

Eine Extremaleigenschaft positiv definiter Hermitescher Matrizes

Von BÉLA GYIRES (Debrecen)

Herrn Professor Béla von Szőkefalvi-Nagy zum 50. Geburtstag gewidmet

I.

Der Verfasser hat in seiner Arbeit [1] als Hilfssatz (Lemma 2) im wesentlichen den folgenden Satz formuliert:

Ist A eine positiv definite Hermitesche Matrix n -ter Ordnung, so ist

$$\inf_{X} \frac{1}{n} \operatorname{Sp}(X^*AX)$$

gleich dem arithmetischen bzw. dem geometrischen Mittel der Eigenwerte der Matrix A , je nachdem X die Gesamtheit der unitären bzw. der unimodularen Matrizes n -ter Ordnung durchläuft.

In dieser Arbeit befassen wir uns mit einer Verallgemeinerung dieses Satzes für den Fall, wo die unitäre, bzw. die unimodulare Eigenschaft verallgemeinert wird, und zwar so, daß X die Menge derjenigen Matrizes durchläuft, für welche das Produkt von irgend k Längenquadraten von Zeilenvektoren ($1 \leq k \leq n$) nicht kleiner als eins ausfällt. Im Laufe der Darlegungen werden gewisse aus positiven Zahlen gebildete Größen H_k ($k=1, \dots, n$) eine Rolle spielen, die sich als Verallgemeinerungen des arithmetischen, bzw. des geometrischen Mittels ansehen lassen, da sie für $k=1$ das arithmetische, für $k=n$ hingegen das geometrische Mittel ergeben.

In Punkt II. werden wir zwischen gewissen positiven Zahlen zugeordneten Größen H_k eine Größenordnung nachweisen. In Punkt IV. befassen wir uns mit den Größenordnungsverhältnissen die zwischen den Größen H_k und den Newtonschen Mitteln bestehen. In Punkt III. geben wir die Verallgemeinerung von Lemma 2 aus der bereits erwähnten Arbeit des Verfassers.

II.

Mit den positiven Zahlen $\lambda_1, \dots, \lambda_n$ bilden wir den folgenden Ausdruck:

$$H_k(\lambda_1, \dots, \lambda_n) = H_k = \frac{k}{n} [\lambda_1 \dots \lambda_{k-1} (\lambda_k + \dots + \lambda_n)]^{\frac{1}{k}} \quad (k=1, \dots, n).$$

Offenbar ist H_1 bzw. H_n das arithmetische, bzw. das geometrische Mittel der gege-

benen Zahlen. Diese Größen sind nur für $k=1$ und für $k=n$ symmetrische Funktionen der Veränderlichen $\lambda_1, \dots, \lambda_n$, und nur in diesen beiden Fällen gilt $H_k(\lambda_1, \dots, \lambda_n) = \lambda_k$.

Wir beweisen den folgenden

Satz 1. *Befriedigen die positiven Zahlen $\lambda_1, \dots, \lambda_n$ die Bedingungen*

$$(1) \quad \lambda_v \geq \lambda_{v+1} + \dots + \lambda_n \quad (v=1, 2, \dots, n-1),$$

dann gilt

$$(2) \quad \lambda_1 > H_1 \geq H_2 > \dots > H_n > \lambda_n.$$

BEWEIS. Auf Grund von (1) ist die erste und die letzte Ungleichung der Behauptung trivial.

Die Behauptung wird offenbar bewiesen sein, falls uns der Nachweis von

$$(3) \quad \frac{[\lambda_1 \dots \lambda_{k-1} (\lambda_k + \dots + \lambda_n)]^{\frac{1}{k}}}{[\lambda_1 \dots \lambda_k (\lambda_{k+1} + \dots + \lambda_n)]^{\frac{1}{k-1}}} \geq 1 + \frac{1}{k}$$

gelingt.

