4. Übungsblatt: Statistische Methoden der Datenanalyse

Aufgabe 11:

a) Mittelwerte: $\vec{\mu}_j = \frac{1}{N_j}\begin{pmatrix}\sum x_{j,1,i} \\ \vdots \\ \sum x_{j,n,i}\end{pmatrix}$

$$\Rightarrow \vec{\mu}_0 = \frac{1}{6}\left[\begin{pmatrix}1\\1\end{pmatrix} + \begin{pmatrix}2\\1\end{pmatrix} + \begin{pmatrix}1,5\\2\end{pmatrix} + \begin{pmatrix}2\\2\end{pmatrix} + \begin{pmatrix}2\\3\end{pmatrix} + \begin{pmatrix}3\\3\end{pmatrix}\right]$$

$$= \frac{1}{6}\begin{pmatrix}11,5\\11\end{pmatrix} = \begin{pmatrix}\frac{23}{12}\\2\end{pmatrix}$$

$$\vec{\mu}_1 = \frac{1}{6}\left[\begin{pmatrix}1,5\\1\end{pmatrix} + \begin{pmatrix}2,5\\1\end{pmatrix} + \begin{pmatrix}3,5\\1\end{pmatrix} + \begin{pmatrix}2,5\\2\end{pmatrix} + \begin{pmatrix}3,5\\2\end{pmatrix} + \begin{pmatrix}4,5\\2\end{pmatrix}\right]$$

$$= \frac{1}{6}\begin{pmatrix}18\\9\end{pmatrix} = \begin{pmatrix}3\\\frac{3}{2}\end{pmatrix}$$

Streumatrizen: $S_j = \sum_i^{N_j} (\vec{x}_i - \vec{\mu}_j)(\vec{x}_i - \vec{\mu}_j)^T$

$$\Rightarrow S_0 = \begin{pmatrix}-\frac{11}{12}\\-1\end{pmatrix}\begin{pmatrix}-\frac{11}{12} & -1\end{pmatrix} + \begin{pmatrix}\frac{1}{12}\\-1\end{pmatrix}\begin{pmatrix}\frac{1}{12} & -1\end{pmatrix} + \begin{pmatrix}-\frac{5}{12}\\0\end{pmatrix}\begin{pmatrix}-\frac{5}{12} & 0\end{pmatrix}$$

$$+ \begin{pmatrix}\frac{1}{12}\\0\end{pmatrix}\begin{pmatrix}\frac{1}{12} & 0\end{pmatrix} + \begin{pmatrix}\frac{1}{12}\\1\end{pmatrix}\begin{pmatrix}\frac{1}{12} & 1\end{pmatrix} + \begin{pmatrix}\frac{13}{12}\\1\end{pmatrix}\begin{pmatrix}\frac{13}{12} & 1\end{pmatrix}$$

$$= \begin{pmatrix}\frac{121}{144} & \frac{11}{12}\\ \frac{11}{12} & 1\end{pmatrix} + \begin{pmatrix}\frac{1}{144} & -\frac{1}{12}\\ -\frac{1}{12} & 1\end{pmatrix} + \begin{pmatrix}\frac{25}{144} & 0\\ 0 & 0\end{pmatrix} + \begin{pmatrix}\frac{1}{144} & 0\\ 0 & 0\end{pmatrix}$$

$$+ \begin{pmatrix}\frac{1}{144} & \frac{1}{12}\\ \frac{1}{12} & 1\end{pmatrix} + \begin{pmatrix}\frac{165}{144} & \frac{13}{12}\\ \frac{13}{12} & 1\end{pmatrix} = \begin{pmatrix}\frac{53}{24} & 2\\ 2 & 4\end{pmatrix}$$

$$S_1 = \begin{pmatrix}-1,5\\-\frac{1}{2}\end{pmatrix}\begin{pmatrix}-\frac{3}{2} & -\frac{1}{2}\end{pmatrix} + \begin{pmatrix}-\frac{1}{2}\\-\frac{1}{2}\end{pmatrix}\begin{pmatrix}-\frac{1}{2} & -\frac{1}{2}\end{pmatrix} + \begin{pmatrix}\frac{1}{2}\\-\frac{1}{2}\end{pmatrix}\begin{pmatrix}\frac{1}{2} & -\frac{1}{2}\end{pmatrix}$$

