

3. Übung zur Vorlesung “Bioinformatische Methoden in der Genomforschung”

Sebastian Böcker, Martin Hoffmann

Ausgabe: 04.11.2019

Abgabe: 11.11.2019

Aufgabe 1 (2 Punkte)

Was sind die Aufgaben der Microarray Datenanalyse? Welche Arbeitsschritte werden häufig dafür gemacht?

Aufgabe 2 (2 Punkte)

Häufig werden bei der statistischen Analyse von Expressionsdaten Signifikanzen (p-values) berechnet. Vervollständigen Sie den folgenden Satz: “Ein p-value (Signifikanz) von 0.01 beim Test auf differentielle Genexpression bedeutet...”.

Aufgabe 3 (6 Punkte)

Gegeben sei die Matrix der Genexpressionswerte gemessen für vier Gene an vier (aufeinanderfolgenden) Tagen:

	Tag 1	Tag 2	Tag 3	Tag 4
Gen 1	0.564	-0.038	-0.561	-1.315
Gen 2	0.606	0.621	-0.83	-1.681
Gen 3	-0.555	-0.224	0.673	0.78
Gen 4	0.238	-0.764	-1.371	-1.868

1. Berechnen Sie die Distanzmatrizen für folgende Distanzen:

(a) Euklidische

$$d_E(x_i, x_j) = \sqrt{\sum_k (x_{ik} - x_{jk})^2}$$

(b) Manhattan

$$d_M(x_i, x_j) = \sum_k |x_{ik} - x_{jk}|$$

(c) Korrelation

$$d_C(x_i, x_j) = 1 - \frac{\sum_k ((x_{ik} - \mu_i)(x_{jk} - \mu_j))}{n\sigma_i\sigma_j}$$

wobei x_{ik} der Expressionswert des i -ten Gens am k -ten Tag ist, μ_i der Mittelwert, σ_i die Standardabweichung der Expressionswerte des i -ten Gens und n die Anzahl der Tage ist.

$$\sigma_i = \sqrt{\frac{1}{n} \sum_k (x_{ik} - \mu_i)^2}$$

2. Normalisieren Sie die Eingabematrix mit dem folgenden *Standardisierungsansatz*:

Schritt 1. Für jeden Wert in der Zeile verwenden Sie die Transformation $x \mapsto \frac{x-\mu}{\sigma}$, wobei μ der Mittelwert und σ die Standardabweichung der Zeilenwerte ist.

Schritt 2. Für jeden Wert in der Spalte verwenden Sie dieselbe Transformation mit dem Unterschied, dass der Mittelwert und die Standardabweichung nun über die Spalten berechnet werden.

Wiederholen Sie den Schritt 1.

Wiederholen Sie den Schritt 2.

usw ...

Wie sieht die Matrix nach einem, zwei, drei Schritten aus? (Zusatzaufgabe: Wie sieht die Matrix nach 10, 100, 1000 Schritten aus?) Was können Sie dabei beobachten?