Albert-Ludwigs-Universität Freiburg Wintersemester 2009

Statistische Methoden der Datenanalyse

Markus Schumacher

Übung VIII

Matthew Beckingham und Henrik Nilsen

17.12.2009

Anwesenheitsaufgaben

Aufgabe 32 *Geradenanpassung mit Matrixmethoden*

Betrachten Sie eine Stichprobe vom Umfang N von Messwerten $(x_1,y_1),(x_2,y_2),\ldots,(x_N,y_N)$, die alle denselben Fehler σ auf die Messung von y_i haben sollen. Benutzen Sie die Matrixnotation der Methode der kleinsten Quadrate, um Schätzer $\hat{\theta}_0$ und $\hat{\theta}_1$ auf die Parameter einer anzupassenden Geraden von der Form

$$\lambda = \theta_0 + \theta_1 x = \sum_{j=0}^{n} a_j(x_i)\theta_j = \sum_{j=0}^{n} A_{ij}\theta_j \tag{1}$$

mit $A_{ij} = a_j(x_i)$ anzugeben.

(i) Schreiben Sie zunächst die Matrix A und die Kovarianzmatrix V auf.

(ii) Berechnen Sie die Matrix für die Schätzer $\hat{\theta}$, indem Sie die Matrizen $A^T V^{-1} A$ und $A^T V^{-1} y$ berechnen.

(iii) Schreiben Sie die Schätzer $\hat{\theta}_0$ und $\hat{\theta}_1$ in Abhängigkeit der Erwartungswerte und Varianzen x, y und xy sowie der Kovarianz von x und y auf.

(iv) Wie unterscheidet sich die Rechnung, wenn jeder Messwert einen nicht korrelierten Fehler σ_i auf y_i hat? Schreiben Sie die Matrizen für V, $A^T V^{-1} A$, $A^T V^{-1} y$ und somit $\hat{\theta}$ auf.

Aufgabe 33 *Kombination zweier Messungen eines Winkels*

Betrachten Sie zwei unabhängige Messungen $s = \sin\theta$ and $c = \cos\theta$ eines Winkels θ. Die Messungen s und c haben gleiche Fehler, σ, und wurden über Mittelung einer großen Anzahl von Messungen erhalten. Wie lautet der "Maximum-Likelihood"-Schätzer für θ und seine Varianz $V[\hat{\theta}]$?

Hausaufgaben

Aufgabe 34 *Unsicherheiten in der Abschätzung von Parametern* **6 Punkte**

Die Lösungen für die Schätzer einer Anpassung auf der Basis der Methode der kleinsten Quadrate $\hat{\vec{\theta}}$ können in Matrixform geschrieben werden als:

$$\hat{\vec{\theta}} = (A^T V^{-1} A)^{-1} A^T V^{-1} \vec{y}. \tag{2}$$

Zeigen Sie durch Anwenden der allgemeinen Formel für Fehlerfortpflanzung, dass die Kovarianzmatrix geschrieben werden kann als

$$V(\hat{\vec{\theta}}) = (A^T V^{-1} A)^{-1} \tag{3}$$

mit $V = V(\vec{y})$.

Aufgabe 35 *Messung einer Asymmetrie* **7 Punkte**

In der Physik werden neue Erkenntnisse oft aus dem Vergleich von Zählraten gewonnen, die durch die Umkehrung bestimmter experimenteller Randbedingungen von einander unterschieden werden. Eine Option ist die Zählraten bezüglich einer vorher ausgezeichneten Raumrichtung zu bestimmen, z.B. in den man die Anzahl von Ereignissen oberhalb und unterhalb einer Ebene vergleicht n_o und n_u. Die gewählte Raumrichtung als Normalenvektor der Ebene definiert dabei „oben “und „unten “.

Die beobachtete Asymmetrie ergibt sich aus den beobachteten Zählraten zu:

$$\alpha = \frac{n_o - n_u}{n_o - n_u}. \tag{4}$$

Die Zählraten sind dabei poissonverteilte Zufallsvariablen mit Mittelwerten ν_L bzw. ν_R.

(i) Benutzen Sie die Tatsache, dass die beobachtete Anzahl der einzelnen Zählraten poissonverteilt ist, um einen „Maximum Likelihood“-Schätzer $\hat{\alpha}$ für die Asymmetrie α aufzustellen.

(ii) Benutzen Sie Fehlerfortpflanzung (und nehmen Sie dabei an, dass die Messungen der Streuereignisse für ober- bzw. unterhalb unabhängig voneinander sind), um die Standardabweichung $\sigma_{\hat{\alpha}}$ auf diesen Schätzer als Funktion von α und $n_{ges.} = n_o + n_u$ zu ermitteln.

(iii) Bevor das Experiment durchgeführt wird, wird aus der Theorie erwartet, dass die Asymmetrie ungefähr 10^{-4} beträgt. Welche Gesamtzahl von Streuereignissen musste man dafür beobachten, um die Asymmetrie mit einer relativen Genauigkeit von 10% zu messen?

Aufgabe 36 *Geradenanpassung* **7 Punkte**

Nehmen Sie an, dass Sie N Messwerte $(x_1,y_1),(x_2,y_2),\ldots,(x_N,y_N)$ mit jeweils verschiedenen, aber bekannten Fehlern σ_i auf den Wert von y_i vorliegen haben. Es soll mittels der Methode der kleinsten Quadrate diejenige Gerade ermittelt werden, die am besten durch diese Punkte geht.

Betrachten Sie die folgende Parametrisierung einer Geraden für die Datenpunkte (x_i,y_i):

$$f_i(\theta_1; \theta_2; x_i) = \theta_1 + \theta_2 x_i, \tag{5}$$

wobei θ_1 der Achsenabschnitt und θ_2 die Steigung der Geraden sind.

(i) Betrachten Sie die kleinsten-Quadrat-Schätzer $\hat{\theta}_1$ und $\hat{\theta}_2$ und ihre Varianzen mit Gewichten $w_i = 1/\sigma_i^2$, indem Sie die Größe

$$X^2 = \sum_{i=1}^{N} w_i (y_i - f_i)^2 = \sum_{i=1}^{N} (\frac{y_i - f_i}{\sigma_i})^2 \tag{6}$$

minimieren.

(ii) Benutzen Sie die Fehlerfortpflanzungsformel, um die Varianzen der beiden Schätzer $\sigma^2_{\hat{\theta}_1}$ und $\sigma^2_{\hat{\theta}_2}$ ermitteln.