

# Statistische Methoden der Datenanalyse

Markus Schumacher

## Übung X

Matthew Beckingham und Henrik Nilsen

14.01.2010

### Anwesenheitsaufgaben

#### Aufgabe 38 Methode der Kleinsten Quadrate mit Zwangsbedingungen: Winkelmessung im Dreieck

In der Vorlesung wurde gezeigt, dass, um einen Satz von Messungen  $\vec{y} = (y_1, \dots, y_N)$  durch eine lineare Funktion  $A\vec{\theta}$  zu fiten, ein Fit mit der Methode der Linearen Kleinsten Quadrate durchgeführt werden kann. Dieser minimiert die Grösse

$$\chi^2 = (\vec{y} - A\vec{\theta})^T V^{-1} (\vec{y} - A\vec{\theta}) \quad (1)$$

$V$  steht für die Kovarianzmatrix der Messungen  $\vec{y}$ .

Weiterhin wurde gezeigt, dass, unter Berücksichtigung eines Satzes von  $K$  Randbedingungen  $\vec{b} = (b_1, \dots, b_K)$ , welche die Gleichungen  $B\vec{\theta} - \vec{b} = 0$  erfüllen, die Methode der kleinsten Quadrate durch Minimierung der Grösse

$$\chi^2 = (\vec{y} - A\vec{\theta})^T V^{-1} (\vec{y} - A\vec{\theta}) + 2\vec{\lambda}^T (B\vec{\theta} - \vec{b}) \quad (2)$$

verbessert werden kann, wobei  $\vec{\lambda} = (\lambda_1, \dots, \lambda_K)$  ein Vektor von Langrange-Multiplikatoren ist. Folglich muss  $\chi^2$  minimiert werden in Bezug auf  $\vec{\theta}$  und  $\vec{\lambda}$ . Die Lösung dieses Minimierungsproblems ergibt die Schätzer

$$\hat{\vec{\theta}} = C^{-1}\vec{c} - C^{-1}B^T V_B^{-1} (BC^{-1}\vec{c} - \vec{b}) = C^{-1}\vec{c} - C^{-1}B^T \hat{\vec{\lambda}} \quad (3)$$

wobei  $C \equiv A^T V^{-1} A$  und  $\vec{c} \equiv A^T V^{-1} \vec{y}$ . Die Varianz errechnet sich daher zu

$$V[\hat{\vec{\theta}}] = C^{-1} - (BC^{-1})^T V_B^{-1} BC^{-1} \quad (4)$$

Die verbesserten Schätzer für die Messungen sind gegeben durch

$$\hat{\vec{n}} = A\hat{\vec{\theta}} = A \left[ C^{-1}\vec{c} - C^{-1}B^T V_B^{-1} (BC^{-1}\vec{c} - \vec{b}) \right] \quad (5)$$

mit der Varianz

$$V[\hat{\vec{n}}] = AV[\hat{\vec{\theta}}]A^T = A \left[ C^{-1} - (BC^{-1})^T V_B^{-1} BC^{-1} \right] A^T \quad (6)$$

- (i) Zeigen Sie, dass  $E(\hat{\vec{\lambda}}) = 0$  und  $E(\hat{\vec{\theta}}) = \vec{\theta}$ . Daraus folgt, dass die Schätzungen der Parameter erwartungstreu sind.
- (ii) Betrachten Sie die Messung von drei Winkeln eines Dreiecks analog zu einem Beispiel aus der Vorlesung. Die drei Messungen sind gegeben durch  $\vec{y} = (y_1, y_2, y_3)$  mit  $\sigma_i = \sigma$  und sollen gefittet werden durch die Funktion  $A\vec{\theta}$  mit  $\vec{\theta} = (\theta_1, \theta_2, \theta_3)$  und  $A = \mathbf{1}_3$ .
  - a) Was sind die Werte für  $B$  und  $\vec{b}$ ?
  - b) Berechnen Sie  $\vec{\theta}$  und  $V[\hat{\vec{\theta}}]$  und somit  $\hat{\vec{n}}$  und  $V[\hat{\vec{n}}]$ .
  - c) Wie verbessert die Zwangsbedingung auf die Messungen die Schätzungen der gemessenen Werte der Dreieckswinkel?

