Iterierte arithmetische Mittelung und eine Verallgemeinerung der Jensenschen Ungleichung

Autor(en): **Wellstein, H.**

Objekttyp: **Article**

Zeitschrift: **Elemente der Mathematik**

Band (Jahr): **37 (1982)**

Heft 3

PDF erstellt am: **21.09.2024**

Persistenter Link: https://doi.org/10.5169/seals-36392

Iterierte arithmetische Mittelung und eine Verallgemeinerung der Jensenschen Ungleichung

1. Iterierte arithmetische Mittelung

Die Bildung des arithmetischen Mittels $(x_1+x_2)/2$ der reellen Zahlen x_1 und x_2 kann als Austausch der Hälften $x_1/2$ und $x_2/2$ aufgefasst werden. Dieser Prozess, auf mehr als zwei Zahlen übertragen, führt zwar nicht unmittelbar, aber doch in der Grenze zum entsprechenden Ergebnis.
Ausgehend vom Verteilungsvektor $(x_1, \ldots, x_n)$ bildet man der Reihe nach durch Mitteln der Koordinaten 1 und 2, 2 und 3, 3 und 4, ... die Vektoren

$$\begin{array}{l}(x_1,\ldots,x_n), \qquad (x_1/2+x_2/2, x_1/2+x_2/2, x_3,\ldots,x_n), \\ (x_1/2+x_2/2, x_1/4+x_2/4+x_3/2, x_1/4+x_2/4+x_3/2, x_4,\ldots,x_n)\end{array}$$

usw., bis nach $n-1$ Schritten alle Koordinaten erfasst sind. Eine günstige Darstellung gelingt mit Hilfe der (n,n)-Matrizen

$$A_i = \frac{1}{2}\cdot \begin{array}{cc} \begin{pmatrix} 2 & & & & & 0 & \\ & \ddots & & & & & \\ & & 2 & & & & \\ & & & 1 & 1 & \cdots & \\ & & & 1 & 1 & & \\ & 0 & & \vdots & & 2 & \\ & & & \vdots & & & \ddots \\ & & & & & & 2 \end{pmatrix} & \begin{array}{c} \\ \\ \\ (i) \\ \\ \\ \\ \\ \end{array} \\ \quad\quad (i) & \end{array}$$

als Folge $(x_1, \ldots, x_n)$, $(x_1, \ldots, x_n)A_1$, $(x_1, \ldots, x_n)A_1A_2$, ..., $(x_1, \ldots, x_n)\prod_{i=1}^{n-1} A_i$.
Die Matrix $A := \prod_{i=1}^{n-1} A_i$ ist als Produkt doppelt-stochastischer Matrizen selbst doppelt-stochastisch. Nach Konstruktion besteht ihre letzte Spalte aus positiven Zahlen, denn die n-te Koordinate des Vektors $(x_1, \ldots, x_n)A$ ist

$$2^{1-n}x_1 + 2^{1-n}x_2 + 2^{2-n}x_3 + \cdots + 2^{-2}x_{n-1} + 2^{-1}x_n .$$

Nach dem Ergodensatz ([1], S. 395) existiert dann die Grenzmatrix $M := \lim_{k\to\infty} A^k$.

Ihre Zeilen stimmen, wieder nach diesem Satz, miteinander überein. Da zudem die Spalten die Summen 1 haben, ist

$$M = \frac{1}{n}\cdot \begin{pmatrix} 1 & \ldots & 1 \\ \vdots & & \vdots \\ 1 & \ldots & 1 \end{pmatrix} .$$

Die Folge $(x_1,\ldots,x_n)A^k$, $k=0,1,2\ldots$ konvergiert daher ebenfalls, und zwar, $\bar{x}=(1/n)\cdot\sum_{i=1}^{n}x_i$ gesetzt, gegen die Gleichverteilung $(x_1,\ldots,x_n)M=(\bar{x},\ldots,\bar{x})$.

