Metabolische Flussanalyse

Verwendung der Maximum-Entropie-Methode

Carola Heinzel

Dr. Katharina Nöh, Johann Fredrik Jadebeck

27.09.2023

Inhaltsverzeichnis

Motivation

- $\mathbf{v} \in \mathbb{R}^d$: Fluss-Vektor
- Ziel: Löse $A\mathbf{v} \leq \mathbf{b}$ und beachte dabei
  (i) die durchschnittliche Wachstumsrate μ_λ
  (ii) μ_λ und die empirische Varianz der Wachstumsrate σ_λ^2
- Dazu: Verwende *Maximum-Entropie (MaxEnt)*

MaxEnt - intuitive Herleitung im Diskreten

- K Urnen
- n_i: Anzahl Bälle in Urne i
- $N = n_1 + \ldots + n_K$: Gesamtanzahl Bälle
- Ω: Anzahl Möglichkeiten N Bälle in K Urnen zu legen
- $p_i = n_i / N$: Wahrscheinlichkeit, dass ein gezogener Ball aus Urne i kommt
- Entropie $H = -\sum_{i=1}^{K} p_i \ln(p_i)$
- $\Omega = \frac{N!}{n_1! \cdot \ldots \cdot n_K!} \approx e^{HN}$ –> Maximale Anzahl an Möglichkeiten $\Leftrightarrow$ Maximale Entropie

MaxEnt - intuitive Herleitung im Diskreten

- Teile N Kugeln zufällig auf K Urnen auf: Die wahrscheinlichsten Werte von $n_1, \ldots, n_K$ sind diejenigen mit der maximalen Entropie
- D.h. Wähle die Dichte $p = (p_1, \ldots, p_K)$ aus, welche die maximale Entropie hat (p_{max}).
- Ohne weitere Bedingungen: p_{max} ist die Dichte der Gleichverteilung
- MaxEnt-Verteilung = Verteilung, welche die Bedingungen erfüllt und "bestmöglich" zu den Daten passt.
- Hier: Betrachte Zellen (Bälle) und ihre Flussvektoren (Urnen) mit zusätzlichen Bedingungen $\mu_\lambda, \sigma_\lambda^2$

MaxEnt - 1. Fall

- Bekannt: μ_λ
- $\lambda(\mathbf{v})$: Wachstumsrate
- $Z(\beta) := \int_{\mathcal{P}} e^{\beta\lambda(\mathbf{v})}\mathrm{d}\mathbf{v}$
- MaxEnt –> Boltzmann-Verteilung

$$p(\mathbf{v}) = \frac{1}{Z(\beta)} e^{\beta\lambda(\mathbf{v})}.$$

- Bestimme β durch Lösen von

$$\frac{1}{\int_{\mathcal{P}} e^{\beta\lambda(\mathbf{v})}\mathrm{d}\mathbf{v}} \int_{\mathcal{P}} \lambda(\mathbf{v}) e^{\beta\lambda(v)}\mathrm{d}\mathbf{v} = \mu_\lambda$$

- Näherung:

$$\frac{1}{\int_{\mathcal{P}} e^{\beta\lambda(\mathbf{v})}\mathrm{d}\mathbf{v}} \int_{\mathcal{P}} \lambda(\mathbf{v}) e^{\beta\lambda(v)}\mathrm{d}\mathbf{v} \approx \frac{\sum_i \lambda(\mathbf{v}^{(i)}) e^{\beta\lambda(\mathbf{v}^{(i)})}}{\sum_i e^{\beta\lambda(\mathbf{v}^{(i)})}}$$