# Hausaufgaben

## Aufgabe 39 Elektron-Positron Kollisionen

3 Punkte

Die Anzahl von Ereignissen in Elektron-Positron Kollisionen mit bestimmten kinematischen Eigenschaften kann als poissonverteilte Variable angesehen werden. Nehmen Sie an, dass bei einer bestimmten integrierten Luminosität (dies wäre ein Datensatz für eine gegebene Strahlintensität) 3.9 Ereignisse aus bekannten Prozessen erwartet werden, aber 16 beobachtet.

Berechnen Sie den P-Wert für die Hypothese, dass kein neuer Prozess zu der Zahl von Ereignissen beiträgt. Um Poisson-Wahrscheinlichkeiten zu summieren kann die Relation

$$\sum_{n=0}^m P(n; \nu) = 1 - F_{\chi^2}(2\nu; n_{\text{dof}}) \quad (7)$$

verwendet werden, wobei  $P(n; \nu)$  die Poisson-Wahrscheinlichkeit für  $n$  (Ereignisse) bei einem Mittelwert von  $\nu$  bezeichnet, sowie  $F_{\chi^2}$  die kumulative  $\chi^2$ -Verteilung mit  $n_{\text{dof}} = 2(m + 1)$  Freiheitsgraden (diese Werte können Sie tabelliert in Standardwerken vorfinden).

## Aufgabe 40 Teilchenidentifikation

4 Punkte

Durchquert ein geladenes Teilchen ein Gasvolumen, erzeugt dieses in dem Medium Ionisation. Die mittlere Menge hängt dabei von der Masse und Impuls des Teilchens ab. Daher können durch Ausarbeitung einer Testhypothese, basierend auf der erfassten Ionisation im Gasvolumen bei bekannten Impuls verschiedene, Teilchen identifiziert werden.

Betrachten Sie einen Strahl von Teilchen, welcher entweder Pionen oder Elektronen enthält. So kann man, als eine Funktion der Ionisation  $t$ , die WDF der Hypothese  $g(t|e)$ , dass das Teilchen ein Elektron ist, und der Hypothese  $g(t|\pi)$ , dass das Teilchen ein Pion ist, aufstellen. Hierzu wählt man eine Menge von Elektronen aus, indem gefordert wird, dass  $t \leq t_{\text{cut}}$ :

Nehmen Sie an, dass beide Hypothesen durch eine um  $t = 0$  zentrierte Gaussverteilung für die Elektronen und eine um  $t = 2$  zentrierte Gaussverteilung für die Background-Pionen konstruiert sind. Beide Gausssschen Hypothesen haben eine Standardabweichung von eins. Ein Test zur Elektronenselektion wird konstruiert, indem eine Ionisation  $t \leq 1$  gefordert wird.

- (i) Was ist der Signifikanzlevel des Tests (d.h. die Wahrscheinlichkeit Elektronen zu akzeptieren)?
- (ii) Wie groß ist die Mächtigkeit des Tests gegen die Hypothese, dass das Teilchen ein Pion ist? Wie groß ist die Wahrscheinlichkeit, dass ein Pion als Elektron akzeptiert wird?
- (iii) Betrachten Sie eine Verhältnis von 99% Pionen and 1% Elektronen im Strahl. Wie groß ist die Reinheit der durch  $t < 1$  selektierten Auswahl?

## Aufgabe 41 Zerfallszeit eines Teilchens unter Verwendung des Likelihood-Quotienten

5 Punkte

In der Vorlesung wurde das Beispiel der Messung der Lebensdauer eines instabilen Teilchens vorgestellt. In einem Experiment wird ein Satz von Zerfallszeiten  $\vec{t} = (t_1, \dots, t_n)$  eines Teilchens aufgenommen. Diese Daten sollen dazu verwendet werden, den Test einer Nullhypothese  $H_0 : \tau = 1$  gegen die zusammengesetzte Alternativhypothese  $H_1 : \tau \neq 1$  zu vergleichen.