2. Verallgemeinerung

Gegeben seien positive Zahlen $p_1,\ldots,p_n$. Zur Abkürzung wird $q_i=p_i/(p_i+p_{i+1})$, $q_i'=p_{i+1}/(p_i+p_{i+1})$, $i=1,\ldots,n-1$, gesetzt. Aus den doppelt-stochastischen Matrizen

$$P_i=\begin{array}{cc} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & 0 & & \\ & & & q_i & q_i' & \cdots & \\ & & & q_i' & q_i & & \\ & 0 & & \vdots & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} & \begin{matrix} \\ \\ \\ (i) \\ \\ \\ \\ \\ \end{matrix} \\ (i) & \end{array}$$

wird die Matrix $P:=\prod_{i=1}^{n-1}P_i$ gebildet. Sie ist doppelt-stochastisch, erfüllt die entscheidende Voraussetzung des Ergodensatzes, und die Folge P^k, $k=0,1,2\ldots$ hat, wie oben A^k, die Grenzmatrix M. Die Folge $(x_1,\ldots,x_n)P^k$, $k=0,1,2\ldots$ konvergiert wieder gegen die Gleichverteilung $(\bar{x},\ldots,\bar{x})$.
Hier wird auch für $n=2$ das arithmetische Mittel i.a. nur approximiert. Man bestätigt leicht, dass sich die Komponentenfolgen $x_1^{(k)},x_2^{(k)}$ der Vektorfolge $(x_1,x_2)P^k$, $k=1,2,\ldots$ durch

$$x_1^{(k)}=\bar{x}+\frac{1}{2}\cdot(2q_1-1)^k\cdot(x_1-x_2),\qquad x_2^{(k)}=\bar{x}-\frac{1}{2}\cdot(2q_1-1)^k\cdot(x_1-x_2)$$

darstellen lassen. Die Konvergenz zeigt sich unmittelbar an $|2q_1-1|<1$. Für $q_1>1/2$ sind beide Folgen monoton, für $q_1<1/2$ schwanken sie um ihren Grenzwert $\bar{x}$.

3. Die Ungleichung

Ist g eine konvexe Funktion auf dem Intervall $I\subseteq\mathbf{R}$, so bedeutet dies: $(1/2)\cdot(g(x_1)+g(x_2))\geq g(\bar{x})$. Setzt man $f(x_1,x_2)=g(x_1)+g(x_2)$, entsteht eine Funktion auf I^2 mit der Eigenschaft $f(x_1,x_2)\geq f(\bar{x},\bar{x})$.
Allgemeiner werden nun Funktionen von n Variablen $x_1,\ldots,x_n$ betrachtet, die bei den Mittelungen P_i gemäss Nr. 2 entsprechenden Ungleichungen genügen.
Die Abbildungen P_i ändern das arithmetische Mittel der Zahlen $x_1,\ldots,x_n$ nicht. Es wird im weiteren als fest vorgegeben anzusehen sein; um dies zu verdeutlichen, steht jetzt $\hat{x}$ für $\bar{x}$.

Satz. *Gegeben seien $n\in\mathbf{N}$ und ein Intervall $I\subseteq\mathbf{R}$. Zu $\hat{x}\in I$ sei $H_n(\hat{x})$ die Hyperebene $\{(x_1,\ldots,x_n)\,|\,(1/n)\cdot\sum_{i=1}^{n}x_i=\hat{x}\}$. Auf $K(\hat{x}):=I^n\cap H_n(\hat{x})$ sei eine reelle Funktion f mit folgenden Eigenschaften gegeben:*

1. Auf $K(\hat{x})$ und für alle $i=1,\dots,n-1$ gilt

$$f(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_n)\geq f(x_1,\dots,x_{i-1},q_ix_i+q_i'x_{i+1},q_i'x_i+q_ix_{i+1},\dots,x_n),$$

wobei Gleichheit nur für $x_i=x_{i+1}$ zugelassen ist.
2. Die Funktion f ist in $(\hat{x},\dots,\hat{x})$ nach unten halbstetig.
Dann gilt $f(x_1,\dots,x_n)\geq f(\hat{x},\dots,\hat{x})$ mit Gleichheit nur für $x_1=\cdots=x_n=\hat{x}$ ·

Es ist klar, dass «nach unten halbstetig» und «$\geq$» durch «nach oben halbstetig» und «$\leq$» ersetzt werden können.