Boltzmann-Verteilung

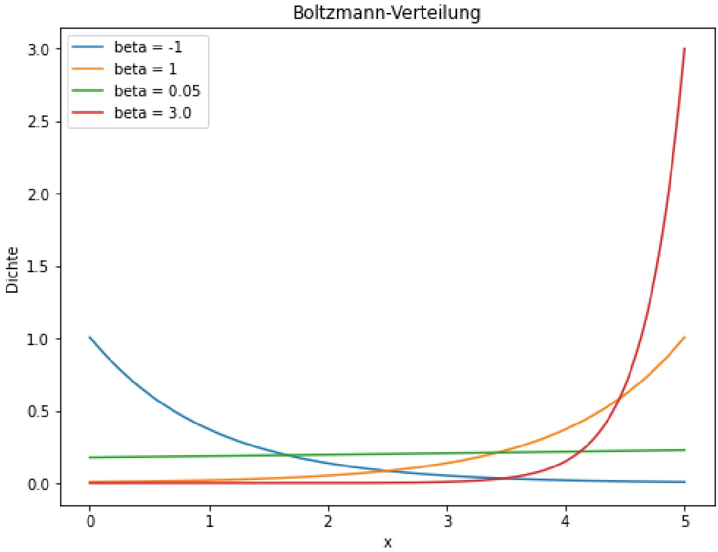

Abb.: Dichte der Boltzmann-Verteilung für verschiedene Parameter

Relativer Monte-Carlo-Fehler

- Stichprobengröße: 1000
- Hier: zu wenig betrachtete Fälle

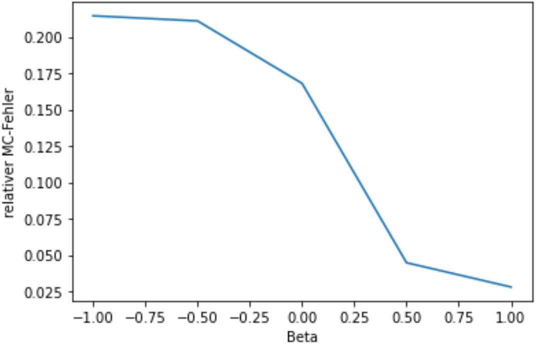

Abb.: Relativer Monte-Carlo-Fehler

MaxEnt - 2. Fall

- Bekannt: $\mu_\lambda, \sigma_\lambda^2$
- Dann gilt

$$p(\mathbf{v}) = c \cdot exp\left(-\frac{(\lambda(\mathbf{v}) - \mu_\lambda)^2}{2\sigma_\lambda^2}\right).$$

Theoretische Anwendung

1. Bestimme $\mu_\lambda, \sigma_\lambda^2$ aus den Daten: *Regression*
2. Bestimme $\lambda(\mathbf{v})$: *Steht im Modell*
3. Bestimme den Parameter der Boltzmann-Verteilung: *Löse die dazugehörige Gleichung*
   Parameter der Normalverteilung: *Bereits bestimmt*
4. Finde **v**, sodass $A\mathbf{v} \leq \mathbf{b}$ mit Verwendung der Informationen $\mu_\lambda, \sigma_\lambda^2$: *hopsy*

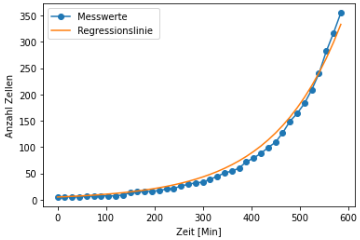

Abb.: Bestimmung der Wachstumsrate durch Regression

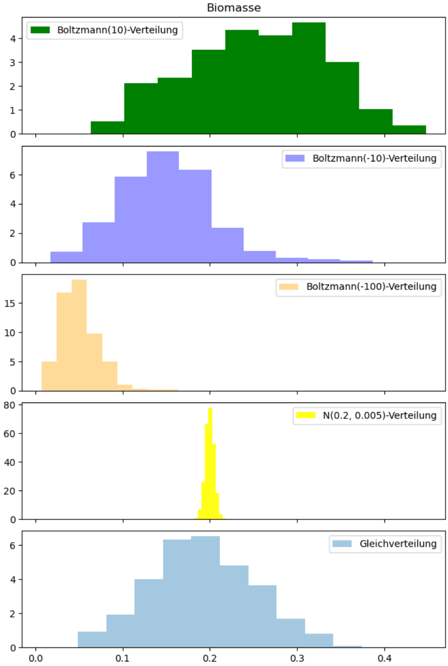

Abb.: Balkendiagramm zur Darstellung der simulierten v

Zusammenfassung und Ausblick

- Haben Infos $\mu_\lambda, \sigma_\lambda^2$ zur Bestimmung von $\mathbf{v}$ verwendet
- ESS ≈ 28, $\hat{R} \approx 1.14$
- Bestimmung der richtigen Einheiten
- Fehlerschätzung der Näherung von β
- Bestimmung des Einflusses von β und $\mu_\lambda, \sigma_\lambda^2$ auf die Ergebnisse: Wasserstein-Distanz, Vergleich von zwei Vektoren, statistische Tests?

Vielen Dank für die Aufmerksamkeit!