

Sessionsprüfung

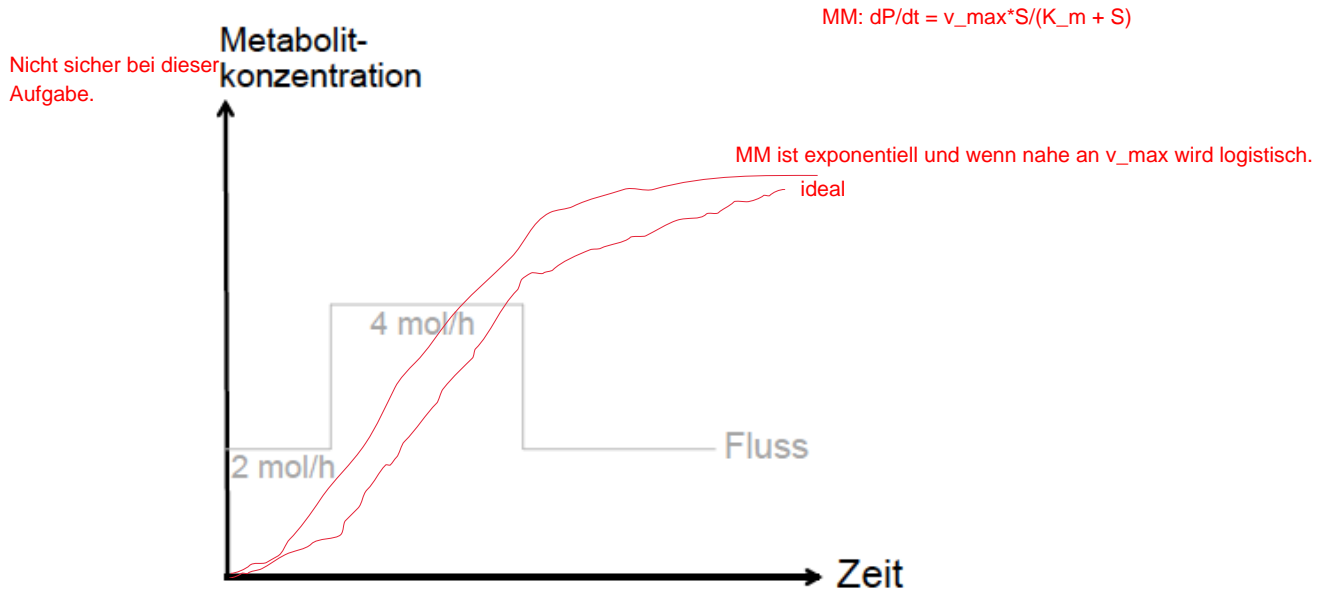
August 2015

Systembiologie: 551-1174-00L

- *Die Prüfungssprache ist deutsch, aber Sie dürfen englische Ausdrücke verwenden.*
- *Insgesamt sind 44 Punkte mit 16 Fragen zu erreichen. 22 P = 4.0; 40 P = 6.0!*
- *24 Punkte im Stelling/Sauer Teil; 20 Punkte im Borgwardt/Zamboni Teil.*
- *Bitte antworten Sie möglichst kurz.*
- *Nach Möglichkeit bitte nicht die Rückseiten verwenden.*
- *Es sind keine Hilfsmittel erlaubt.*

Prüfungsteil Sauer/Stelling

1. Alle Enzyme eines hypothetischen Stoffwechselweges reagieren ausschliesslich mit Michaelis-Menten Kinetiken. Der Fluss in diesen Stoffwechselweg erhöht sich von 2 mol/h als Puls auf 4 mol/h, um dann wieder auf 2 mol/h zurückzugehen wie abgebildet. A) Bitte skizzieren Sie den zu erwartenden Verlauf der Konzentrationen EINES Metaboliten im Stoffwechselweg bei „realer“ Michaelis-Menten Kinetik und das „ideale“ Verhalten, welches von der Zelle am Besten zu tolerieren wäre? (1 P)



- B) Nennen Sie ein Beispiel für ein aus diesem realen Verhalten resultierendes Problem?(1 P)

[Zu hohe Konzentration eines intermediären Metaboliten könnte zu Toxizität führen, bzw.] der ATP Vorrat für die nächste Reaktion könnte bereits aufgebracht sein welche nicht mehr stattfinden

- C) Wie könnte eine Zelle das Problem umgehen/reduzieren?

