ÜBER DIE VARIATIONSPRINZIPIEN DER INFORMATIONSTHEORIE

Von

K. SEITZ

Lehrstuhl für Mathematik, Technische Universität Budapest

Eingegangen am 28. März 1980

Vorgelegt von Prof. G. Szász

Diese Arbeit befaßt sich mit der informationstheoretischen Anwendung der Variationsrechnung, insbesondere mit der Kapazitätbestimmung einiger spezieller Informationskanäle.

Eines der Hauptresultate dieser Arbeit ist die Verallgemeinerung eines Variationsprinzips von Shannon durch die Anwendung des Rényischen Informationsmaßes. Es wird bewiesen, daß jede kontinuierliche, positive Wahrscheinlichkeitsdichtefunktion mit ihren Nebenbedingungen eine Lösung eines Entropie-Maximalisierungsproblems darstellt.

1. Einleitung und Problemstellung

Die Anwendung der Variationsrechnung in der Informationstheorie ermöglichte die Optimierung der Entropie von kontinuierlichen Wahrscheinlichkeitsvariablen, und zwar bei speziellen Nebenbedingungen [1—4].

Dies ermöglicht die Charakterisierung einiger kontinuierlicher Wahrscheinlichkeitsdichtefunktionen als Lösungen von Entropie-Maximalisierungsproblemen, welche mit entsprechenden Nebenbedingungen versehen sind.

Betrachten wir jetzt ein Optimierungsproblem, welches Shannon in seiner »A mathematical theory of communication« [5] behandelt:

$$-\int_0^\infty f(x)\ln f(x)\,dx = \text{Max}! \tag{1}$$

$$\begin{aligned} f(x) &\geqq 0;\ \text{falls } x > 0 \\ f(x) &= 0,\ \text{falls } x < 0 \end{aligned} \tag{2}$$

$$\int_0^\infty f(x)\,dx = 1 \quad (3), \qquad \int_0^\infty x f(x)\,dx = m \tag{4}$$

Die Lösung obiger Aufgabe — wie das Shannon bewiesen hat — ist die folgende Dichtefunktion:

$$f(x) = \begin{cases} 0 \ \text{falls } x < 0 \\ \dfrac{1}{m} e^{-\frac{x}{m}} \ \text{falls } x > 0 \end{cases} \tag{5}$$

Gestaltet man die Variationsaufgabe (1—4) so, daß diese statt der Shannonschen die Rényische Entropie bzw. deren Maximum enthalte, so gelangt man zum folgenden Variationsproblem:

$$\frac{1}{1-\alpha}\, 2 \log \left[\int_0^\infty f^\alpha(x)\,dx\right] = \text{Max!} \tag{1*}$$

$$\left.\begin{aligned} f(x) &\geq 0, \text{ falls } x > 0 \\ f(x) &= 0, \text{ falls } x < 0 \end{aligned}\right\} \tag{2*}$$

$$\int_0^\infty f(x)\,dx = 1 \qquad (3^*), \qquad \int_0^\infty x f(x)\,dx = m \tag{4*}$$

Nach Lösung des obigen Problems wird bewesien, daß jede kontinuierliche und positive Wahrscheinlichkeitsdichtefunktion als die Lösung eines Entropie-Maximalisierungsproblems — mit entsprechenden Nebenbedingungen — aufgefaßt werden kann. Schließlich wird die Kapazität eines speziellen Informationskanals ermittelt.

2. Die Maximalisierung der Rényischen Entropie unter isoperimetrischen Nebenbedingungen

Die Extremalen der Variationsaufgabe (1*) — (4*) sind beweisbar aus der Gleichung

$$\frac{1}{(1-\alpha)\ln 2}\, \frac{1}{\int_0^\infty f^\alpha(x)\,dx}\, \alpha f^{\alpha-1}(x) + \lambda + \mu x = 0 \tag{6}$$

zu bestimmen. Die Werte der Parameter λ und μ sind aus den Nebenbedingungen (3*), (4*) zu berechnen.

