

Themenvorschläge für die kleinen Übungen am 13. April 2010

- a) Zeigen Sie, daß das folgende lineare Gleichungssystem unlösbar ist:

$$x + y = 1, \quad x + 2y = 2 \quad \text{und} \quad 2x + 3y = 4 \quad (*)$$

Lösung: Klar, denn die Summe der ersten beiden Gleichungen ist gleich der dritten mit durch drei ersetzter rechter Seite.

- b) Finden Sie reelle Zahlen x, y , so daß $(*)$ mit diesen Zahlen im Sinne der kleinsten Quadrate möglichst wenig falsch ist!

Lösung: Die Matrix des Gleichungssystems ist

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 3 \end{pmatrix} \quad \text{mit} \quad A^T = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix}.$$

Wir müssen das lineare Gleichungssystem $(A^T A)x = A^T b$ lösen, wobei b die rechte Seite des gegebenen Gleichungssystems ist. Wegen

$$A^T A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 6 & 9 \\ 9 & 14 \end{pmatrix} \quad \text{und} \quad {}^t A \vec{b} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 11 \\ 17 \end{pmatrix}$$

ist dies das Gleichungssystem

$$6x + 9y = 11 \quad \text{und} \quad 9x + 14y = 17.$$

Subtraktion von $1\frac{1}{2}$ mal der ersten Gleichung von der zweiten führt auf $\frac{1}{2}y = \frac{1}{2}$, also $y = 1$, und damit folgt aus jeder der beiden Gleichungen schnell, daß $x = \frac{1}{3}$ sein muß.

- c) Gegeben seien N Datenpaare (x_i, y_i) , die ungefähr proportional zueinander sein sollten. Finden sie das im Sinne der Methode der kleinsten Quadrate bestmögliche $a \in \mathbb{R}$, so daß $y_i \approx ax_i$ ist!

Lösung: Schreiben wir die Bedingungen $y_i = ax_i$ um zu einem Gleichungssystem für a , hat dieses als Matrix den Spaltenvektor mit Einträgen x_i und als rechte Seite den mit Einträgen y_i . Multiplikation mit dem Zeilenvektor (x_1, \dots, x_n) macht daraus die Gleichung

$$\left(\sum_{i=1}^N x_i^2 \right) a = \sum_{i=1}^N x_i y_i;$$

falls nicht alle x_i verschwinden, ist dies eindeutig lösbar durch

$$a = \frac{\sum_{i=1}^N x_i y_i}{\sum_{i=1}^N x_i^2}.$$

- d) Wie können Sie vorgehen, wenn ein Zusammenhang der Form $x_i = A \cos(t_i + \varphi)$ zu erwarten ist?

Lösung: Das Problem hier ist, daß die Gleichungen nicht linear in φ sind. Nach der Additionsformel für den Kosinus ist aber

$$\cos(t_i + \varphi) = \cos t_i \cos \varphi - \sin t_i \sin \varphi,$$

d.h. $x_i = a \cos t_i + b \sin t_i$ mit $a = A \cos \varphi$ und $b = A \sin \varphi$.

Dieses Gleichungssystem ist linear in a und b , kann also nach der Methode aus der Vorlesung gelöst werden: Multiplikation mit der Transponierten der Matrix des Gleichungssystems führt auf

$$\left(\sum_{i=1}^N \cos^2 t_i \right) a + \left(\sum_{i=1}^N \sin t_i \cos t_i \right) b = \sum_{i=1}^N x_i \cos t_i$$

und

$$\left(\sum_{i=1}^N \sin t_i \cos t_i \right) a + \left(\sum_{i=1}^N \sin^2 t_i \right) b = \sum_{i=1}^N x_i \sin t_i.$$

Dieses Gleichungssystem liefert a und b ; daraus läßt sich A berechnen als

$$A = \sqrt{a^2 + b^2}$$

und φ durch die beiden Bedingungen

$$\cos \varphi = \frac{a}{A} \quad \text{und} \quad \sin \varphi = \frac{b}{A}.$$

(Man benötigt beide Bedingungen um φ modulo 2π zu kennen; eine allein liefert nur φ modulo π .)

- e) Die Datenpaare (t_i, x_i) , $i = 1, \dots, N$ beschreiben den zeitlichen Verlauf einer zu verschiedenen Zeitpunkten t_i gemessenen Größe. Es steht zu erwarten, daß diese Größe einerseits periodischen (z.B. saisonalen) Schwankungen unterworfen ist, andererseits aber im langfristigen Mittel linear ansteigt. Damit bietet sich ein Ansatz der Form

$$x_i \approx a + bt_i + c \sin \omega t_i$$

an, wobei ω eine von der Periode abhängige bekannte Konstante bezeichnet. Stellen Sie ein lineares Gleichungssystem für die bestmöglichen Schätzwerte a, b und c auf!

