Übung zur Vorlesung

Maschinelles Lernen 1

Wintersemester 2008/2009

Abteilung Maschinelles Lernen
Institut für Softwaretechnik und theoretische Informatik
Fakultät IV, Technische Universität Berlin
Prof. Dr. Klaus-Robert Müller
Email: krm@cs.tu-berlin.de

Blatt 3

Abgabe bis Mittwoch, 5. November 2008, 14:00 Uhr
im Sekretariat FR6052, oder bei Mikio Braun, FR6058 (notfalls unter der Tür durchschieben).

Aufgaben

1. **Maximum-Likelihood-Schätzer für die Exponentialverteilung (7 Punkte)**
   Wir betrachten Zufallsvariablen, welche die Exponentialverteilung besitzen:
   $$p(x;\theta) = \begin{cases} \theta e^{-\theta x} & x \geq 0 \\ 0 & \text{sonst.} \end{cases}$$
   Zeige, dass die Maximum-Likelihood-Schätzung von θ gegeben n unabhängiger Samples $X_1, \ldots, X_n$
   $$\hat{\theta} = \frac{1}{\frac{1}{n}\sum_{k=1}^{n} X_k}.$$
   lautet.

2. **Maximum-Likelihood-Schätzer für die Gleichverteilung auf einem Intervall (7 Punkte)**
   Die Gleichverteilung auf $[0, \theta]$ hat die Dichte
   $$p(x;\theta) = \begin{cases} 1/\theta & 0 \leq x \leq \theta \\ 0 & \text{sonst.} \end{cases}$$
   Zeige, dass die Maximum-Likelihood-Schätzung für θ gegeben n unabhängiger Samples $X_1, \ldots, X_n$
   $$\hat{\theta} = \max\{X_1, \ldots, X_n\}.$$
   lautet.

3. **Maximum-Likelihood-Schätzer für die Binomialverteilung (7 Punkte)**
   Die Bionmialverteilung $B(\pi, n)$ hat auf $\{0, \ldots, n\}$ die Gewichte
   $$p(k;\pi,n) = \binom{n}{k} \pi^k (1-\pi)^{n-k}.$$
   Zeige, dass der Maximum-Likelihood-Schätzer von p für eine Stichprobe $X_1, \ldots, X_n$
   $$\hat{\pi} = \frac{1}{n}\sum_{i=1}^{n} X_i$$
   ist.

4. **Diskriminanzfunktion für normalverteilte Klassen mit beliebigen Kovarianzmatrizen (9 Punkte)**
   Wir betrachten den Fall zweier normalverteilter Klassen mit allgemeinen Kovarianzmatrizen Σ_1, Σ_2 und Mittelwerten μ_1, μ_2. Zeige, dass die Diskriminanzfunktionen
   $$g_i(x) = x^t W_i x + w_i^t x + b_i$$
   lauten, mit
   $$W_i = -\frac{1}{2}\Sigma_i^{-1}, \qquad w_i = \Sigma_i^{-1}\mu_i, \qquad b_i = -\frac{1}{2}\mu_i^t \Sigma_i^{-1}\mu_i - \frac{1}{2}\ln|\Sigma_i| + \ln P(Y=i).$$
   ($|\Sigma_i|$ ist hierbei die Determinante von Σ_i, und $P(Y=i)$ die a-priori-Wahrscheinlichkeit der Klasse i)