

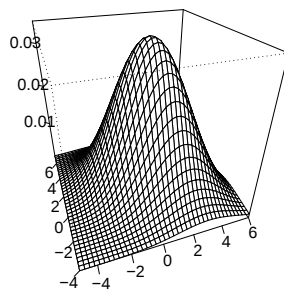
Aufgabe 4:

Gegeben sei eine Stichprobe mit drei Beobachtungen aus einer zweidimensionalen Normalverteilung mit folgender Datenmatrix:

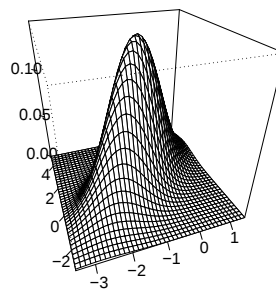
$$\mathbf{X} = \begin{pmatrix} -0.85 & 1.62 \\ -2.80 & -0.85 \\ -1.56 & 0.43 \end{pmatrix}$$

Berechnen Sie für die vorliegende Stichprobe folgende Kenngrößen:

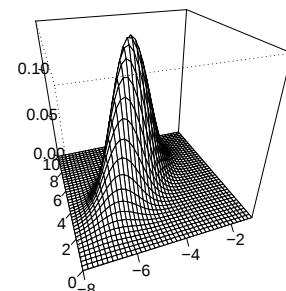
- den empirischen Mittelwertsvektor $\bar{\mathbf{x}}$
- die zentrierte Datenmatrix $\mathbf{X}^* = \mathbf{H} \cdot \mathbf{X}$ mit der Zentrierungsmatrix $\mathbf{H}$ wie in Aufgabe 3
- die empirische Kovarianzmatrix $\mathbf{S}$
- die empirische Korrelationsmatrix $\mathbf{R}$
- Die folgenden Abbildungen A, B und C zeigen Dichte und Konturplot von zweidimensional normalverteilten Zufallsvariablen – darunter auch die Abbildung der wahren Dichte, die der gegebenen Stichprobe zugrunde liegt. Welche der drei Dichten erscheint Ihnen dafür am plausibelsten?



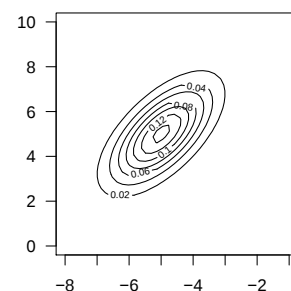
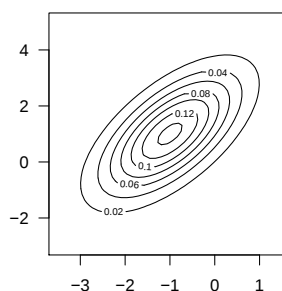
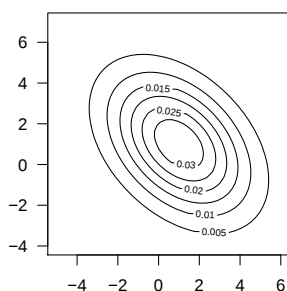
A



B



C



Aufgabe 5:

In dieser Aufgabe wird ein Datensatz mit 3103 Wettkampfergebnissen der weltbesten Zehnkämpfer aus den Jahren 1998 bis 2009 betrachtet (siehe www.decathlon2000.com). Der Datensatz enthält 3103 Zeilen mit “rohen” Resultaten der Athleten in den 10 Disziplinen. Die rohen Resultate wurden mit Hilfe der Formeln von www.iaaf.org zusätzlich in das im Zehnkampf übliche Punktesystem umgewandelt. Die folgende Tabelle gibt einen Überblick über die Variablen im Datensatz – die zu den Resultaten gehörigen Punktezahlen sind jeweils in Klammern dargestellt:

name	description [unit]	summary of raw athletes' performances (points)				
		min	q25	mean	q75	max
M100	100m race [s]	12.19(615)	11.41(772)	11.23(812)	11.04(852)	10.31(1021)
LJ	Long jump [m]	6.06(600)	6.78(762)	7(814)	7.21(864)	8.11(1089)
SP	Shot put [m]	9.04(429)	12.37(629)	13.23(682)	14.1(735)	17.78(962)
HJ	High jump [cm]	165(504)	188(697)	194(749)	200(804)	222(1012)
M400	400m race [s]	56.57(538)	51.5(747)	50.54(792)	49.52(837)	46.15(1001)
MH110	110m hurdles [s]	18.05(520)	15.51(789)	15.14(834)	14.77(878)	13.67(1018)
Disc	Discus [m]	17.78(231)	36.83(601)	39.79(661)	42.64(719)	54.08(956)
PV	Pole vault [cm]	0(0)	420(674)	446(752)	470(820)	555(1084)
Jav	Javelin [m]	19.1(150)	49.94(588)	54.06(650)	58.23(712)	77.47(1004)
M1500	1500m race [s]	355(285)	293(601)	284.15(658)	274(719)	248(896)

Auf Seite 3 sind die Ergebnisse zusätzlich grafisch dargestellt.

- (a) Beschreiben und interpretieren Sie diese Art der Datenvisualisierung.
- (b) Können Sie grobe Zusammenhänge zwischen verschiedenen Disziplinen bzw. zwischen Gruppen von Disziplinen erkennen?