Lösung: Wenn wir lauter exakte Gleichungen hätten, würden a, b, c das lineare Gleichungssystem

$$1 \cdot a + t_i \cdot b + \sin \omega t_i \cdot c = x_i, \quad i = 1, \dots, N$$

erfüllen. Seine Matrix und rechte Seite sind

$$A = \begin{pmatrix} 1 & t_1 & \sin \omega t_1 \\ \vdots & \vdots & \vdots \\ 1 & t_N & \sin \omega t_N \end{pmatrix} \quad \text{und} \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}.$$

Multiplikation mit der transponierten Matrix

$$A^T = \begin{pmatrix} 1 & \dots & 1 \\ t_1 & \dots & t_N \\ \sin \omega t_1 & \dots & \sin \omega t_N \end{pmatrix}$$

führt auf

$$A^T A = \begin{pmatrix} N & \sum_{i=1}^N t_i & \sum_{i=1}^N \sin \omega t_i \\ \sum_{i=1}^N t_i & \sum_{i=1}^N t_i^2 & \sum_{i=1}^N t_i \sin \omega t_i \\ \sum_{i=1}^N \sin \omega t_i & \sum_{i=1}^N t_i \sin \omega t_i & \sum_{i=1}^N \sin^2 \omega t_i \end{pmatrix} \quad \text{und} \quad A^T x = \begin{pmatrix} \sum_{i=1}^N x_i \\ \sum_{i=1}^N t_i x_i \\ \sum_{i=1}^N x_i \sin \omega t_i \end{pmatrix}.$$

Das zu lösende Gleichungssystem ist also

$$\begin{aligned} N \cdot a + \sum_{i=1}^N t_i \cdot b + \sum_{i=1}^N \sin \omega t_i \cdot c &= \sum_{i=1}^N x_i \\ \sum_{i=1}^N t_i \cdot a + \sum_{i=1}^N t_i^2 \cdot b + \sum_{i=1}^N t_i \sin \omega t_i \cdot c &= \sum_{i=1}^N t_i x_i \\ \sum_{i=1}^N \sin \omega t_i \cdot a + \sum_{i=1}^N t_i \sin \omega t_i \cdot b + \sum_{i=1}^N \sin^2 \omega t_i \cdot c &= \sum_{i=1}^N x_i \sin \omega t_i \end{aligned}$$

- f) Beim vorigen Problem war der periodische Anteil zum Zeitpunkt $t = 0$ stets gleich Null; maximal war er unter anderem zum Zeitpunkt $\pi/2\omega$. Wie können Sie vorgehen, wenn Sie nicht wissen, zu welchem Zeitpunkt Sie verschwindende, minimale und maximale Beiträge erwarten sollten?

Lösung: Die offensichtliche Lösung besteht natürlich darin, im Sinus eine Phasenverschiebung einzuführen:

$$x_i \approx a + bt_i + c \sin(\omega t_i + \varphi)$$

mit einem ebenfalls zu schätzenden Parameter φ . Da die Gleichungen nicht linear in φ sind, läßt sich das allerdings nicht mit den uns bekannten Methoden durchführen. Wie wir bei AAufgabe d) gesehen haben, erfüllt ein Ansatz der Form

$$x_i \approx a + bt_i + c \sin \omega t_i + d \cos \omega t_i$$

denselben Zweck, und hier können wir mit nur unwesentlich größerer Mühe als beim gerade gelösten Problem ein lineares Gleichungssystem für die vier Koeffizienten a, b, c, d aufstellen.

- g) Gegeben seien N Wertepaare (x_i, t_i) , wobei theoretisch ein Zusammenhang der Form

$$x_i = a \sin t_i + b \sin 2t_i + c \sin 3t_i + d \sin 4t_i$$

bestehen sollte. Stellen Sie ein lineares Gleichungssystem auf, dessen Lösungen im Sinne der kleinsten Quadrate die beste Schätzung für a, b, c, d liefern!

Lösung: Die gesuchten Größen a, b, c, d sollten theoretisch das lineare Gleichungssystem aus den N Gleichungen

$$(\sin t_i) \cdot a + (\sin 2t_i) \cdot b + (\sin 3t_i) \cdot c + (\sin 4t_i) \cdot d = x_i$$

erfüllen. Dessen Matrix A hat vier Spalten, wobei die Einträge der j -ten Spalte gleich den Zahlen $\sin jt_i$ sind. Damit ist $A^T A$ eine 4×4 -Matrix mit Einträgen

$$a_{k\ell} = \sum_{i=1}^N \sin kt_i \cdot \sin \ell t_i.$$

Als rechte Seite des Gleichungssystem haben wir den Vektor

$${}^t A \vec{x} = \begin{pmatrix} \sum_{i=1}^N x_i \sin t_i \\ \vdots \\ \sum_{i=1}^N x_i \sin 4t_i \end{pmatrix},$$

das Gleichungssystem besteht also aus den vier Gleichungen

$$\begin{aligned} \left(\sum_{i=1}^N \sin kt_i \cdot \sin t_i \right) a + \left(\sum_{i=1}^N \sin kt_i \cdot \sin 2t_i \right) b + \left(\sum_{i=1}^N \sin kt_i \cdot \sin 3t_i \right) c \\ + \left(\sum_{i=1}^N \sin kt_i \cdot \sin 4t_i \right) d = \sum_{i=1}^N x_i \sin kt_i \end{aligned}$$

für $k = 1, \dots, 4$.

