



Blatt 9

Aufgabe 9.1 Zeigen Sie:

- (a) wenn \vec{p} ein Zufallsvektor und P eine stochastische Matrix ist, so ist auch $\vec{q} = \vec{p}P$ ein Zufallsvektor.
- (b) wenn P und Q stochastische Matrizen sind, so sind auch PQ und QP stochastische Matrizen.

Bemerkungen:

- (a) \vec{x} ist ein D -dimensionaler Zufallsvektor, gdw. $x_d \geq 0 \forall d = 1, \dots, D$ und $\sum_{d=1}^D x_d = 1$
- (b) X ist eine stochastische $D \times D$ Matrix, gdw. $x_{ij} \geq 0 \forall i, j = 1, \dots, D$ und $\sum_{j=1}^D x_{ij} = 1 \forall i = 1, \dots, D$

Aufgabe 9.2 Zeigen Sie:

- (a) wenn eine stochastische $D \times D$ Matrix X nur identische Zeilen $\vec{x} = (x_{i1}, \dots, x_{iD})$, $i = 1, \dots, D$ besitzt, so ist die stationäre Verteilung dieser Matrix gleich \vec{x} .
- (b) jede Anfangsverteilung \vec{y} mit positiven Elementen $y_i > 0 \forall i = 1, \dots, D$ und $\sum_{d=1}^D y_d = 1$ konvergiert durch einmaliges Anwenden von X zur stationären Verteilung \vec{x} .

Aufgabe 9.3 Der Datensatz "sigma70.txt" auf der Homepage zur Vorlesung besteht aus 238 Sigma-70 Bindungsstellen der Länge 12.

- a) Modellieren Sie eine Sigma-70 Bindungsstellen 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.
- b) Bestimmen Sie die zu dieser Gewichtsmatrix (und zu diesem Datensatz) gehörende Konsensussequenz.
- c) Wie häufig taucht die Konsensussequenz im gegebenen Datensatz auf?
- d) Welche Sequenz taucht am häufigsten im gegebenen Datensatz auf, und wie häufig?