Unter Berücksichtigung von

$$\lambda_k (\lambda_{k+1} + \dots + \lambda_n) \leq \frac{1}{4} (\lambda_k + \lambda_{k+1} + \dots + \lambda_n)^2,$$

erhalten wir

$$(4) \quad \frac{[\lambda_1 \dots \lambda_{k-1} (\lambda_k + \dots + \lambda_n)]^{\frac{1}{k}}}{[\lambda_1 \dots \lambda_k (\lambda_{k+1} + \dots + \lambda_n)]^{\frac{1}{k-1}}} \geq \frac{[\lambda_1 \dots \lambda_{k-1} (\lambda_k + \dots + \lambda_n)]^{\frac{1}{k}}}{[\frac{1}{4} \lambda_1 \dots \lambda_{k-1} (\lambda_k + \dots + \lambda_n)^2]^{\frac{1}{k-1}}} =$$

$$= [\lambda_1 \dots \lambda_{k-1} (\lambda_k + \dots + \lambda_n)]^{\frac{1}{k(k-1)}} \left[\frac{4}{\lambda_k + \dots + \lambda_n} \right]^{\frac{1}{k-1}}.$$

Da ferner aus den Bedingungen (1)

$$\lambda_v \geq (k-v)(\lambda_k + \dots + \lambda_n) \quad (v=1, \dots, k-1)$$

folgt, haben wir

$$[\lambda_1 \dots \lambda_{k-1} (\lambda_k + \dots + \lambda_n)]^{\frac{1}{k}} \geq (\lambda_k + \dots + \lambda_n) [(k-1)!]^{\frac{1}{k}}$$

und somit

$$[\lambda_1 \dots \lambda_{k-1} (\lambda_k + \dots + \lambda_n)]^{\frac{1}{k(k-1)}} \geq (\lambda_k + \dots + \lambda_n)^{\frac{1}{k-1}} [(k-1)!]^{\frac{1}{k(k-1)}}.$$

Indem wir dies auf die Ungleichung (4) anwenden erhalten wir

$$(5) \quad \frac{[\lambda_1 \dots \lambda_{k-1} (\lambda_k + \dots + \lambda_n)]^{\frac{1}{k}}}{[\lambda_1 \dots \lambda_k (\lambda_{k+1} + \dots + \lambda_n)]^{\frac{1}{k-1}}} \geq [4^k (k-1)!]^{\frac{1}{k(k-1)}}.$$

Berücksichtigen wir nun die Ungleichung

$$\left(\frac{k+1}{k}\right)^{k(k+1)} = \left[\left(1 + \frac{1}{k}\right)^k\right]^{k+1} < e^{k+1},$$

so ist in Hinblick auf (5) die Ungleichung (3) gewiß erfüllt, falls nur

$$(k-1)! \geq e \left(\frac{e}{4}\right)^k$$

gilt. Diese Ungleichung ist aber für $k \geq 3$ bereits als strenge Ungleichung gültig. Andererseits ist die Ungleichung

$$[4^k (k-1)!]^{1/(k(k+1))} \geq \frac{k+1}{k}$$

für $k=1$ und für $k=2$ erfüllt, und zwar gilt für $k=1$ Gleichheit, für $k=2$ hingegen Ungleichheit. Somit kann Gleichheit nur zwischen H_1 und H_2 bestehen. Gleich sind H_1 und H_2 z. B. für $n=2$ und für $\lambda_1 = \lambda_2$.

Wir berechnen jetzt die Mittel H_k für den Fall, wo $\lambda_n = a > 0$ ist, und gefordert wird, daß in den Bedingungen (1) überall Gleichheit bestehen soll. Offenbar gilt dann

$$\lambda_n = a, \quad \lambda_v = 2^{n-v-1} a \quad (v = 1, \dots, n-1)$$

und

$$H_k = \frac{ka}{n} 2^{n-k} \sqrt[k]{2^{\binom{k-1}{2}}}.$$

In diesem Fall ist wiederum $H_1 = H_2$.

Ist nun

$$\lambda_n = a > 0, \quad \lambda_v = a(1+\theta)^{n-v-1} \theta \quad (v = 1, \dots, n-1),$$

wo θ eine beliebige festbleibende positive Zahl kleiner als eins ist, dann sind die Bedingungen (1) nicht erfüllt. Da in diesem Falle

$$H_k = \frac{ka\theta}{n} (1+\theta)^{n-k} \sqrt[k]{(1+\theta)^{\binom{k-1}{2}}}$$

gilt, ist $H_1 < H_2$, und so sind bereits die Ungleichungen (2) nicht im allgemeinen erfüllt.

