

Übungen zur Microarrayanalyse: Spezielle Themen (19.2.2004)

- 1) Laden Sie das Paket GeneTS (Methoden zur Analyse von genetischen Zeitreihen)
- 2) Laden den Beispieldatensatz "caulobacter".
Finden Sie mehr über diesen Datensatz heraus:
 - wie viele Arrays?
 - wie viele Gene?

Identifizierung periodischer Gene:

- 3) Plotten und interpretieren Sie das "Average Peridogram"
- 4) Testen Sie jedes Gen mit Hilfe von Fischers exaktem g-Test auf Periodizität und berechnen Sie den entsprechenden p-Wert
- 5) Bestimmen Sie die Zahl der signifikanten Gene mit Hilfe der FDR Methode (multiples Testen für den Level $Q=0.05$) und konstruieren Sie eine entsprechende Datenmatrix

Schätzung des genetischen Korrelationsnetzwerkes (GGM):

- 6) Schätzen Sie die partiellen Korrelationen der Caulobacte-Daten mit Hilfe von drei Methoden:
 - beobachtete partielle Korrelation (π_1)
 - partielle gebaggte Korrelation (π_2)
 - gebaggte partielle Korrelation (π_3)Vergleichen Sie die Resultate.
- 7) Benutzen Sie π_2 um die Nullverteilung der partiellen Korrelationen zu bestimmen und damit alle möglichen 990 Kanten zu testen. Wieviele signifikante Gene gibt es bei einem FDR cut-off von $Q=0.05$?
- 8) Wiederholen Sie die Analyse von 7) für die anderen beiden Schätzer π_1 und π_3 .
Was stellen Sie fest?
- 9) Simulieren Sie ein zufälliges Netzwerk mit 40 Knoten (wahre partielle Korrelationen). Ziehen Sie dann eine Stichprobe der Größe 40 aus der dazugehörigen Verteilung und schätzen Sie die partiellen Korrelationen mit der Methode π_1 und π_3 . Welcher Schätzer ist besser?

Schätzung von genetischen Interaktionsmustern ("emerging patterns"):

- 10) Lesen Sie den R Code für "Emerging Patterns" (EPs) sowie den reduzierten Beispieldatensatz "Leukemie" ein.

- 11) In dieser Frage wird das EP $\{X_{3455} < -0.37\}$ and $\{X_{2285} < 0.55\}$ betrachtet.
Plotten Sie auf derselben Graphik X_{2285} gegen X_{3455} für die Patienten von Klasse 1 (in Schwarz) und die Patienten von Klasse 2 (in Rot).
- 12) Suchen Sie nach Emerging Patterns mit Parametern $p_g = 1e-17$ and $p_s = 1e-5$.
Wieviele EPs erhalten Sie?
- 13) Suchen Sie nach EPs mit Parametern $p_g = 1e-13$ and $p_s = 1e-9$.
Wieviele von den erhaltenen EPs sind Interaktionen zwischen 2 Genen ?