Birkhoffs Ergodensatz (oder: Wie wir die Zukunft vorhersagen)

Der fundamentale Ergodensatz besagt etwas über die relative Häufigkeit, mit welcher ein Orbit einer Abbildung T in eine bestimmte Menge trifft. In Anwendungen ist es häufig eher so, dass auf X (dem gesamten Raum) eine reelle Funktion f vorgegeben ist – nämlich eine **Messgröße**, welche in einem konkreten physikalischen Modell gemessen werden kann – und der Mittelwert dieser Funktion f längs des Orbits von T ist zu bestimmen. Dies können wir auch wirklich tun – und zwar in den wichtigen Fällen sogar dann, wenn wir fast nichts wissen über f, über T und über das Orbit. Kurz: Wir wissen praktisch nichts und schlussfolgern trotzdem praktisch alles!

Der folgende Ergodensatz von Birkhoff übersetzt den fundamentalen Ergodensatz in die Sprache von reellen Funktionen auf X:

THEOREM. ***Birkhoff-Ergodensatz:*** *Wenn (X, μ) ein Wahrscheinlichkeitsraum ist, $T : X \to X$ maßerhaltend bezüglich μ, und $f \in L^1(\mu)$ (d.h. $f : X \to \mathbb{R}$ integrierbar), dann existiert der Limes*

$$A_f(x) := \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x))$$

für fast alle (bezüglich μ) Punkte $x \in X$, und der Limes existiert in $L^1(\mu)$.

BEWEIS. Den Fall $f = \chi_B$ kennen wir schon, denn da ist die Behauptung gerade die Aussage des fundamentalen Ergodensatzes. Der Term, der beim fundamentalen Ergodensatz S_n hieß, ist gerade unser Term $\sum_{i=0}^{n-1} f(T^i(x))$. Außerdem gilt die Behauptung sicherlich auch für eine endliche Linearkombination von solchen charakteristischen Funktionen ist, also $f = \sum_{k=1}^{N} a_k \chi_{C_k}$ mit $a_k \in \mathbb{R}$ und C_k disjunkte messbare Mengen.

Wir wissen aus der Analysis, dass es zu jeder L^1-Funktion f und für alle $\varepsilon > 0$ eine Funktion f_ε gibt, deren Werte eine diskrete Menge in $\mathbb{R}$ sind und so dass $|f(x) - f_\varepsilon(x)| < \varepsilon$ für alle $x \in X$ gilt, also

$$f_\varepsilon = \sum_k a_k \chi_{C_k}.$$

Für den Fall, dass f essentiell beschränkt ist, gilt dann, dass f_ε nur endlich viele Werte annimmt, also $f_\varepsilon = \sum_{k=1}^{N} a_k \chi_{C_k}$ wie vorhin, also existiert $A_{f_\varepsilon}(x)$ für fast alle x. Wegen der Abschätzung $|f(x) - f_\varepsilon(x)| < \varepsilon$ gilt dann

$$\overline{A_f}(x) := \limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)) \le \overline{A_f}(x) + \varepsilon$$

und

$$\underline{A_f}(x) := \liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) \geq \overline{A_f}(x) - \varepsilon,$$

also existiert der Limes $A_f(x)$. Für den Fall, dass f nicht essentiell beschränkt ist, überlegen wir uns erst, dass wegen der Integrierbarkeit von f gilt, dass $\sum_k a_k\mu(C_k)$ endlich ist, denn

$$\sum_k |a_k|\mu(C_k) = \int_X |f_\varepsilon| d\mu < \varepsilon + \int_X |f| d\mu < \infty.$$

Deswegen gilt

$$\overline{A_f}(x) \leq \sum_k a_k A_{\chi C_k}(x) + \varepsilon$$

und

$$\underline{A_f}(x) \geq \sum_k a_k A_{\chi C_k}(x) - \varepsilon,$$

also existiert auch hier der Limes $A_f(x)$ für fast alle $x \in X$. Das zeigt die erste Behauptung.

Nun zeigen wir, dass A_f in $L^1(\mu)$ liegt. Der Einfachheit halber nehmen wir an, dass $f \geq 0$ ist; ansonsten ersetzen wir f durch $|f|$, und wenn $A_{|f|} \in L^1(\mu)$ gilt, dann gilt wegen $|A_f| \leq A_{|f|}$ auch $A_f \in L^1(\mu)$.

Es gilt

$$\begin{aligned}\int_X A_f d\mu &\leq \int_X \overline{A_f} d\mu \\ &\leq \sum_k a_k \int_X A_{\chi C_k} d\mu + \varepsilon \\ &\leq \sum_k a_k\mu(C_k) + \varepsilon \\ &< \infty,\end{aligned}$$

also $A_f \in L^1(\mu)$ wie behauptet. □

Übrigens benutzen wir hierbei die Notation $f \in L^1(\mu)$ und verstehen damit, dass f automatisch eine Funktion auf X ist; dies ist gerechtfertigt, denn wenn wir ein Maß μ haben, dann haben wir automatisch auch dessen Definitionsbereich (eine σ-Algebra) gegeben, da diese Information in μ enthalten ist, und natürlich ist die Menge X durch die σ-Algebra bestimmt, weil X das größte Element davon ist.

Bei diesem Beweis sehen wir: In Beweisen der Ergodentheorie benötigt es oft geraume Zeit, um ganz elementare (und nicht besonders interessant klingende) Dinge zu zeigen, z.B. dass das Zeitmittel A_f überhaupt definiert ist. Dagegen sind diejenigen Fakten, die uns brennend interessieren, z.B. dass das Zeitmittel fast überall von x

unabhängig ist und gleich dem Raummittel ist, im Beweis ganz einfache Folgerungen aus den vorher langwierig gezeigten elementaren Dingen. Hier nun diese wichtige und einfache Folgerung:

COROLLARY. *Wenn (T, μ) eine ergodische Transformation (X, μ) auf einem Wahrscheinlichkeitsraum (X, μ) ist (somit μ invariant unter $T : X \to X$), und $f \in L^1(\mu)$ ist, dann existiert für μ-fast alle $x \in X$ der Limes*

$$A_f(x) := \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)),$$

genannt das ***Zeitmittel*** *von f (längs dem Orbit von x), und für μ-fast alle $x \in X$ ist das Zeitmittel von f längs dem Orbit von x gleich und hat den (konstanten) Wert*

$$\int_X f d\mu,$$

genannt das ***Raummittel*** *von f (über den gesamten Raum X).*

BEWEIS. Die Funktion A_f ist T-invariant, also wegen der Ergodizität von (T, μ) essentiell konstant bezüglich μ. □

Das ist nun ein reichlich erstaunliches Ergebnis! Es besagt Folgendes: Wir wollen den Wert einer Funktion (Messgröße) f längs Orbits einer Transformation T vorhersagen. Dabei wissen wir *praktisch nichts* über die Funktion f (denn L^1 ist eine Messgrößen eigentlich immer), wir wissen *sehr wenig* über die Transformation T (denn Ergodizität ist eine schwache Bedingung), und wir wisssen *überhaupt nichts* über das (möglicherweise extrem komplizierte) Orbit von x, ja wir kennen *nicht einmal den Startwert* des Orbits. Und trotz alledem können wir *mit hundertprozentiger Wahrscheinlichkeit* und *ganz exakt* vorhersagen, wie der Wert von f im Mittel über das Orbit ist!