Läßt man die sehr langwierigen Detailrechnungen weg, so erhält man als Lösung die Funktion

$$f(x) = \begin{cases} 0 & , \text{ falls } x < 0 \\ \dfrac{\alpha}{(2\alpha-1)}\left[1 + \dfrac{1-\alpha}{(2\alpha-1)m}\, x\right]^{\frac{1}{\alpha-1}}, & \text{falls } x > 0 \end{cases} \tag{7}$$

wenn

$$\frac{\alpha}{\alpha-1} < -1$$

Die Lösung des Shannonschen Optimierungsproblems kann aus (7) durch Limesbildung abgeleitet werden: Aus (7) folgt nämlich:

$$\lim_{\alpha\to 1} f(x) = \begin{cases} 0 & , \text{ falls } x < 0 \\ \dfrac{1}{m}\, l^{-\frac{x}{m}}, & \text{falls } x > 0. \end{cases} \tag{9}$$

3. Ableitung von positiven Dichtefunktionen durch Informationsoptimierung

Es wird bewiesen, daß jede kontinuierliche und positive Wahrscheinlichkeitsdichtefunktion als eine Lösung eines Entropie-Maximalisierungsproblems aufgefaßt werden kann.

Wenn nämlich $\varphi(x)$ eine positive Dichtefunktion darstellt, so sind die Extremalen der hier folgenden Variationsaufgabe

$$I = -\int_{-\infty}^{\infty} f(x) \ln f(x)\, dx = \text{Max!} \tag{10}$$

$$f(x) \geq 0 \quad (11), \qquad \int_{-\infty}^{\infty} f(x)\, dx = 1 \quad (12), \qquad \int_{-\infty}^{\infty} \ln\varphi \cdot f(x)\, dx = \int_{-\infty}^{\infty} \varphi \ln \varphi\, dx \quad (13)$$

beweisbar der Form

$$f(x) = \omega \varphi^{\mu}(x) \tag{14}$$

wobei ω und μ Konstanten sind.

Aus (12) folgt:

$$\omega = [\int_{-\infty}^{\infty} \varphi^{\mu}(x)\, dx]^{-1} \tag{15}$$

Im Sinne der Zusammenhänge (13) und (14):

$$\int_{-\infty}^{\infty} \omega\varphi^{\mu} \ln\varphi\, dx = \int_{-\infty}^{\infty} \varphi \ln\varphi\, dx \tag{16}$$

Daraus folgt, wegen (10), (15) und (16):

$$I(\mu) = -\ln\omega - \mu \int_{-\infty}^{\infty} \varphi \ln\varphi\, dx \tag{17}$$

Die erste Ableitung nach μ der Funktion $I(\mu)$ ist

$$I'(\mu) = -\int_{-\infty}^{\infty} \varphi \ln\varphi\, dx + \frac{1}{\int_{-\infty}^{\infty} \varphi^{\mu}\, dx} \int_{-\infty}^{\infty} \varphi^{\mu} \ln\varphi\, dx \tag{18}$$

welche nur an der Stelle $\mu = 1$ gleich Null ist.

Die zweite Ableitung nach der Funktion $I(\mu)$ ist an der Stelle

$$I''(1) = \int_{-\infty}^{\infty} \varphi \ln\varphi\, dx - (\int_{-\infty}^{\infty} \varphi \ln\varphi\, dx)^2 < 0$$

So besitzt $I(\mu)$ an der Stelle $\mu = 1$ ein Maximum.

Aus Gleichung (15) folgt im Falle von $\mu = 1$ und $\omega = 1$ und schließlich wegen (14):

$$f(x) = \varphi(x) \tag{20}$$

Die Ungleichung (19) soll noch bewiesen werden.

Für beliebige reale Zahl λ gilt

$$\int_{-\infty}^{\infty} \varphi(x)\,[\ln\varphi - \lambda]^2\,dx > 0 \tag{21}$$

Es wird ausgenutzt, daß in (21) $\varphi(x)$ eine Wahrscheinlichkeitsdichtefunktion ist, und daraus folgt, daß

$$\int_{-\infty}^{\infty} \varphi \ln^2 \varphi\,dx - 2 \int_{-\infty}^{\infty} \varphi \ln \varphi\,dx\,\lambda + \lambda^2 > 0 \tag{22}$$

bzw.

$$(\lambda - \int_{-\infty}^{\infty} \varphi \ln \varphi\,dx)^2 + \int_{-\infty}^{\infty} \varphi \ln^2 \varphi\,dx - (\int_{-\infty}^{\infty} \varphi \ln \varphi\,dx)^2 > 0 \tag{23}$$

Durch die Wahl folgt aus (23) (19), daß

$$\lambda = \int_{-\infty}^{\infty} \varphi \ln \varphi\,dx \tag{24}$$

4. Bestimmung der Kanalkapazität durch die Optimierung von Doppelintegralen

Betrachten wir einen Informationskanal mit ξ Eingangs- und η Ausgangssignalen, bei welchen die gemeinsame Dichtefunktion von ξ und η, $f(x, y)$ ist.

