Zu den Ungleichungen der Informationstheorie

Von W. Kämmerer

Die Ungleichungen für Verbundquellen

$$H(Y/X) \leqq H(Y)$$

und

$$H(X, Y) \leqq H(X)+H(Y)$$

werden üblicher Weise mittels der für konvexe Funktionen geltenden Jensenschen Ungleichung bewiesen.

Hier wird ein direkter Beweis über Extremwertbestimmung dargeboten. Dieser Weg erscheint didaktisch empfehlenswert, da er die gleiche Methode benutzt, die im allgemeinen bei der Bestimmung des Extremwertes der Entropie einer Quelle schon zuvor verwendet wird. Betrachtet werden zwei Quellen **X und Y** mit den Wahrscheinlichkeitsfeldern p_i und q_j. Gefragt wird nach dem Extremwert der bedingten Entropie $H(Y/X) = -\sum_i \sum_j p_i q_{j/i} \ln q_{j/i}$ bezüglich der zwischen den Quellen bestehenden Abhängigkeit — sie wird durch das bedingte Wahrscheinlichkeitsfeld $q_{j/i}$ charakterisiert — bei festgehaltenem p_i und q_j.

Die Extremwertbestimmung ist somit unter den Nebenbedingungen

$$\sum_i p_i q_{j/i} = q_j \qquad (j=1, \ldots, m)$$

und

$$\sum_j q_{j/i} = 1 \qquad (i=1, \ldots, n)$$

durchzuführen.

Mittels der Lagrangeschen Multiplikatoren

$$\lambda_1, \ldots, \lambda_m \quad \text{und} \quad \mu_1, \ldots, \mu_n$$

erhält man als Funktional

$$L = -\sum_i \sum_j p_i q_{j/i} \ln q_{j/i} - \sum_j \lambda_j \Big(\sum_i p_i q_{j/i} - q_j\Big) - \sum_i \mu_i \Big(\sum_j q_{j/i} - 1\Big).$$

Partielle Differentiation nach $q_{j/i}$ liefert das Gleichungssystem

$$\frac{\delta L}{\delta q_{j/i}} = -p_i \ln q_{j/i} - p_i - \lambda_j p_i - \mu_i = 0, \quad \text{für} \quad i=1, \ldots, n \quad \text{und} \quad j=1, \ldots, m.$$

Die Auflösung nach $q_{j/i}$ liefert

$$q_{j/i} = e^{-1-\lambda_j-\frac{\mu_i}{p_i}} = e^{-1-\lambda_j} \cdot e^{-\frac{\mu_i}{p_i}}.$$

Summation über j liefert unter Berücksichtigung der Nebenbedingungen

$$\sum_j q_{j/i} = 1$$

$$e^{-\frac{\mu_i}{p_i}} \sum_j e^{-1-\lambda_j} = 1.$$

Daraus resultiert

$$e^{-\frac{\mu_i}{p_i}} = \text{const.} \qquad \text{und weiter, daß auch}$$

$$q_{j/i} = \text{const} \cdot e^{-1-\lambda_j} \qquad \text{unabhängig von } i \text{ ist.}$$

Aus den weiteren Nebenbedingungen

$$\sum_i p_i q_{j/i} = q_j \qquad \text{folgt dann}$$

$$q_{j/i} = q_j.$$

Dies ist aber die Aussage, daß der Extremwert dann angenommen wird, wenn die Quelle Y nicht von der Quelle X abhängt. Die bedingte Entropie geht dann in die Entropie $H(Y)$ über. Aus den zweiten partiellen Ableitungen von L ist zu entnehmen, daß der erhaltene Extremwert ein Maximum ist. Somit gilt

$$H(Y/X) \leqq H(Y).$$

Addiert man beiderseits $H(X)$, so erhält man, da

$$H(X) + H(Y/X) = H(X, Y)$$

ist,

$$H(X, Y) \leqq H(X) + H(Y).$$

(Eingegangen am 19. Dez. 1970)