(1 P)

Indem es z.B. einen Kontrollmechanismus besitzt, der einen alternativen pathway aktiviert (z.B. ein pathway mit cAMP anstatt ATP, sodass sich das ATP regenerieren kann - vgl. Turbo Design aus der VL)

2. Sie möchten die dynamische Antwort eines Stoffwechselweges auf die Hemmung durch ein Antibiotikum verstehen.
- A) Welche Art von Modell verwenden sie dazu? (1 P)
- B) Durch welche 2 generellen Faktoren wird das Modellverhalten (Model Output) bestimmt?(2 P)
- C) Was können Sie tun, wenn Ihr Modell trotz aller Versuche geeignete Modellparameter zu finden nicht mit den gemessenen Daten übereinstimmt? (1 P)

a) mass-balance kinetics + michaelis menten kinetics

b) durch die Reaktionsraten und die jeweiligen Konzentrationen der Metaboliten zu einem Zeitpunkt t

c) Im Prinzip eine Korrekturfunktion anwenden: $\phi(K_m) = \text{abs}((x_{\text{measured}} - x_{\text{computed}})/\text{std_deviation})$ um mindestens mit Annäherungen arbeiten zu können

3. A) Was ist Ultrasensitivität und wie unterscheidet sie sich von einer Michaelis-Menten Kinetik? (1 P)
- B) Nennen Sie ein Beispiel wie eine ultrasensitive Systemantwort erreicht werden kann.(1 P)

a) Def. Ultrasensitivität: die Antwort der Zelle zu einem gegebenen flux ist stärker wie man es von normaler MM Kinetik erwarten würde. Mathematisch gesehen ist der Hill Koeffizient n relevant: $dP/dt = v_{\text{max}} * S^{**n}/(K_m^{**n} + S^{**n})$ mit $n \geq 2$

b) Kooperativität, Kaskaden, Multistep Ultrasensitivity

Model Topology

4. Ein biologisches System mit irreversiblen Reaktionen hat die in der Abbildung gezeigte Topologie (nehmen Sie Michaelis-Menten Kinetik für Enzyme E1 und E2 und einen Carbon Influx mit konstanter Rate r_{in} an).



- A) Schreiben Sie die Bilanzgleichungen für die Metabolitkonzentrationen. (1 P)
 B) Ändern sich die Steady State Konzentrationen von Metabolit A und/oder B wenn die Startkonzentration von Metabolit A von 4 mM auf 0.01 mM geändert wird? (1 P)
 C) Ändern sich die Steady State Konzentrationen von Metabolit A und/oder B wenn der K_m -Wert von Enzym 2 von 0.5 mM auf 0.25 mM geändert wird? (1 P)
 D) Der Carbon Influx r_{in} ist konstant und der Fluss durch Enzym E1 wird mit r_1 bezeichnet. Skizzieren Sie graphisch die Raten r_{in} und r_1 als Funktion der Konzentration von Metabolit A. (2 P)
 E) Existieren Fälle, in denen das System keinen stationären Zustand erreichen kann? Wenn ja, wann (Argumentieren Sie mit der Skizze aus D)? (1 P)

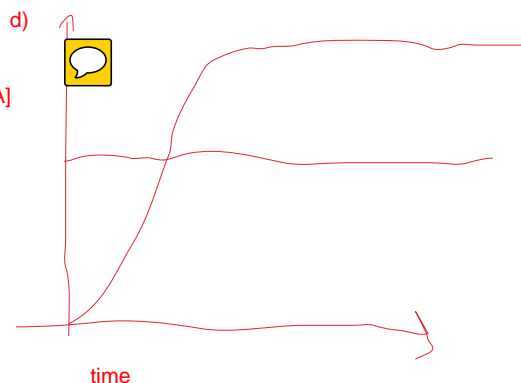
a) $\frac{dA}{dt} = r_{in} - E1 \cdot A$

