Selbst-organisierende, adaptive Systeme (WS 16/17)

Übungsblatt 02 (Bearbeitung bis: 02.11.2016, 23:59 Uhr)

Chaostheorie, Statistisches Testen und Entropie

Im Rahmen dieser Übung werden die Grundlagen komplexer Systeme sowie die informationstheoretische Entropie vertieft. Außerdem verbessern Sie den Umgang mit der Simulationsumgebung NetLogo. Zur Erinnerung: Entsprechende Dokumentation finden Sie unter `http://ccl.northwestern.edu/netlogo/docs/`.

Aufgabe 1. ***(Chaostheorie)***

Zunächst entwickeln Sie NetLogo-Modelle zur Untersuchung der in der Vorlesung behandelten Phänomene. Veranschaulichen Sie sich das Verhalten von chaotischen Systemen am Beispiel des logistischen Modells bzw. der logistischen Abbildung. Verwenden Sie dazu die Datei "B02A01-LogisticMap.nlogo" aus Digicampus.

a) Vervollständigen Sie dazu das User-Interface: fügen Sie einen Slider für die globale Variable `R` hinzu, der auf Werte zwischen 0 und 4.0 gesetzt werden kann. Fügen Sie zudem einen Slider für die Initialbedingung (globale Variable `y0`) hinzu für Werte zwischen 0 und 1.

b) Platzieren Sie Monitore für die Werte von `y-current` und `y-new`, welche für $y(t)$ und $y(t+1)$ aus der Vorlesung stehen. Zudem benötigen Sie noch einen Plot für die Werte von `y-current`, um den Orbit der erzeugten Folge zu betrachten.

c) Vervollständigen Sie den Code, indem Sie die kommentierten Prozeduren gemäß den Vor- und Nachbedingungen implementieren. Achten Sie auf die Schreibweise der global bereits definierten Variablen.

d) Testen Sie Ihre Implementierung mit verschiedenen Parametern von R. Können Sie Bifurkationen beobachten? Wenn ja, bei welchen Werten?

Aufgabe 2. ***(Entropie)***

In der Vorlesung wurde die informationstheoretische Entropie nach Shannon [2] als Maß für die (Un)-Ordnung eines Systems besprochen. Da die Entropie in vielfältigen Informatik-Anwendungen wie Bildverarbeitung, Signalverarbeitung, Maschinelles Lernen etc. eine Rolle spielt, lohnt es sich, ein paar mathematische Eigenschaften der Entropie zu verstehen[1].
Für eine endliche Zufallsvariable $X : \Omega \to \mathbb{R}$ für eine geeigneten Wahrscheinlichkeitsraum $(\Omega, \mathfrak{P}(\Omega), \mathbb{P})$ (siehe Folien) ist die Entropie wie folgt definiert:

$$H^X = -\sum_{x \in X(\Omega)} \mathbb{P}^X(x) \log_2(\mathbb{P}^X(x))$$

wobei die Wahl des Zweierlogarithmus völlig beliebig ist, uns aber eine Intuition in Form von Bits liefert. Beachten Sie, dass die Entropie über dem induzierten Wahrscheinlichkeitsmaß $\mathbb{P}^X$ definiert ist, *nicht* über die Elementarergebnisse $\omega \in \Omega$ oder die Werte $x \in X(\Omega)$.

a) Nehmen Sie zunächst an, dass $\mathbb{P}^X(x) = 1$ für ein $x \in X(\Omega) \subseteq \mathbb{R}$ ist und $\mathbb{P}^X(x') = 0$, falls $x \neq x'$. Wie hoch ist die Entropie H_X? [2]

[1] Sollten Sie nachschlagen wollen, bietet [1] eine exzellente Einführung

[2] Per stetiger Fortsetzung vereinbaren wir $0 \log_2(0) = 0$

b) Betrachten wir nun einen einfachen Münzwurf. Definieren Sie einen geeigneten Wahrscheinlichkeitsraum und eine dazu passende Zufallsvariable. Bezeichnen wir die Wahrscheinlichkeit, Kopf zu erhalten mit h (folglich hat Zahl welche Wahrscheinlichkeit?): Berechnen Sie H^X für $h = 0$, $h = \frac{1}{2}$ und $h = 1$, was stellen Sie fest?