Man sieht leicht ein, daß im Falle wo die positiven Zahlen $\lambda_1, \dots, \lambda_n$ die Bedingungen (1) nicht befriedigen, (2) im allgemeinen bereits nicht gültig ist. Es sei nämlich $\lambda_1 = \dots = \lambda_n = \lambda$; dann sind die Bedingungen (1) offenbar nicht erfüllt. Würde nunmehr (2) gelten, dann folgte aus der Gleichheit $H_1 = H_n = \lambda$ die Gleichung

$$H_k = \frac{k}{n} \lambda (n-k+1)^{\frac{1}{k}} = \lambda.$$

bzw.

$$\frac{k}{n} (n-k+1)^{\frac{1}{k}} = 1.$$

Dies kann aber nur für $k=1$ und für $k=n$ sein.

III.

Es bezeichne $\mathfrak{M}_k$ die Menge derjenigen quadratischen Matrizes n -ter Ordnung, für welche eine beliebige Kombination k -ter Ordnung ohne Wiederholung der Längenquadrate der Zeilenvektoren ein Produkt $\cong 1$ ergibt.

Wir beweisen den folgenden

Satz 2. Die Eigenwerte der Diagonalmatrix Λ seien die positiven Zahlen $\lambda_1, \dots, \lambda_n$, und es sei

$$(6) \quad \lambda_1 \cong \lambda_2 \cong \dots \cong \lambda_n > 0.$$

Von den Zahlen $1, \dots, n$ sei k eine beliebige festgewählte. Wir behaupten, daß

$$\inf_{X \in \mathfrak{M}_k} \frac{1}{n} \text{Sp}(X^* \Lambda X) \cong H_k(\lambda_1 \dots \lambda_n)$$

gilt, und daß dabei Gleichheit bzw. Ungleichheit besteht, je nachdem

$$(7) \quad \lambda_{k-1} \cong \lambda_k + \dots + \lambda_n,$$

bzw.

$$(8) \quad \lambda_{k-1} < \lambda_k + \dots + \lambda_n$$

gilt.

Bei der Formulierung dieses Satzes ergibt sich keine Verallgemeinerung, falls wir statt der Diagonalmatrix Λ eine positiv definite Hermitesche Matrix nehmen, deren Eigenwerte $\lambda_1, \dots, \lambda_n$ die entsprechenden Bedingungen befriedigen. Ist nämlich Λ eine solche Matrix und gilt $\Lambda = U^* \Lambda U$, dann ist statt X die Matrix UX in $\mathfrak{M}_k$ enthalten.

BEWEIS VON SATZ 2. Es bezeichne s'_α das Längenquadrat des α -ten Zeilenvektors der Matrix X . Dann gilt

$$\frac{1}{n} \text{Sp}(X^* \Lambda X) = \frac{1}{n} \text{Sp}(\Lambda X X^*) = \frac{1}{n} \sum_{\alpha=1}^n \lambda_\alpha s'_\alpha.$$

Ordnen wir die Zahlen $s'_1, \dots, s'_n$ in die monoton wachsende Reihenfolge $s_1, \dots, s_n$ um, d. h. es gelte

$$(9) \quad 0 < s_1 \leq s_2 \leq \dots \leq s_n,$$

wobei auf Grund der gemachten Annahmen

$$(10) \quad s_1 s_2 \dots s_k \cong 1$$

ist. Mit Rücksicht auf die Reihenfolgen (6) und (10) ([2], Theorem 368) haben wir

$$\frac{1}{n} \text{Sp}(X^* \Lambda X) \cong \frac{1}{n} \sum_{\alpha=1}^n \lambda_\alpha s_\alpha \cong \frac{1}{n} [\lambda_1 s_1 + \dots + \lambda_{k-1} s_{k-1} + \dots + (\lambda_k + \dots + \lambda_n) s_k].$$

Gilt in (10) das Ungleichheitszeichen, so können wir in unserer letzten Formel eine weitere Abnahme erwirken, falls wir die Minderung (9) entsprechend so vornehmen, daß in (10) Gleichheit bestehen soll.

