



Blatt 12

Aufgabe 12.1 Der Datensatz `sigma70.txt` auf der Homepage zur Vorlesung besteht aus 238 Sigma-70 Bindungsstellen der Länge 12.

- Modellieren Sie eine Sigma-70 Bindungsstelle als $iMM(0)$ und bestimmen Sie die ML-Schätzwerte der Parameter $p_{il} := P(X_l = a_i)$. Diese p_{il} bilden eine Gewichtsmatrix (PWM), wobei die l -te Spalte die Verteilung von X_l beschreibt.
- Bestimmen Sie die zu dieser Gewichtsmatrix (und zu diesem Datensatz) gehörende Konsensussequenz.
- Wie häufig taucht die Konsensussequenz im gegebenen Datensatz auf?
- Welche Sequenz taucht am häufigsten im gegebenen Datensatz auf, und wie häufig?

Aufgabe 12.2 Zeigen Sie, dass ein Produkt von Dirichletdichten konjugiert zur Likelihood eines homogenen Markovmodells nullter und erster Ordnung ist. Wie transformieren sich die Hyperparameter? Zeigen Sie dies ebenfalls für inhomogenes Markovmodell nullter und erster Ordnung. Wie transformieren sich hier die Hyperparameter? Leiten Sie für alle vier Fälle den ML, MAP und MP Schätzer der Modellparameter her.

Zusatzaufgaben:

Aufgabe 12.3 Geben den Viterbi-Algorithmus im Pseudocode an.

Aufgabe 12.4 Füllen Sie den Fragebogen zu Ablauf, Inhalt, Nutzen, etc. von Vorlesung und Übung, sowie mit einer Einschätzung von Dozent und Tutorin aus. Sie finden den Fragebogen im Stud.IP auf der Homepage zur Vorlesung oder Übung.

Hinweis:

Die Lösung dieser "Aufgabe" ist weder prüfungsrelevant noch hilfreich für das Verständnis des Stoffes der Vorlesung. . . **Aber**, wir würden uns trotzdem freuen, wenn wir auf diesem Wege Feedback zur Veranstaltung bekämen und erfahren würden, wie Sie über "Algorithmen der Bioinformatik II" denken und was wir in Zukunft noch verbessern können.

Vielen Dank!