Vorlesung 12b

Relative Entropie

Teil 3:

Entropieschranken

In den folgenden Beispielen

benutzen wir den eben bewiesenen Satz in der Gestalt

$$(*) \quad \mathbf{H}[\rho] \leq -\sum_a \rho(a) \log \pi(a)$$

mit Gleichheit genau für $\rho = \pi$.

Wir sehen:

Jede Wahl von π liefert in $(*)$ eine Schranke für $\mathbf{H}[\rho]$,

mit Gleichheit genau für $\rho = \pi$.

Für jede Wahl von π wird die rechte Seite von $(*)$
zum Erwartungswert der Zufallsvariablen $g(X) := -\log \pi(X)$.

$(*) \quad \mathbf{H}[\rho] \leq -\sum_a \rho(a) \log \pi(a)$ mit Gleichheit genau für $\rho = \pi$.

Beispiel 1: Vergleich mit der uniformen Verteilung:

Sei S endlich mit n Elementen

und sei $\pi(a) = 1/n$ für alle $a \in S$.

Dann folgt aus $(*)$ für jede Verteilung ρ auf S:

$$\mathbf{H}[\rho] \leq -\sum_a \rho(a) \log \left(\frac{1}{n}\right) = \log n \; .$$

$$\boxed{\mathbf{H}[\rho] \leq \log n.}$$

Gleichheit gilt genau im Fall der uniformen Verteilung,

sie maximiert auf S die Entropie. $\square$

$(*) \quad \mathbf{H}[\rho] \leq -\sum_a \rho(a) \log \pi(a)$ mit Gleichheit genau für $\rho = \pi$.

Beispiel 2: Vergleich mit der um -1 verschobenen geometrischen Verteilung:

Sei nun $S = \{0, 1, 2, \ldots\}$, und $\pi(k) := 2^{-k-1}$.

Dann folgt aus $(*)$ für alle Verteilungen ρ mit EW $\mu(\rho)$:

$$\mathbf{H}_2[\rho] \leq -\sum_k \rho(k) \log_2(2^{-k-1})$$

$$= \sum_{k=0}^{\infty} \rho(k)(k+1) = \mu(\rho) + 1 \, .$$

Gleichheit gilt für $\rho = \pi$, dann ist $\mathbf{H}_2[\rho] = 2$. Also:
Unter allen Verteilungen auf $\mathbb{N}_0$ mit EW ≤ 1

hat die Verteilung π die größte binäre Entropie, nämlich 2.$\square$

Im nächsten Beispiel betrachten wir

(für eine Abbildung $u : S \to \mathbb{R}$)

die Frage:

Wie sieht unter allen Verteilungen von X

mit vorgegebenem Wert v für $\mathbf{E}[u(X)]$

diejenige mit der größten Entropie aus?

(Das obige Beispiel 2 passt in diesem Rahmen mit $u(k) := k$)

Wieder verwenden wir die Informationsungleichung

in der Form

$$(*) \quad \mathbf{H}[\rho] \le -\sum_a \rho(a) \log \pi(a)$$

mit Gleichheit genau für $\rho = \pi$.

Beispiel 3: Vergleich mit einer "Boltzmann-Gibbs-Verteilung":

Gegeben sei $u : S \to \mathbb{R}$, $\beta \geq 0$.

Wir definieren die Gewichte $\pi(a) := e^{-\beta u(a)}/z$ mit

$$z := \sum_{a \in S} e^{-\beta u(a)} \quad \text{(Annahme: } z < \infty.\text{)}$$

Die Abschätzung $(*)$ ergibt für alle Verteilungen ρ auf S:

$$\mathbf{H}_e[\rho] \leq - \sum_{a \in S} \rho(a) \ln \pi(a) = - \sum_{a \in S} \rho(a)(-\beta u(a) - \ln z)$$

$$= \beta \sum_{a \in S} \rho(a) u(a) + \ln z$$

mit Gleichheit genau für $\rho = \pi$.

Anders gewendet:

Sei π wie auf der vorigen Folie, und $v := \sum_{a \in S} u(a)\pi(a)$.

Unter allen Zufallsvariablen X mit

vorgegebenem Erwartungswert $\mathbf{E}[u(X)] = v$

hat diejenige die größte Entropie,

die die Verteilungsgewichte $\pi(a), \;\; a \in S$, hat.

Die Verteilung mit den Gewichten $e^{-\beta u(a)}/z$

heißt *Boltzmann-Gibbsverteilung*

zum Potenzial u mit Parameter β.