Dementsprechend suchen wir im folgenden die Extremalwerte der Funktion

$$(11) \quad f(x_1, \dots, x_n) = \lambda_1 x_1 + \dots + \lambda_{k-1} x_{k-1} + \frac{\lambda_k + \dots + \lambda_n}{x_1 \dots x_{k-1}}.$$

Da sich aus dem Gleichungssystem

$$(12) \quad \frac{\partial f}{\partial x_v} = \lambda_v - \frac{\lambda_k + \dots + \lambda_n}{x_1 \dots x_{k-1}} \frac{1}{x_v} = 0 \quad (v = 1, \dots, k-1)$$

die Beziehung

$$y_v = x_1 \dots x_{k-1} x_v = \frac{\lambda_k + \dots + \lambda_n}{\lambda_v},$$

und daraus wiederum

$$y_1 \dots y_{k-1} = (x_1 \dots x_{k-1})^k = \frac{(\lambda_k + \dots + \lambda_n)^{k-1}}{\lambda_1 \dots \lambda_{k-1}}$$

folgt, erhalten wir unter Einführung der Bezeichnung

$$q = [\lambda_1 \dots \lambda_{k-1} (\lambda_k + \dots + \lambda_n)]^{\frac{1}{k}}$$

die Gleichheit

$$(13) \quad \frac{1}{x_1 \dots x_{k-1}} = \frac{q}{\lambda_k + \dots + \lambda_n}$$

und somit

$$(14) \quad x_v = s_v = \frac{q}{\lambda_v} \quad (v = 1, \dots, k-1).$$

Indem wir jetzt die v -te Hesse'sche Determinante D_v der Funktion (11) an der Stelle (14) berechnen, erhalten wir

$$D_v = \frac{(\lambda_1 \dots \lambda_v)^2}{q^v} (v+1) > 0 \quad (v = 1, \dots, k-1).$$

Darum, und weil das System (12) genau ein Lösungssystem besitzt, ist die Stelle (14) für die Funktion (11) eine absolute Minimumstelle.

Bei der Konstruktion der Funktion (11) haben wir nur auf $s_1 \dots s_k = 1$ Rücksicht genommen, nicht aber darauf, daß die Anordnung (9) bestehen soll. Auf Grund von (6) und (14) ist es klar, daß diese Anordnung in den Fällen $v = 1, 2, \dots, k-1$ bereits erfüllt ist. Da gemäß (13)

$$(15) \quad s_k = \dots = s_n = \frac{q}{\lambda_k + \dots + \lambda_n}$$

gilt, ist die Anordnung (9) dann und nur dann erfüllt, falls (7) erfüllt ist. Für X wählen wir jetzt eine Diagonalmatrix, deren Eigenwerte die Quadratwurzeln der Größen (14) und (15) sind. Dann gilt

$$(16) \quad \frac{1}{n} \text{Sp}(X^* \Lambda X) = H_k(\lambda_1, \dots, \lambda_n).$$

Damit haben wir die erste Behauptung unseres Satzes bewiesen.

Ist (8) erfüllt, dann gilt $k < n$, und es ist zwar $s_1 \dots s_k = 1$, aber $s_2 \dots s_{k+1} < 1$, d. h. für diejenige Matrix X , in welcher die Längenquadrate der aufeinanderfolgenden Zeilenvektoren gleich den Zahlen $s_1, \dots, s_n$ sind, ist auch jetzt (16) erfüllt, es ist aber X kein Element von $\mathfrak{M}_k$.

Es bezeichne T die Menge derjenigen Punkte des $k-1$ -dimensionalen Raumes, für welche

$$0 < x_1 \leq x_2 \leq \dots \leq x_{k-1}, \quad x_1 \dots x_{k-2} x_{k-1}^2 \leq 1$$

gilt; dann ist die Gesamtheit der Koordinaten eines beliebigen Punktes von T ein Element von $\mathfrak{M}_k$, und umgekehrt. Bilden wir jetzt die Abschließung der Menge T dadurch, daß wir auch die Relation $0 \leq x_1$ zulassen. Derjenige Punkt P des $k-1$ -dimensionalen Raumes, mit Koordinaten (14), welche die Ungleichung (8) befriedigen, ist in Hinblick auf die sich aus (8) ergebenden Relationen