- (i) Was sind sowohl Likelihood-Funktion, als auch Schätzer der Lebensdauer  $\hat{\tau}$  für  $n$  Messungen eines exponentiellen Zerfalls eines Teilchens der Lebensdauer  $\tau$ ?
- (ii) Was ist somit die Likelihood-Funktion  $L(\vec{x}, \hat{\lambda}(\Omega))$  des gesamten Parameterraumes  $\Omega$  (gegeben durch  $H_0 + H_1$ )? Was ist die Likelihood-Funktion  $L(\vec{x}, \hat{\lambda}(\omega))$  des Raumes  $\omega$ , das heißt der Hypothese  $H_0$ ?
- (iii) Zeigen Sie, dass der Likelihood-Quotient gegeben ist durch

$$T = \bar{t}^n \exp(-n(\bar{t} - 1)) \quad (8)$$

mit  $\bar{t} = \sum t/n$ .

- (iv) Für große  $n$  ist  $\bar{t}$  gemäß einer gaußschen WDF  $N(1, 1/n)$  verteilt. Zeigen Sie, dass für eine Signifikanz von  $\alpha = 0.05$  die Grenze des Signifikanzbereiches  $\lambda_{0.05}$  gegeben ist durch

$$\lambda_{0.05} = \left(1 + \frac{1.645}{\sqrt{n}}\right)^n \exp(-1.645\sqrt{n}) \quad (9)$$

**Aufgabe 42** *Kombination von Messungen mit der Methode der kleinsten Quadrate*

**8 Punkte**

Es ist möglich, einen Spezialfall der Methode der kleinsten Quadrate zu benutzen, um Messungen mit der selben Qualität zu kombinieren. Betrachten Sie  $N$  Messungen,  $y_i$ , welche den wahren, aber unbekanntem Wert  $\lambda$  bestimmen sollen. Jede Messung  $y_i$  hat einen geschätzten Fehler von  $\sigma_i$ . Da  $\lambda$  für alle Ereignisse konstant ist, ergibt sich die  $\chi^2$ -Variable zu:

$$\chi^2(\lambda) = \sum_{i=1}^N \frac{(y_i - \lambda)^2}{\sigma_i^2} \quad (10)$$

- (i) Wenn jedoch die Messungen von  $y_i$  nicht unabhängig sind, sondern eine Korrelation, gegeben durch die Kovarianzmatrix  $V$ , besitzen, ergibt sich:

$$\chi^2(\lambda) = \sum_{i,j=1}^N (y_i - \lambda)(V^{-1})_{ij}(y_j - \lambda) \quad (11)$$

Zeigen Sie, dass in diesem Fall der Schätzer der Methode der kleinsten Quadrate für  $\lambda$  gegeben ist durch

$$\hat{\lambda} = \sum_{i=1}^N w_i y_i \quad (12)$$

wobei die Gewichtungen  $w_i$  gegeben sind durch

$$w_i = \frac{\sum_{j=1}^N (V^{-1})_{ij}}{\sum_{k,l=1}^N (V^{-1})_{kl}}, \quad (13)$$

und dass die Varianz gegeben ist durch

$$V[\hat{\lambda}] = \sum_{i,j=1}^N w_i V_{ij} w_j. \quad (14)$$

- (ii) Betrachten Sie jetzt zwei Messungen,  $y_1$  und  $y_2$ , mit einer zugehörigen Kovarianzmatrix

$$V = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \quad (15)$$

mit dem Korrelationskoeffizienten  $\rho = V_{12}/(\sigma_1\sigma_2)$ . Zeigen Sie durch Berechnung des Inversen der Kovarianzmatrix  $V^{-1}$ , dass der Schätzer für  $\lambda$  gegeben ist durch

$$\hat{\lambda} = w y_1 + (1 - w) y_2 \quad (16)$$

mit

$$w = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}. \quad (17)$$

- (iii) Zeigen Sie, dass die Varianz von  $\hat{\lambda}$  gegeben ist durch

$$\frac{1}{V[\hat{\lambda}]} = \frac{1}{1 - \rho^2} \left[ \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} - \frac{2\rho}{\sigma_1\sigma_2} \right] = \frac{1}{\sigma^2} \quad (18)$$

und zeigen Sie dass folglich

$$\frac{1}{\sigma^2} - \frac{1}{\sigma_1^2} \geq 0 \quad (19)$$

was bedeutet, dass die Kombination von zwei Messungen immer zu einer Verbesserung der Varianz des Schätzers führt.