$\frac{dB}{dt} = A \cdot E1 - E2 \cdot B$

wobei hier $E1/2$ als Reaktionsraten uminterpretiert wurden, alternativ hätte man auch k_1 und k_2 setzen können

b) die steady-state Konzentrationen ändern sich nicht, es dauert vermutlich etwas länger bis die steady-state Konzentrationen erreicht werden. r_{in} ist nach wie vor unverändert, sowie E1 und E2, da diese Parameter entscheidend für steady state sind.

c) ja, die steady-state Konzentrationen ändern sich, da die K_m bzw. E2 nicht mehr so viel umsetzen kann.

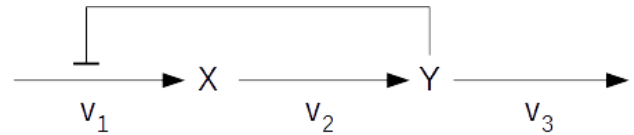


nicht sicher bei d)

r_{in} unterer Graph, gerade Linie, da konstant und unabhängig von $[A]$
 r_1 oberer Graph (darf aber auch unterhalb sein, falls $r_{in} < r_1$)

e) es existiert kein steady state, genau dann wenn z.B. $r_{in} \gg r_1$, dann kriegt man immer mehr influx, sodass $[A]$ immer grösser wird. Dann kann E1 nie genug A umsetzen ohne das immer mehr A nachkommt.

5. In dem abgebildeten Stoffwechselweg wird der Fluss v_1 durch das Endprodukt Y inhibiert; Fluss v_3 repräsentiert den zellulären Bedarf an Y.



- A) Welche Aussagen können Sie über die tatsächliche Konzentration von Y machen wenn das System im stationären Zustand ist und Sie nicht mehr zu dem Pathway wissen? **(1 P)**
 B) In Wirklichkeit ist der Pathway komplizierter, mit vielen Reaktionsschritten und Metaboliten zwischen X und Y. Wenn der Bedarf an Y stark fluktuiert, ist die Konzentration von Y auch mit Feedback immer nahezu konstant? **(1 P)**
 C) Wäre im Fall (B) ein negativer Feedback durch allosterische Regulation oder durch Genregulation vorteilhafter und wieso? **(1 P)**

a) je nachdem ob $v_3 > v_2$ ist, wird Y = 0 sein und der Pathway ist nie inhibiert.

b) ja wenn man $\langle Y \rangle_t$ betrachtet ($\langle Y \rangle_t = 1/T \cdot \text{INTEGRAL}(0, T, Y(t))$), dann $\langle Y \rangle_t$ ableiten \Rightarrow muss 0 geben \Leftrightarrow konstant
 Das Integral kann man mit sinus und cosinus ausdrücken, d.h. eine Fourieranalyse auf Messdaten ausführen (verwende fft - fast fourier transform in einem Programm)

c) Genregulation, da Y als transcription factor and die DNA-sequenz binden kann und sofort die Produktion der Enzyme/des Enzymes inhibiert. Bei Bedarf wird Y entfernt und die Transcription kann ungehindert losgehen.

6. Sie analysieren ein metabolisches Netzwerk im stationären Zustand. Das Netzwerk hat drei interne Metabolite A, B und C, fünf Reaktionen R_1 - R_5 mit Flüssen v_1 - v_5 und alle Reaktionen sind irreversibel. Die stöchiometrische Matrix ist rechts gegeben.

$$S = \begin{matrix} & \begin{matrix} r1 & r2 & r3 & r4 & r5 \end{matrix} \\ \begin{matrix} A \\ B \\ C \end{matrix} & \begin{pmatrix} +1 & 0 & +1 & -1 & 0 \\ 0 & +1 & -1 & -1 & 0 \\ 0 & 0 & 0 & +1 & -1 \end{pmatrix} \end{matrix}$$

Hinweis: Die Lösung von A) ist nicht Voraussetzung für die anderen Teile der Aufgabe.