$$s_1 \leq s_2 \leq \dots \leq s_{k-1}, \quad s_1 \dots s_{k-2} s_{k-1}^2 > 1$$

kein Element von T . So ist der Abstand dieses Punktes von der Menge T eine positive Zahl. Mit anderen Worten: besitzt dieser Punkt P eine Umgebung, welche nicht zu T gehört. Da nun aber die Funktion (11) an der Stelle (14) ein absolutes Minimum hat, und da diese Funktion auf der Menge T beim unbeschränkten Wachsen der Veränderlichen x_v selbst unbeschränkt wächst, kann sich auf T der Wertevorrat der Funktion (11) dem für (14) angenommenen absoluten Minimum nicht unbeschränkt nähern. Damit haben wir auch für die zweite Behauptung unseres Satzes den Nachweis erbracht.

In Satz 2. ergibt sich für $k=1$ weder aus (7) noch aus (8) eine Bedingung. Im Falle $k=1$ folgt die Ungleichung (7) bereits aus (6). Somit gelten die folgenden beiden Sätze:

Satz 3. Sind die Eigenwerte $\lambda_1, \dots, \lambda_n$ der Diagonalmatrix Λ positive Zahlen, dann gilt

$$\inf_{X \in \mathfrak{M}_1} \frac{1}{n} \operatorname{Sp}(X^* \Lambda X) = \frac{1}{n} [\lambda_1 + \dots + \lambda_n].$$

Satz 4. Sind die Eigenwerte $\lambda_1, \dots, \lambda_n$ der Diagonalmatrix Λ positiv, dann gilt

$$\inf_{X \in \mathfrak{M}_n} \frac{1}{n} \operatorname{Sp}(X^* \Lambda X) = (\lambda_1 \dots \lambda_n)^{\frac{1}{n}}$$

Setzen wir bezüglich der Matrizes X nur Unimodularität voraus, so bleibt Satz 4. immer noch gültig. Einerseits gilt nämlich, wenn wir die Gesamtheit der unimodularen Matrizes der Ordnung n mit $\mathfrak{M}_n^{(u)}$ bezeichnen, offenbar $\mathfrak{M}_n^{(u)} \subset \mathfrak{M}_n$. Andererseits ist gemäß Bedingung (6) diejenige Diagonalmatrix, deren Eigenwerte die aus (14) und aus (15) für $k=n$ erhältlichen Werte sind, ebenfalls unimodular. Der sich so ergebende Satz ist der bereits in der Arbeit [1] des Verfassers auftretende, und in der Einleitung erwähnte Lemma.

Die k -te derivierte Matrix der Matrix A der Ordnung n ([4], 5. 03) sei mit $R_k(A)$ bezeichnet. Es sei $\mathfrak{M}_k$ die Menge derjenigen Matrizes X der Ordnung n , für welche die Länge der Zeilenvektoren der Matrizes $R_k(X)$ nicht kleiner als eins ist. Wir zeigen, daß $\mathfrak{M}_k \subset \mathfrak{M}_k$ gilt.

Es bezeichne s_v das Längenquadrat des v -ten Zeilenvektors der Matrix $X = (a_{\mu k}) \in \mathfrak{H}_k$. Für $1 \leq \alpha_1 < \dots < \alpha_k \leq n$ und

$$A = \begin{pmatrix} a_{\alpha_1 1} & \dots & a_{\alpha_1 n} \\ \vdots & \dots & \vdots \\ a_{\alpha_k 1} & \dots & a_{\alpha_k n} \end{pmatrix}$$

ist nun $\text{Det } AA^*$ nach der Determinantenidentität von Binet—Cauchy ([3], 8) das Längenquadrat eines Zeilenvektors von $R_k(X)$. Nach der Bedingung, welche den Matrizes der Menge $\mathfrak{H}_k$ auferlegt wurde, und nach dem Satz von Hadamard ([2], Theorem 30) gilt nun

$$1 \leq \text{Det } AA^* \leq s_{\alpha_1} \dots s_{\alpha_k}$$

und dies ergibt bereits unsere Behauptung.