Gegeben seien die bedingte Dichtefunktion der Wahrscheinlichkeitsvariablen η unter der Bedingung $\xi = x$, $f(y|x)$, sowie die auf die Variable x bezogene profizierte Dichtefunktion $q(y)$ von $f(x, y)$, und schließlich das Shannonsche Informationsmaß der auf die Variable y bezogenen, projizierten Dichtefunktion $p(x)$ von $f(x, y)$.

So wird für die Bestimmung der Kapazität des betrachteten Informationskanals die Lösung des folgenden Variationsproblems benötigt:

$$C = -\int_0^{\infty}\int_0^{\infty} f(x, y) \ln \frac{f(x, y)}{p(x)\,q(y)}\,dx\,dy = \text{Max!} \tag{25}$$

$$f(y|x) = \frac{f(x, y)}{p(x)} = \omega(x, y) \tag{26}$$

$$-\int_{-\infty}^{\infty} p(x)\ln p(x)\,dx = m \tag{27}$$

$$\int_{0}^{\infty} p(x)\,dx = 1, \quad (28) \qquad p(x) \geqq 0, \tag{29}$$

falls $f(x, y) = 0$ oder x oder y = Null.

Es ist sofort resichtlich, daß durch die Einführung der Bezeichnung

$$v(x) = \int_{0}^{\infty} \omega(x, y)\ln\frac{\omega(x, y)}{q(y)}\,dy \tag{30}$$

die Variationsaufgabe in folgender Form aufgeschrieben werden kann:

$$C = \int_{-\infty}^{\infty} v(x)p(x)\,dx = \text{Max!} \tag{31}$$

$$\int_{-\infty}^{\infty} p(x)\,dx = 1 \tag{32}$$

$$p(x) \geqq 0 \tag{33}$$

$$-\int_{-\infty}^{\infty} p(x)\ln p(x)\,dx = m \tag{34}$$

Durch Anwendung der Lagrangeschen Multiplikatoren läßt sich verhältnismäßig einfach beweisen, daß die Lösung des Optimierungsproblems (31)—(34) die folgende ist:

$$p(x) = \begin{cases} e^{\frac{v(x)-\lambda}{\mu}-1}, & \text{falls } x > 0, \\ 0 & , \text{falls } x < 0 \end{cases} \tag{35}$$

wobei der Werte der Parameter λ und μ aus dem folgenden Gleichungssystem zu ermitteln sind:

$$\int_{0}^{\infty} e^{\frac{v(x)}{\mu}}\,dx = e^{1-\frac{\lambda}{\mu}}, \tag{36}$$

$$\int_{0}^{\infty} v(x)e^{\frac{v(x)}{\mu}}\,dx = (\mu - \lambda - \mu m)e^{1-\frac{\lambda}{\mu}} \tag{37}$$

Wegen (31), (35), (37) und (38) beträgt die Kapazität des Informationskanals

$$C = \mu(1 - m) - \lambda$$

Literatur

1. Rényi, A.: Wahrscheinlichkeitsrechnung. Berlin (1962).
2. Richter, H.: Wahrscheinlichkeitstheorie Berlin—Heidelberg—New York (1966), 2. Auflage.
3. Lamperti, J.: Probability. New York—Amsterdam (1966).
4. Hestenes, M.: Calculus of Variations and Optimal Control Theory, Wiley (1966).
5. Shannon, C.: A mathematical theory of communications Bell System Techn. J., 27 (1948) m. No. 3.

Zusammenfassung

Es gelang das Variationsprinzip von Shannon durch Anwendung des Rényischen Informationsmaßes zu verallgemeinern. Es wurde bewiesen, daß jede kontinuierliche positive Wahrscheinlichkeitsdichtefunktion mit ihren Nebenbedingungen eine Lösung eines Entropie-Maximalisierungsproblems darstellt.

dr. Károly Seitz H-1521 Budapest