Beweis: Nach Voraussetzung wird der Funktionswert bei keiner der Mittelungen P_i, die von $(x_1,\dots,x_n)$ zu $(x_1,\dots,x_n)P$ führen, vergrössert. Dasselbe gilt dann auch beim Übergang von $(x_1,\dots,x_n)$ zu $(x_1,\dots,x_n)P^k$, $k=1,2,\dots$ Für die Folge der Funktionswerte $f((x_1,\dots,x_n)P^k)$, $k=0,1,2,\dots$ kann man schliessen:

$$f(x_1,\dots,x_n)\geq f((x_1,\dots,x_n)P^k)\geq \lim_{k\to\infty} f((x_1,\dots,x_n)P^k)$$
$$\geq f(\lim_{k\to\infty}(x_1,\dots,x_n)P^k)=f((x_1,\dots,x_n)M)=f(\hat{x},\dots,\hat{x})\cdot$$

In der letzten Abschätzung ist die Halbstetigkeit ausgenutzt. Gilt Gleichheit, so auch $f(x_1,\dots,x_n)=f((x_1,\dots,x_n)P)$. Aus $(x_1,\dots,x_n)\neq(x_1,\dots,x_n)P$ würde folgen, dass mindestens eine der Mittelungen P_i ein Koordinatenpaar abändern und so den Funktionswert verkleinern müsste. Also ist $(x_1,\dots,x_n)$ Fixvektor von P und damit auch von M. Diese Matrix hat aber $(\hat{x},\dots,\hat{x})$ als einzigen Fixvektor.
In den folgenden Anwendungen wird nur die einfache Mittelung A vorkommen.

4. Spezialisierung auf die Jensensche Ungleichung

Die Funktion $g: I\to \mathbf{R}$ sei stetig und streng konvex, also

$$(g(x_1)+g(x_2))/2\geq g(x_1/2+x_2/2)$$

mit Gleichheit nur für $x_1=x_2$.
Setzt man $f(x_1,\dots,x_n):=\sum_{i=1}^n g(x_i)$, wird für $x_i\neq x_{i+1}$

$$f(x_1,\dots,x_n)-f(x_1,\dots,\ x_i/2\ +x_{i+1}/2,x_i/2+x_{i+1}/2,\dots,x_n)$$
$$=g(x_i)+g(x_{i+1})-2\cdot g(x_i/2+x_{i+1}/2)>0\cdot$$

Der Satz besagt dann:

$$\frac{1}{n}\cdot\sum_{i=1}^n g(x_i)\geq g\left(\frac{1}{n}\cdot\sum_{i=1}^n x_i\right)$$

mit Gleichheit nur für $x_1=\cdots=x_n=(1/n)\cdot\sum_{i=1}^n x_i$.
Dies ist die Jensensche Ungleichung (siehe z. B. [2], S. 51).

5. Beispiel

In [3] wird die Aufgabe gestellt, die Ungleichung

$$\sum_{1\le i<j\le n} x_i x_j/(1-x_i)(1-x_j) \ge \binom{n}{2}/(n-1)^2$$

für $n \ge 3$ und positive Zahlen x_i mit $\sum_{i=1}^n x_i = 1$ zu beweisen; Gleichheit solle nur für $x_1 = \cdots = x_n = 1/n$ eintreten.

Der bewiesene Satz liefert etwas mehr: Für $n \ge 3$ und $\hat{x} \ge 1/n$ gilt auf $[0,1)^n \cap H_n(\hat{x})$

$$f(x_1, \ldots, x_n) := \sum_{1\le i<j\le n} x_i x_j/(1-x_i)(1-x_j) \ge \binom{n}{2} \hat{x}^2/(1-\hat{x})^2$$

mit Gleichheit nur für $x_1 = \cdots = x_n = \hat{x}$.