$$+ \begin{pmatrix}-\frac{1}{2}\\\frac{1}{2}\end{pmatrix}\begin{pmatrix}-\frac{1}{2} & \frac{1}{2}\end{pmatrix} + \begin{pmatrix}\frac{1}{2}\\\frac{1}{2}\end{pmatrix}\begin{pmatrix}\frac{1}{2} & \frac{1}{2}\end{pmatrix} + \begin{pmatrix}\frac{3}{2}\\\frac{1}{2}\end{pmatrix}\begin{pmatrix}\frac{3}{2} & \frac{1}{2}\end{pmatrix}$$

$$= \begin{pmatrix}\frac{9}{4} & \frac{3}{4}\\ \frac{3}{4} & \frac{1}{4}\end{pmatrix} + \begin{pmatrix}\frac{1}{4} & \frac{1}{4}\\ \frac{1}{4} & \frac{1}{4}\end{pmatrix} + 2\begin{pmatrix}\frac{1}{4} & -\frac{1}{4}\\ -\frac{1}{4} & \frac{1}{4}\end{pmatrix} + \begin{pmatrix}\frac{1}{4} & \frac{1}{4}\\ \frac{1}{4} & \frac{1}{4}\end{pmatrix} + \begin{pmatrix}\frac{9}{4} & \frac{3}{4}\\ \frac{3}{4} & \frac{1}{4}\end{pmatrix}$$

$$= \begin{pmatrix}\frac{11}{2} & \frac{3}{2}\\ \frac{3}{2} & \frac{3}{2}\end{pmatrix}$$

$$\Rightarrow S_w = S_0 + S_1 = \begin{pmatrix} \frac{185}{24} & \frac{7}{2} \\ \frac{7}{2} & \frac{11}{2} \end{pmatrix}$$

b) $\vec{\lambda}^* = S_w^{-1}(\vec{\mu}_0 - \vec{\mu}_1)$

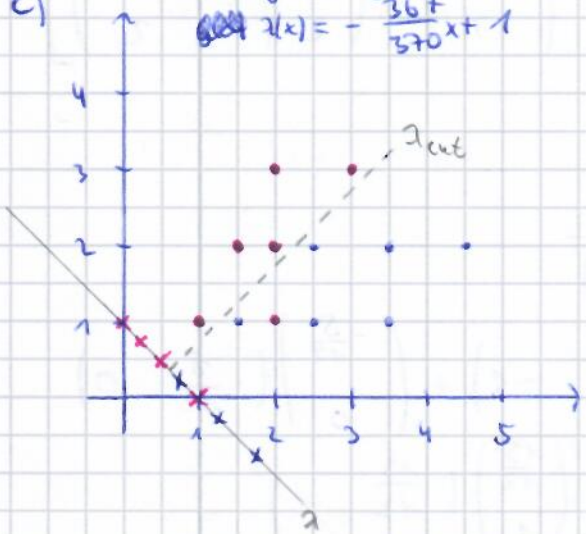
$$S_w^{-1} = \frac{1}{\frac{185}{24}\,\frac{11}{2} - \frac{49}{4}} \begin{pmatrix} \frac{11}{2} & -\frac{7}{2} \\ -\frac{7}{2} & \frac{185}{24} \end{pmatrix}$$

$$\Rightarrow \vec{\lambda}^* = \frac{48}{1447} \begin{pmatrix} \frac{11}{2} & -\frac{7}{2} \\ -\frac{7}{2} & \frac{185}{24} \end{pmatrix} \begin{pmatrix} -\frac{13}{12} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} -\frac{370}{1447} \\ \frac{367}{1447} \end{pmatrix}$$

c)

Geradengleichung:

$\lambda(x) = -\frac{367}{370}x + 1$



λ_{cut} wird so gewählt, dass möglichst wenige Werte falsch zugeordnet sind

d)

positiv

Reinheit:

Bzgl. Population 0:

$$R_0 = \frac{5}{5} = 1$$

Bzgl. Population 1:

$$R_1 = \frac{6}{7}$$

Effizienz:

Bzgl. Population 0:

$$E_0 = \frac{5}{6}$$

Bzgl. Population 1:

$$E_1 = \frac{6}{6} = 1$$