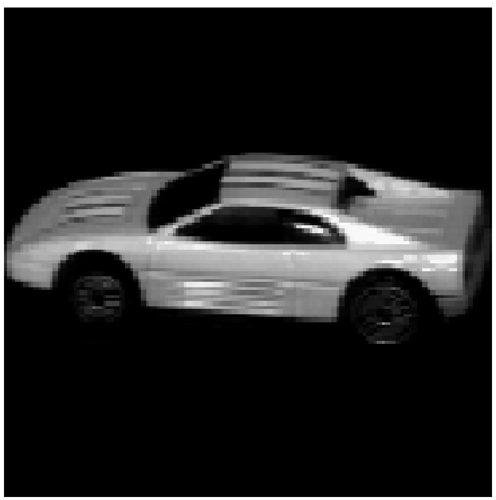
Eigenobjects



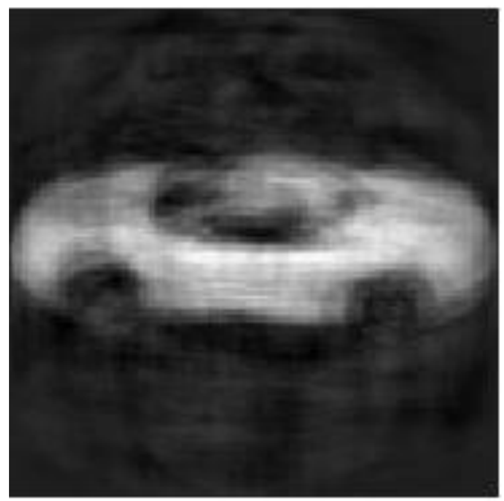
Rekonstruktion

- von einem Objekt aus dem Trainingsdatensatz

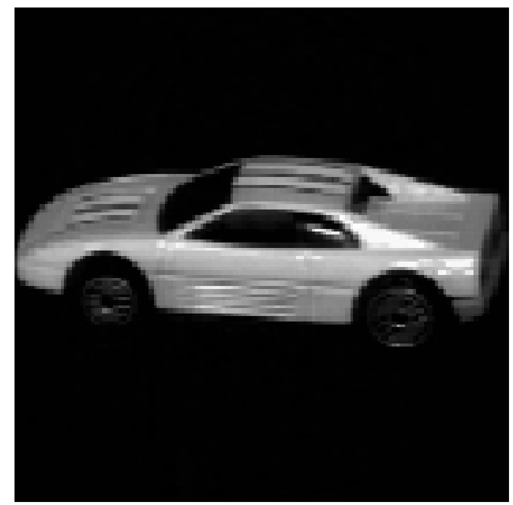
Original



Rekonstruktion mit reduzierter Basis



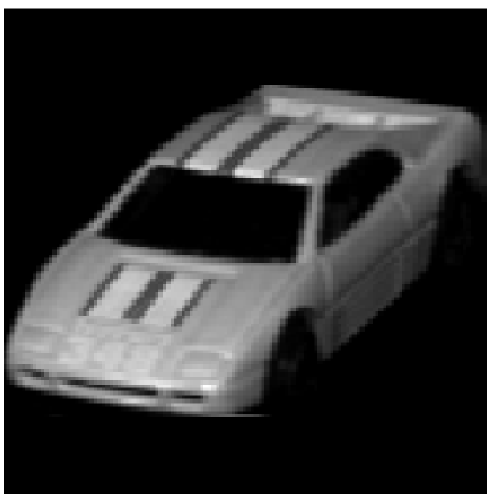
Rekonstruktion mit gesamter Basis



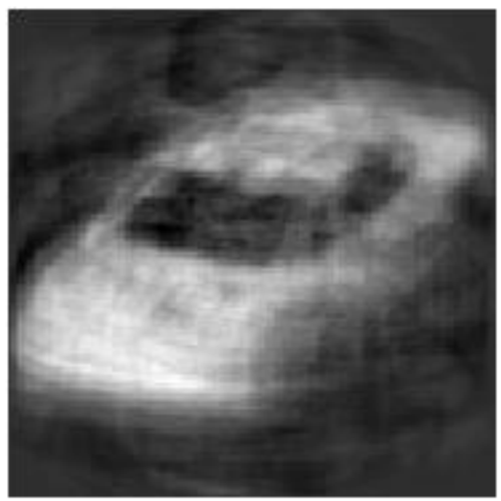
Rekonstruktion

- von einem unbekannten Objekt

Original



Rekonstruktion mit reduzierter Basis



Rekonstruktion mit gesamter Basis