Zunächst wird ein trivialer Sonderfall erledigt: $x_1 + x_2 = 1$ und $x_3 = \cdots = x_n = 0$. Dann ist $\hat{x} = 1/n$ und

$$f(x_1, \ldots, x_n) = 1 > \frac{n}{2(n-1)} = \binom{n}{2} \hat{x}^2/(1-\hat{x})^2 \,.$$

Um den Satz ins Spiel zu bringen, berechnet man die Differenz D der Funktionswerte aus Bedingung 1; wegen der Symmetrie von f genügt es, dies für $i = 1$ zu tun. Mit $m := x_1/2 + x_2/2$ und $x_1 \neq x_2$ erhält man

$$\begin{aligned} D &= f(x_1, \ldots, x_n) - f(m, m, x_3, \ldots, x_n) \\ &= (x_1 x_2/(1-x_1)(1-x_2) - m^2/(1-m)^2) \\ &+ (x_1/(1-x_1) + x_2/(1-x_2) - 2m/(1-m)) \cdot \sum_{i=3}^n x_i/(1-x_i) \,. \end{aligned}$$

Nach einiger Rechnung ergibt sich für die erste Klammer

$$K_1 = -(x_1 - x_2)^2 (1 - x_1 - x_2)/(1-x_1)(1-x_2)(2-x_1-x_2)^2$$

und für die zweite

$$K_2 = (x_1 - x_2)^2/(1-x_1)(1-x_2)(2-x_1-x_2) \,.$$

Beachtet man $K_2 > 0$ und

$$\sum_{i=3}^n x_i/(1-x_i) \ge \sum_{i=3}^n x_i = n\hat{x} - x_1 - x_2 \ge 1 - x_1 - x_2 \,,$$

wo aber Gleichheit zwischen erstem und letztem Term nur im Sonderfall eintreten kann, folgt ohne Rücksicht auf das Vorzeichen von $(1 - x_1 - x_2)$

$$D = K_1 + K_2 \cdot \sum_{i=3}^{n} x_i/(1-x_i) > K_1 + K_2 \cdot (1-x_1-x_2)$$
$$= (x_1-x_2)^2(1-x_1-x_2)^2/(1-x_1)(1-x_2)(2-x_1-x_2)^2 \geq 0 \cdot$$

Damit ist der Satz anwendbar und die Behauptung bewiesen. Übrigens gilt für $\hat{x} < 1/n$ die Abschätzung nicht mehr, wie die Gleichung $f(n\hat{x}, 0, \ldots, 0) = 0$ zeigt.

6. Bemerkungen

Bekanntlich kann die Ungleichung zwischen arithmetischem und geometrischem Mittel positiver Zahlen bewiesen werden, indem man auf $g(x) = \log x$ die Jensensche Ungleichung anwendet. Zum Vergleich sei $f(x_1, \ldots, x_n) = \prod_{i=1}^{n} x_i$ betrachtet:

$$f(x_1, \ldots, x_n) - f(x_1/2 + x_2/2, x_1/2 + x_2/2, x_3, \ldots, x_n)$$
$$= (x_1 x_2 - ((x_1+x_2)/2)^2) \cdot \prod_{i=3}^{n} x_i < 0 \quad \text{für} \quad x_1 \neq x_2 .$$

Das gibt sofort

$$\left(\prod_{i=1}^{n} x_i\right)^{1/n} \leq \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \cdot$$

Offensichtlich ist es zweckmässig, f als symmetrische Funktion zu wählen; analog zum Beispiel in Nr. 5 etwa

$$f(x_1, \ldots, x_n) = \sum_{1 \leq i < j \leq n} g(x_i) g(x_j) \quad \text{mit einer stetigen Funktion} \quad g: I \to \mathbf{R} \cdot$$

