INSTITUT FÜR GEOPHYSIK UND METEOROLOGIE WS 2008/09
UNIVERSITÄT ZU KÖLN

H. Elbern mit A. Strunk 10. November 2008

Übung zur Vorlesung

Objektive Analyse

Lösungen zu Übungsblatt 1

Lösung zu Aufgabe 1:
Sei A = Grenzwertüberschreitung und B = Hagelschlag, dann ergeben sich folgende Informationen:

$$\begin{aligned} P(A) &= 0.09 & \Longleftrightarrow && P(\bar{A}) &= 0.91 \\ P(A|B) &= 0.94 & \Longleftrightarrow && P(\bar{A}|B) &= 0.06 \\ P(\bar{B}|A) &= 0.04 & \Longleftrightarrow && P(B|A) &= 0.96 \end{aligned}$$

Mit der Regel von Bayes

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})}$$

ergibt sich

$$\begin{aligned} P(B|\bar{A}) &= \frac{P(B|A)P(A)}{P(A|B)P(\bar{A})} - \frac{P(B|A)P(A)}{P(\bar{A})} \\ &= \frac{0.96 \cdot 0.09}{0.94 \cdot 0.91} - \frac{0.96 \cdot 0.09}{0.91} \approx 0.006 = 0.6\% \end{aligned}$$

Lösung zu Aufgabe 2:
N unabhängige Messungen $Y = \{y_1, ..., y_N\}$ und eine Hintergrundinformation x_b, die hier der Einfachheit halber zusammengefasst werden sollen zu $N + 1$ Beobachtungen $Y = \{y_1, ..., y_N, x_b\}$. Aus der Unabhängigkeit der Beobachtungen und des Hintergrundwissens folgt, dass ihre Fehler unkorreliert sind ($< . >$ bezeichne den Erwartungswert einer Größe):

$$< \epsilon_i \epsilon_j >= 0, \quad \forall i \neq j$$

Zusätzlich sei davon ausgegangen, dass alle Größen unbiased (BIAS-frei) sind:

$$< \epsilon_i >= 0, \quad \forall i$$

Damit gilt für den *best linear unbiased estimate* (BLUE) $x_a = \sum_i^{N+1} w_i y_i$. Gleichung

(1.11) der Vorlesung ergibt sich damit zu

$$\begin{aligned}
< x_a > &= < \sum_i^{N+1} w_i y_i > \\
&= < \sum_i^{N+1} (x_t + \epsilon_i) w_i > \\
&= < \sum_i^{N+1} (x_t w_i + \epsilon_i w_i) > \\
&= < \sum_i^{N+1} x_t w_i > + \quad \underbrace{< \sum_i^{N+1} \epsilon_i w_i >}_{\text{=0, weil Größen BIAS-frei}} .
\end{aligned}$$

Soll $< x_a > = < x_t >$ sein, so folgt sofort

$$\sum_i w_i = 1$$

Die Bestimmung der Gewichte w_i ist nun auf zwei Arten möglich:

1. mit dem *minimum-variance* Schätzer und
2. mit dem *maximum-likelihood* Schätzer

zu 1. (minimum-variance) Gleichung (1.13) der Vorlesung lautet nun

$$\begin{aligned}
V_a = \sigma_a^2 &= < (x_a - < x_a >)^2 > \\
&= < (\sum_i^{N+1} w_i y_i - x_t)^2 > \\
&= < (\sum_i^{N+1} w_i y_i - (\sum_i^{N+1} w_i) x_t)^2 > \\
&= < (\sum_i^{N+1} w_i (y_i - x_t))^2 > \\
&= < (\sum_i^{N+1} w_i \epsilon_i)^2 > \\
&= < \sum_i^{N+1} \sum_{j=i}^{N+1} w_i w_j \epsilon_i \epsilon_j > \\
&= \sum_i^{N+1} w_i^2 < \epsilon_i^2 > \\
&= \sum_i^{N+1} w_i^2 \sigma_i^2 ,
\end{aligned}$$

wobei die Unabhängigkeit der Fehler verwendet wurde (die Mischterme $\epsilon_i \epsilon_j$ sind weggefallen).

Nun hat man eine Optimierungsaufgabe $\sigma_a^2 \stackrel{!}{=} \min$ unter der Nebenbedingung $\sum_i w_i = 1$. Dies lässt sich - wie in der Vorlesung gezeigt - mit Hilfe der Lagrange-Multikplikatoren erzielen:
Man minimiere die skalare Funktion J mit

$$J(W) = \sigma_a^2 + \lambda(1 - \sum_i w_i) = \sum_i w_i \sigma_i^2 + \lambda(1 - \sum_i w_i) \stackrel{!}{=} \min,$$

wobei $W = \{w_1, w_2, ..., w_N, w_{N+1}\}$ sei. Man setzt nun die erste Ableitung nach w_i null und erhält:

$$w_i = \frac{\lambda}{2\sigma_i^2}$$

Mit $\sum w_i = 1$ folgt $\lambda = \frac{2}{\sum \sigma_i^{-2}}$. Für die Gewichte erhalten wir damit abschließend:

$$w_i = \frac{\sigma_i^{-2}}{\sum_i^{N+1} \sigma_i^{-2}}$$

Der BLUE x_a und die Varianz σ_a^2 ergeben sich zu

$$x_a = \sum_i^{N+1} \frac{\sigma_i^{-2}}{\sum_i^{N+1} \sigma_i^{-2}} \quad \text{und} \quad \sigma_a^2 = \frac{1}{\sum_i^{N+1} \sigma_i^{-2}}.$$

zu 2. (maximum-likelihood) Die a-posteriori-Wahrscheinlichkeit unter Berücksichtigung aller Informationen lautet

$$\begin{aligned} p(x, Y) &= \prod_{i=1}^{N+1} p(x, y_i) \\ &= \prod_{i=1}^{N+1} \left[\frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(y_i - x)^2}{2\sigma_i^2} \right) \right] \\ &= \prod_{i=1}^{N+1} \left[\frac{1}{\sigma_i \sqrt{2\pi}} \right] \exp\left(-\sum_{i=1}^{N+1} \frac{(y_i - x)^2}{2\sigma_i^2} \right) \end{aligned}$$

Nun finde $x = x_a$, das diese Wahrscheinlichkeit maximiert (*maximum likelihood*)! $p(x, Y)$ wird maximal, wenn das Argument der Exponentialfunktion minimal wird.[1]. Daher differenzieren wir

$$-\ln p(x, Y) = \sum_{i=1}^{N+1} \frac{(y_i - x)^2}{2\sigma_i^2} + \text{const}$$

nach x und setzen null: $\frac{-\partial \ln p(x,Y)}{\partial x} \stackrel{!}{=} 0$. Wir erhalten dann

$$x = x_a = \sum_{i=1}^{N+1} \frac{\sigma_i^{-2}}{\sum_i \sigma_i^{-2}} y_i$$

und somit die gleiche Schätzung wie die der minimalen Varianz (für Normalverteilungen).

[1] Da der Logarithmus eine monotone Funktion ist und das Folgende daher immer gültig ist, findet man für $\ln(p)$ auch häufig die Bezeichnung *log-likelihood*-Schätzer.