



Blatt 7

Aufgabe 7.1

(3 Punkte)

Wir betrachten Ternärsequenzen der Länge N über dem Alphabet $\{A, B, C\}$ und die dazugehörigen Likelihoodfunktionen eines homogenen Markovmodells nullter Ordnung. Wie lautet für eine gegebene Ternärsequenz mit k_1 As, k_2 Bs und $N - k_1 - k_2$ Cs die Likelihood als Funktion von $\theta_1 = P(A)$ und $\theta_2 = P(B)$. Betrachten Sie nun die Parametertransformation $f : \underline{\theta} \rightarrow \underline{\phi}$ mit $\phi_a = \ln(\theta_a/\theta_3)$ für $a \in \{1, 2\}$ und $\theta_3 = 1 - \theta_1 - \theta_2$. Ist f invertierbar? Wie lautet die inverse Parametertransformation $f^{-1} : \underline{\phi} \rightarrow \underline{\theta}$? Wie lauten der Definitions- bzw. Wertebereich von f ? Wie lauten die Likelihood und Loglikelihood als Funktion von $\underline{\phi}$? Plotten Sie die Loglikelihood für $k_1 = 3$, $k_2 = 6$, $N = 10$ als Funktion von θ_1 und θ_2 sowie als Funktion von ϕ_1 und ϕ_2 . Leiten Sie den Maximum-Likelihood-Schätzer für $\underline{\phi}$ her und tragen Sie den Schätzwert in die entsprechenden Plots ein. Wie groß ist die maximale Likelihood? Vergleichen Sie die Schätzwerte als auch die maximale Likelihood mit den entsprechenden Werten für $\underline{\theta}$.

Aufgabe 7.2

(2 Punkte)

Wiederholen Sie Aufgabe 7.1, mit Ausnahme der Plots und der Berechnung der maximalen Likelihoodwerte, für Sequenzen der Länge N über einem Alphabet mit $A > 3$ Symbolen.

Aufgabe 7.3

(5 Punkte)

Analog zu Aufgabe 6.2 betrachten wir wieder Binärsequenzen der Länge $N = 2$ über dem Alphabet $\{Z, W\}$ und ein durch $\theta = P(Z)$ parametrisiertes homogenes Markovmodell nullter Ordnung. Gegeben sei die A-Priori-Dichte $P_1(\theta) = 1$. Plotten Sie die A-Priori-Dichte sowie für jede der vier Binärsequenzen (ZZ , ZW , WZ , WW) die A-Posteriori-Dichte als Funktionen von θ . Betrachten Sie nun die fünf Parametertransformationen

(a) $\phi = \theta^2$,

(b) $\phi = \sqrt{\theta}$,

(c) $\phi = \ln \theta$,

(d) $\phi = \ln(-\ln \theta)$,

(e) $\phi = \ln \frac{\theta}{1-\theta}$,

und berechnen und plotten Sie für jede der fünf Parametrisierungen die A-Priori-Dichte sowie für jede der vier Binärsequenzen (ZZ , ZW , WZ , WW) die A-Posteriori-Dichte als Funktion von ϕ . Geben Sie in allen fünf Fällen den Definitionsbereich der Dichten, also den Wertebereich von ϕ , an. Leiten Sie die fünf Maximum-A-Posteriori-Schätzer für ϕ her und tragen Sie jeden der sechs Schätzwerte in jeden der sechs Plots ein. Wie groß sind die (vier mal sechs mal sechs) Maximalwerte der A-Posteriori-Dichten? Sind sie identisch? Falls nicht, welche Parametrisierung würden Sie empfehlen? Wäre Ihre Empfehlung dieselbe für jeden der vier Datensätze (ZZ , ZW , WZ , WW)?

Abgabetermin: 19. Dezember