Die Ungleichung $f(x_1, \ldots, x_n) \geq f(\hat{x}, \ldots, \hat{x})$ kann für diesen Funktionstyp umgeformt werden zu

$$\frac{1}{2} \cdot \left(\sum_{i=1}^{n} g(x_i)\right)^2 - \frac{1}{2} \cdot \sum_{i=1}^{n} (g(x_i))^2 \geq \binom{n}{2} \cdot \left(g\left(\frac{1}{n} \cdot \sum_{i=1}^{n} x_i\right)\right)^2 ,$$

also zu einer Ungleichung, die drei Mittel verknüpft. Ist g positiv, konvex und logarithmisch konvex (bzw. konkav und logarithmisch konkav), lässt sich die entscheidende Bedingung (o.E. für $i = 1$ aufgeschrieben)

$$g(x_1) g(x_2) - (g(x_1/2 + x_2/2))^2$$
$$+ (g(x_1) + g(x_2) - 2 \cdot g(x_1/2 + x_2/2)) \cdot \sum_{i=3}^{n} g(x_i) \geq 0 \qquad (\text{bzw.} \leq 0)$$

sofort verifizieren.

Ein typisches Beispiel ist $g(x)=\log x$ mit $x>e$:

$$\left(\log\prod_{i=1}^{n} x_i\right)^2 \le n(n-1)\left(\log\left(\frac{1}{n}\cdot\sum_{i=1}^{n} x_i\right)\right)^2+\sum_{i=1}^{n}(\log x_i)^2\,.$$

Man kann sich überzeugen, dass diese Ungleichung für $x_i>e$ die arithmetisch-geometrische Ungleichung verschärft.

Die Verallgemeinerung dieser Konstruktion von 2-Auswahlen auf k-Auswahlen liegt auf der Hand und führt z. B. für die Gammafunktion zu

$$\sum_{k\text{-Auswahlen aus } n} \Gamma(x_{i_1})\cdot\cdots\cdot\Gamma(x_{i_k})\ge\binom{n}{k}\cdot\left(\Gamma\left(\frac{1}{n}\cdot\sum_{i=1}^{n} x_i\right)\right)^k.$$

Interessanter sind Funktionen wie $g(x)=x/(1-x)$ in Nr. 5, die zwar konvex, aber nicht logarithmisch konvex sind. Zur Anwendung des Satzes benötigt man daher eine zusätzliche Bedingung; im Beispiel war dies $\hat{x}\ge 1/n$.

H. Wellstein, Flensburg, BRD

LITERATURVERZEICHNIS

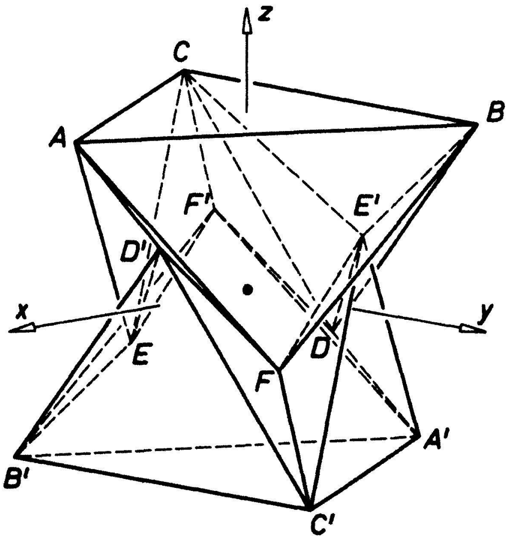
1 A. Rényi: Wahrscheinlichkeitsrechnung. Berlin 1962.
2 G. Pólya und G. Szegö: Aufgaben und Lehrsätze aus der Analysis, Bd. 1, 3. Aufl. Berlin, Göttingen, Heidelberg 1964.
3 S. Gabler: Aufgabe 830, in: El. Math. *34*, 126 (1979).

 0013-6018/82/030079-06$1.50+0.20/0

Kipp-Ikosaeder II[1])

II. Ikosaeder mit dreizähliger Symmetrieachse

5. Betrachtet werden jetzt Ikosaeder mit Mittelpunkt, welche die 120°-Drehungen um die z-Achse vertragen und deren Aussehen an eine Sanduhr erinnert (Fig. 4).



Figur 4. Wackelikosaeder mit dreizähliger Symmetrieachse.

1) Kipp-Ikosaeder I erschien in El. Math. 36, 153–158 (1981).