h) Berechnen Sie den Korrelationskoeffizienten der Datenpaare $(\sin^2 k, \cos^2 k)$, $k = 1, \dots, 100$!

Lösung: Da $\cos^2 k = 1 - \sin^2 k$ ist für alle k , liegen die Datenpaare auf einer Geraden mit Steigung $-1 < 0$; der Korrelationskoeffizient ist also -1 .

i) Zeigen Sie, daß die offenen Intervalle $(0, \frac{1}{n})$ für $n = 1, 2, 3, \dots$ eine offene Überdeckung des Intervalls $[\frac{1}{4}, \frac{1}{3}]$ bilden und finden Sie eine endliche Teilüberdeckung!

Lösung: Das erste der offenen Intervalle ist $(0, 1)$ und enthält bereits sowohl $[\frac{1}{4}, \frac{1}{3}]$ als auch alle folgenden offenen Intervalle. Damit ist bereits $\{(0, 1)\}$ eine endliche Teilüberdeckung. (Das Intervall $(0, \frac{1}{2})$ für sich alleine würde ebenfalls ausreichen, nicht aber $(0, \frac{1}{3})$, denn dieses Intervall enthält nicht den Punkt $\frac{1}{3}$.)

j) Die Menge \mathcal{U} bestehe aus den offenen Mengen

$$U_{a,b} = \{(x, y) \in \mathbb{R}^2 \mid (x - a)^2 + (y - b)^2 < 1\} \quad \text{mit} \quad -1 \leq a, b \leq 1.$$

Zeigen Sie, daß dies eine offene Überdeckung des Rechtecks

$$R = \{(x, y) \in \mathbb{R}^2 \mid -\frac{3}{2} \leq x, y \leq \frac{3}{2}\}$$

ist und finden Sie eine endliche Teilüberdeckung!

Lösung: Für $(x, y) \in R$ setzen wir

$$a = \begin{cases} -1 & \text{falls } x < -1 \\ x & \text{falls } -1 \leq x \leq 1 \\ 1 & \text{falls } x > 1 \end{cases} \quad \text{und} \quad b = \begin{cases} -1 & \text{falls } y < -1 \\ y & \text{falls } -1 \leq y \leq 1 \\ 1 & \text{falls } y > 1 \end{cases}.$$

Da $-\frac{3}{2} \leq x, y \leq \frac{3}{2}$ ist, kann $(x - a)^2 + (y - b)^2$ nicht größer werden als $(\frac{1}{2})^2 + (\frac{1}{2})^2 = \frac{1}{2}$, der Punkt (x, y) liegt also in $U_{a,b}$. Somit ist \mathcal{U} eine Überdeckung von R .

Als endliche Teilüberdeckung können wir beispielsweise die Mengen $U_{a,b}$ mit $a, b \in \{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\}$ nehmen: Wenn wir bei der obigen Definition von a und b zusätzlich noch beide Zahlen zum nächsten Vielfachen von $\frac{1}{2}$ runden, machen wir einen zusätzlichen Fehler von höchstens $\frac{1}{4}$. Bei einer Koordinate aus $[-1, 1]$ machen wir damit insgesamt einen Fehler von höchstens $\frac{1}{4}$, bei einer Koordinate außerhalb dieses Intervalls kann der Fehler bis zu $\frac{1}{2}$ betragen. Insgesamt gilt weiterhin dieselbe Abschätzung wie oben.

k) Zeigen Sie, daß die offene Kreisscheibe

$$D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 2\}$$

nicht kompakt ist!

Lösung: Wir können beispielsweise die Überdeckung aus den offenen Kreisringen

$$U_n = \{(x, y) \in \mathbb{R}^2 \mid 2 - \frac{1}{n} < x^2 + y^2 < 2 - \frac{1}{2n}\}$$

zusammen mit der Kreisscheibe

$$U_0 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < \frac{3}{2}\}$$

betrachten. Eine endliche Auswahl von Kreisringen U_{n_i} mit $n_1 < \dots < n_r$ enthält nur Punkte mit $x^2 + y^2 < 2 - \frac{1}{2n_r}$, kann also weder mit noch ohne U_0 eine Überdeckung von D sein.