- A) Skizzieren Sie das Reaktionsnetzwerk mit allen Metaboliten, Reaktionen und Systemgrenze. **(2 P)**
- B) Welches Paar von Reaktionen bildet einen sogenannten ‚enzyme subset‘ (begründen Sie die Antwort)? **(1 P)**
- C) Sie haben den Fluss für Reaktion R_1 gemessen: $v_1 = 2$ mol/h. Bestimmen Sie, wie v_4 von dem unbekannten Fluss v_2 abhängt und stellen Sie diese Abhängigkeit graphisch dar. **(2 P)**
- D) In einem neuen Experiment messen Sie $v_1 = 2$ mol/h und $v_2 = 1$ mol/h. Warum ist entweder die Messung oder das Modell nicht korrekt? **(1 P)**

a) siehe blatt

b) ? (3,4)

c) vermutlich ist v_4 (bzw. r_4) einfach 1 mol/h; graph=?

d) steady state ist nicht mehr garantiert da der Eintrag in (C, r_4) nicht mehr so stimmen kann
(I think this answer is not correct)

Prüfungsteil Borgwardt/Zamboni

7. Nennen Sie zwei Gründe, warum univariate Featureselektion in der Systembiologie ein sehr limitierter Ansatz ist! (2 P)

es kann nur ein feature identifizieren, meist gibt es viele weitere features.

Ein feature kann in mehreren Gruppen eine Rolle spielen, das kann durch univariate feature selection nicht erkannt werden.

Ebenfalls: berücksichtigt nicht die additivität von features, die Interaktion von diesen und die Beziehung von diesen.

8. Wozu wird (A) die L1-Norm bzw. (B) die L2-Norm eingesetzt bei der multivariaten Featureselektion? (2 P)

a) minimieren von β , sodass möglichst viele einträge in β ungleich 0 sind, β ist ein n -dim spaltenvektor => erklären so ganz viele SNPs als irrelevant

b) minimieren von $\sqrt{\sum \beta^2}$, sodass möglichst viele einträge gleiche weights haben, welche miteinander korrelieren

(an der Prüfung auch die Formeln aufschreiben: $\text{norm}(\beta)_1$ bzw. $_2$)

9. Was versteht man unter Initialisierungsabhängigkeit bei k-means Clustering? (1 P)

man wählt im ersten Schritt k zufällige Datenpunkte welche unsere centroids sind und lässt den Algorithmus laufen.
Da in einer Wiederholung wieder k zufällige Datenpunkte ausgewählt werden, wird es zu (leicht oder sehr) verschiedenen Klustern kommen.
Eventuell gibt es unterschiedliche Clusters, abhängig davon was k für einen Wert hat.

10. Nennen Sie die zwei zentralen Schritte beim Lloyds-Algorithmus für k-means! (2 P)

- 1) assign to each point its nearest cluster.
 - 2) recompute the cluster mean point (which needs not to be an actual data point)
- [repeat 1) and 2)]

11. Was ist der Single-Link-Effekt beim graphbasierten Clustering? (1 P)

?

Hat man einen Knoten falsch berechnet/gemessen und man würde einen Link zwischen zwei Clustern an diesem Knoten inferieren, dann würde ein Cluster entstehen, das nur durch diesen einen Knoten gegeben ist. Jedoch ist das Ergebnis in dem Fall verfälscht, daher der Name Single Link Effect.

-
- 12. Was ist der zentrale Grund, warum häufig die AUC bei der Interaktionsvorhersage gut aussieht, während Precision und Recall ernüchternd niedrig sind? (2 P)**

?

(An der Prüfung hätte ich geschrieben: beide Methoden sind äquivalent, lediglich die optische Präsentation sieht bei AUC eindrücklicher aus, aber wirklich wissen tu ich das nicht.)

