

Aufgabe 1

- (a) Der Beginn der Zeitrechnung wird hier willkürlich auf 12:00 Uhr festgelegt. Damit werden die Additionen einfacher und für die Bestimmung des Startzeitpunkts (b) ist keine Subtraktion nötig.

Zeitpunkt	t	korrigierte Prüfungen
12:00 Uhr	0	
14:21 Uhr	141	3
15:03 Uhr	183	5
16:07 Uhr	247	9
17:08 Uhr	308	14

Modell: $\hat{y} = \beta_0 + \beta_1 t$

$$(b) \quad X = \begin{pmatrix} 1 & 141 \\ 1 & 183 \\ 1 & 247 \\ 1 & 308 \end{pmatrix}; y = \begin{pmatrix} 3 \\ 5 \\ 9 \\ 14 \end{pmatrix} \Rightarrow \beta = \begin{pmatrix} -6.744 \\ 0.06596 \end{pmatrix}$$

$$\hat{y} = -6.744 + 0.06596t$$

$$R^2 = 0.9889 \text{ (gut)}$$

$$(c) \text{ Start: } 0 = -6.744 + 0.06596t \Rightarrow t_{\text{Start}} \approx 102 \text{ Min}$$

$$\text{Ende: } 20 = -6.744 + 0.06596t \Rightarrow t_{\text{Ende}} \approx 405 \text{ Min}$$

$$\text{Korrekturbeginn: } 12:00 \text{ Uhr} + 102 \text{ Min} = 13:42 \text{ Uhr}$$

$$\text{Korrekturende: } 12:00 \text{ Uhr} + 405 \text{ Min} = 18:45 \text{ Uhr}$$

Aufgabe 2

$$(a) \quad \hat{y}_a = \beta_0 + \beta_1 x$$

$$X_a = \begin{pmatrix} 1 & 3 \\ 1 & 1 \\ 1 & 2 \\ 1 & 4 \end{pmatrix}; y = \begin{pmatrix} 4 \\ 7 \\ 5 \\ 6 \end{pmatrix}$$

$$\text{prgmLSQM: } \hat{y}_a = 6.5 - 0.4x \quad (R_a^2 = 0.16)$$

$$(b) \quad \hat{y}_b = \beta_0 + \beta_1 x + \beta_2 x^2$$

$$X_b = \begin{pmatrix} 1 & 3 & 9 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 16 \end{pmatrix}; y = \begin{pmatrix} 4 \\ 7 \\ 5 \\ 6 \end{pmatrix}$$

$$\text{prgmLSQM: } \hat{y}_b = 11.5 - 5.4x + x^2 \quad (R_b^2 = 0.96)$$

$$(c) \hat{y}_c = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$$

$$X_c = \begin{pmatrix} 1 & 3 & 9 & 27 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 4 & 16 & 64 \end{pmatrix}; y = \begin{pmatrix} 4 \\ 7 \\ 5 \\ 6 \end{pmatrix}$$

$$\text{prgmLSQM: } \hat{y}_c = 8 + 0.1667x - 1.5x^2 + 0.3333x^3 \quad (R_c^2 = 1)$$

Beobachtung: Mit jeder zusätzlichen Variable vergrößert (verbessert) sich das Bestimmtheitsmass R^2 . Für das Modell (c) erhält man sogar eine perfekte Übereinstimmung, da es dort 4 Variablen für 4 Gleichungen gibt.

Dieses Phänomen wird *Overfitting* (Überanpassung) genannt und entsteht dann, wenn es zu viele (überflüssige) Modellparameter oder zu wenig Beispieldaten gibt.

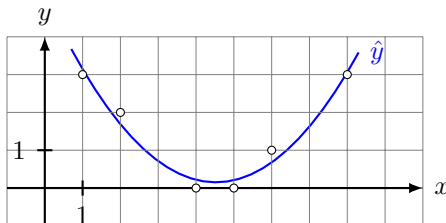
Aufgabe 3

Die Punkte liegen näherungsweise auf einer Parabel.

$$\text{quadratisches Modell: } \hat{y} = \beta_0 + \beta_1 x + \beta_2 x^2$$

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \\ 1 & 8 & 64 \end{pmatrix}; y = \begin{pmatrix} 3 \\ 2 \\ 0 \\ 0 \\ 1 \\ 3 \end{pmatrix}$$

$$\text{prgmLSQM: } \hat{y} = 0.2410x^2 - 2.181x + 5.086; R^2 = 0.9666 \text{ (gut)}$$



Aufgabe 4

- Es muss mindestens so viele Datensätze geben wie Modellparameter.
- Die Spalten in der Matrix der Beispieldaten müssen *linear unabhängig* sein; d. h. keine der Spalten darf durch eine Linearkombination der anderen Spalten darstellbar sein.

Aufgabe 5

(a) Modell: $\hat{y} = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2$ (x_1 : Leistung; x_2 : Hubraum)

$$X = \begin{pmatrix} 1 & 55 & 1360 \\ 1 & 120 & 1796 \\ \dots & \dots & \dots \\ 1 & 103 & 2435 \end{pmatrix}; y = \begin{pmatrix} 7.5 \\ 8.5 \\ \dots \\ 8.9 \end{pmatrix}$$

$$\text{prgmLSQM: } \beta = \begin{pmatrix} 4.6847 \\ 0.01919 \\ 0.009039 \end{pmatrix}; R^2 = 0.6569 \text{ (mässig)}$$

$$\hat{y} = 4.6847 + 0.01919x_1 + 0.009039x_2$$

(b) $\hat{y}(85, 1984) = 4.6847 + 0.01919 \cdot 85 + 0.009039 \cdot 1983 \approx 8.1 \ell/100 \text{ km}$

Aufgabe 6

Modell: $\hat{y} = \beta_0 + \beta_1 \sqrt{x}$

$$A = \begin{pmatrix} 1 & \sqrt{1} \\ 1 & \sqrt{2} \\ 1 & \sqrt{5} \\ 1 & \sqrt{10} \end{pmatrix}; y = \begin{pmatrix} 3 \\ 4 \\ 5 \\ 6 \end{pmatrix}$$

prgmLSQM: $\hat{y} = 1.896 + 1.333\sqrt{x}$; $R^2 = 0.9744$ (gut)

