

5-6

Kai

22 Dezember 2016

Kategoriable Variablen müssen kodiert werden, damit sie nicht stetig regressiert werden. Es gibt 2 gängige Methoden. Die Dummy- und die Effekt-Kodierung, wobei die Dummy-Kodierung weiter verbreitet ist und auch in R normalerweise die Standardeinstellung.

i)

$$\mathbf{x} = \begin{pmatrix} 1 & 0 & 0 & 0 & x_{13} \\ 1 & 0 & 1 & 0 & x_{23} \\ 1 & 0 & 0 & 1 & x_{33} \\ 1 & 1 & 0 & 0 & x_{43} \\ 1 & 1 & 1 & 0 & x_{53} \\ 1 & 1 & 0 & 1 & x_{63} \end{pmatrix}$$

1. Spalte: Intercept, 2. Spalte: x_1 , 3. Spalte: $x_2(2)$, 4. Spalte: $x_2(3)$, 5. Spalte: x_3 \ In den Spalten 2 und 4, Zeile 1 steht die Referenz Kategorie.

ii)

$$\mathbf{x} = \begin{pmatrix} 1 & -1 & -1 & -1 & x_{13} \\ 1 & -1 & 1 & 0 & x_{23} \\ 1 & -1 & 0 & 1 & x_{33} \\ 1 & 1 & -1 & -1 & x_{43} \\ 1 & 1 & 1 & 0 & x_{53} \\ 1 & 1 & 0 & 1 & x_{63} \end{pmatrix}$$

1. Spalte: Intercept, 2. Spalte: x_1 , 3. Spalte: $x_2(2)$, 4. Spalte: $x_2(3)$, 5. Spalte: x_3 \ In den Spalten 2 und 4, Zeile 1 steht die Referenz Kategorie.