- 13. Nennen Sie jeweils ein Anwendungsbeispiel für Clustering von (A) Proben und von (B) Variablen in der Systembiologie! (2 P)**

- a) Genom: welche Gene gehören in dasselbe Cluster (verwende 3-D hierarchical clustering)
b) feature selection with latent variables

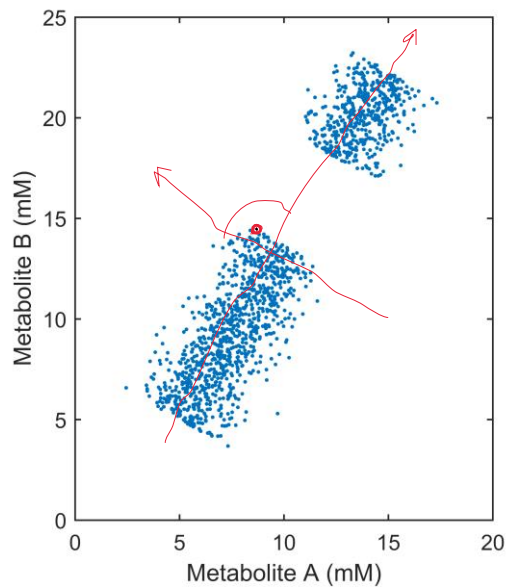
im allgemeinen nicht sicher bei den Antworten hier, explizites BSP für b)?

- 14. Wie wird eine Anreicherungsanalyse (Gene Set Enrichment Analysis) durchgeführt? (3 P)**

- 1) zuerst feature selection um features zu identifizieren.
- 2) für jede feature testen wir ob sie überrepräsentiert sind:
calculate mean of feature_i in group A and group B
calculate fold change $FC = (\text{mean_of_groupA} / \text{mean_of_groupB})$ and plot it versus the significance values p.
Also, set $p=0.05$ or 0.01
If necessary, perform bonferroni correction of FWER on p.
build contingency table: (make a drawing of the contingency table)
- 3) use fisher's exact test with the correct inputs from the contingency tables

15. Sie haben die Konzentration von 2 Metaboliten im Blut von ca. 1000 Individuen gemessen und bekommen das folgende Bild.

(A) Zeichnen sie auf der Abbildung die Richtung der zweiten Hauptkomponente (PC). **(1P)**



(B) Wieviele Hauptkomponenten würde Sie nehmen um den Datensatz ausreichend zu repräsentieren? Warum? **(1P)**

1 hauptkomponente, denn die zweite hauptkomponente ist immer rechtwinklig auf der ersten und ebenfalls gilt:

Sei P die Projektionsfunktion, welche die hauptkomponenten auf die x-Achse projiziert. Dann gilt:

$\{P(2.\text{Hauptkomponente})\}$ ist element von $\{P(1.\text{Hauptkomponente})\}$

(C) Hauptkomponenten sind durch Vektoren definiert. Erklären Sie warum es für jede einzelne Hauptkomponente immer zwei gleichwertige Definitionen gibt. **(1P)**

?

(Welche zwei definitionen gibt es denn überhaupt?)

(An der Prüfung hätte ich geschrieben: erste Definition ist über die Maximale Varianz der Datenpunkte, zweite definition ist: PC2 ist rechtwinklig auf PC1 mit maximaler Varianz der Datenpunkte, jedoch nicht sicher bei dieser Aufgabe.)

An answer: one can either the PC or use its mirror PC, same is true for PC2

16. Nennen zwei Anwendungsbeispiele für die Vorhersage von biologischen Netzwerken.(2 P)

Vorhersage ob ein Transcriptionsfactor mit einem locus interagiert.

Vorhersage ob es zur allosterischen regulation von Enzymen kommt bei bekannten small molecules

Vorhersage der Interaktion von small molecules mit Proteinen im Allgemeinen (z.B. im Sinne von gene-knockdowns)

Vorhersage von protein-protein interaktionen.

Krankheitsmechanismen verstehen => wie wirkt ein Medikament

Wie wirkt sich eine Mutation auf das Wachstum aus