c) Visualisieren ("plotten") Sie die Entropie für verschiedene Werte von h im Interval $[0, 1]$ (X-Achse gibt h an, Y-Achse H^X) mit einer Programmiersprache Ihrer Wahl (z.B. Python mit `matplotlib`).

d) Daraufhin zeigen Sie, dass eine endliche Zufallsvariable mit einer Gleichverteilung maximale Entropie aufweist. Dieser Beweis kann in zwei Schritten geführt werden. Verwenden Sie hierzu die relative Entropie (oder Kullback-Leibler-Divergenz) als Maß für die „Distanz“[3] zwischen Wahrscheinlichkeitsmaßen $\mathbb{P}_1$ und $\mathbb{P}_2$:

$$D(\mathbb{P}_1 \parallel \mathbb{P}_2) = \sum_{x \in X(\Omega)} \mathbb{P}_1^X(x) \log \frac{\mathbb{P}_1^X(x)}{\mathbb{P}_2^X(x)}$$

1. (Extra-Aufgabe) Als erster Schritt ist zu zeigen, dass $D(\mathbb{P}_1 \parallel \mathbb{P}_2) \geq 0$ gilt. Hierzu kann die Jensensche Ungleichung (unter Zuhilfenahme des Umstandes, dass $\log(x)$ konkav ist und folglich $-\log(x)$ konvex) verwendet werden. Sei $f : \mathbb{R} \to \mathbb{R}$ eine konvexe Funktion ($f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$ für $0 \leq \lambda \leq 1$), und $Y = \{y_1, \ldots, y_n\}$ eine endliche Menge von Werten in der Domäne von f, dann gilt für nichtnegative λ_i mit $\sum\limits_{i=1}^{n} \lambda_i = 1$:

$$f\left(\sum_{i=1}^{n} \lambda_i y_i\right) \leq \sum_{i=1}^{n} \lambda_i f(y_i)$$

bzw.

$$f\left(\sum_{i=1}^{n} \lambda_i y_i\right) \geq \sum_{i=1}^{n} \lambda_i f(y_i)$$

falls f konkav ist.

2. Unter Berücksichtigung von $D(\mathbb{P}_1 \parallel \mathbb{P}_2) \geq 0$ ist nun zu zeigen, dass ein Wahrscheinlichkeitsmaß $\mathbb{P}_2$, das zu $\mathbb{P}_2^X(x) = \frac{1}{|X|}$, $\forall x \in X(\Omega)$ führt, die maximale Entropie besitzt, d.h., $H^X_{\mathbb{P}_2} \geq H^X_{\mathbb{P}_1}$ für alle Wahrscheinlichkeitsmaße $\mathbb{P}_1$. Wir vereinbaren dafür einen Wahrscheinlichkeitsraum $(\Omega, \mathfrak{P}(\Omega), \mathbb{P}_2)$[4], sodass $\Omega = M$ für eine endliche Menge M und $\mathbb{P}_2(\varepsilon) = \frac{|\varepsilon|}{|M|}$ für $\varepsilon \subseteq \Omega$. Die Zufallsvariable X nummeriert einfach alle Elemente von M (ist also *injektiv*, $X(m) \neq X(n)$, wenn $m \neq n$).

Literatur

[1] Thomas M Cover and Joy A Thomas. *Elements of information theory.* John Wiley & Sons, 2012.

[2] Claude Elwood Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 1948.

[3] Stefan Tappe. *Einführung in die Wahrscheinlichkeitstheorie.* Springer, 2013.

[3] Die KL-Divergenz ist im strengen Sinne keine Distanz, da sie im Allgemeinen nicht symmetrisch ist und die Dreiecksungleichung nicht erfüllen muss.

[4] Beachten Sie bitte, dass die Notation $\mathfrak{P}(\Omega)$ wie in [3] für die Potenzmenge von Ω steht, also $\{x \mid x \subseteq \Omega\}$