Nehmen wir jetzt noch darauf Rücksicht, daß $X \in \mathfrak{H}_k$ gilt, falls X eine Diagonalmatrix ist, deren Eigenwerte die den Bedingungen (6) und (7) genügenden Zahlen (14) und (15) sind. Auf Grund des gesagten ergibt sich nun als Korollar zu Satz 2. der folgende

Satz 5. Ist Λ eine Diagonalmatrix deren Eigenwerte der Bedingungen (6) genügende Zahlen sind, dann gilt

$$\inf_{X \in \mathfrak{H}_k} \frac{1}{n} \text{Sp}(X^* \Lambda X) \cong H_k(\lambda_1, \dots, \lambda_n)$$

und es besteht Gleichheit bzw. Ungleichheit, je nachdem (7) oder (8) erfüllt ist.

IV.

Wir bezeichnen das k -te Newtonsche Mittel ($k=1, 2, \dots, n$) der positiven Zahlen $\lambda_1, \dots, \lambda_n$ mit $C_k(\lambda_1, \dots, \lambda_n)$. Also gilt

$$C_k(\lambda_1, \dots, \lambda_n) = C_k = \left[\frac{1}{\binom{n}{k}} \sum \lambda_{i_1} \dots \lambda_{i_k} \right]^{\frac{1}{k}}, \quad 1 \leq i_1 < \dots < i_k \leq n.$$

Bekanntlich hat man ([2], Theorem 52)

$$(17) \quad C_1 \cong C_2 \cong \dots \cong C_n,$$

und Gleichheit gilt dann und nur dann, falls $\lambda_1 = \dots = \lambda_n$ ist. Offenbar bedeutet C_1 bzw. C_n das arithmetische bzw. das geometrische Mittel der gegebenen Zahlen. Wir beweisen den folgenden

Satz 6. Befriedigen die Zahlen $\lambda_1, \dots, \lambda_n$ die Bedingungen (6) und (7), dann gilt

$$(18) \quad H_k(\lambda_1, \dots, \lambda_n) \cong C_k(\lambda_1, \dots, \lambda_n).$$

Dabei kann Gleichheit nur in den Fällen $k=1$ und $k=n$ gelten.

BEWEIS. Es sei A diejenige Diagonalmatrix, deren Eigenwerte die Zahlen $\lambda_1, \dots, \lambda_n$ sind. Dann gilt für die Wertsysteme (14) und (15) und auf Grund von (17)

$$(19) \quad H_k(\lambda_1, \dots, \lambda_n) = C_1(\lambda_1 s_1, \dots, \lambda_n s_n) \cong C_k(\lambda_1 s_1, \dots, \lambda_n s_n).$$

Da nun das Produkt von irgend k der Zahlen $s_1, \dots, s_n$ nicht kleiner als eins ist, haben wir

$$(20) \quad C_k(\lambda_1 s_1, \dots, \lambda_n s_n) \cong C_k(\lambda_1, \dots, \lambda_n)$$

und dies ergibt, zusammen mit der Ungleichung (19) bereits unsere Behauptung (18).

Nun gilt in (20) Gleichheit nur im Falle wo das Produkt von irgend k der Zahlen $s_1, \dots, s_n$ immer gleich eins ist, und für $k < n$ ist dies mit $s_1 = \dots = s_n = 1$ äquivalent; andererseits gilt in (19) für $k > 1$ Gleichheit nur dann, wenn $\lambda_1 s_1 = \dots = \lambda_n s_n$ ist. Dementsprechend haben wir in (18) im Falle $1 < k < n$ Gleichheit dann und nur dann falls $\lambda_1 = \dots = \lambda_n$ ist. Diese Bedingung ist aber für $1 < k < n$ mit der Bedingung (7) nicht verträglich. Somit kann in (18) in den Fällen $1 < k < n$ nur Ungleichheit bestehen.

Literatur

- [1] B. GYIRES, A generalization of a theorem of Szegő, *Publ. Math. Inst. Hung. Acad. Sci.* 7 Ser. A, (1962), 43–51.
- [2] G. H. HARDY—J. E. LITTLEWOOD—G. PÓLYA, *Inequalities*, Cambridge, 1952.
- [3] F. R. GANTMACHER, *Matrizenrechnung*, I. Berlin, 1958.
- [4] J. H. M. WEDDERNBURN, *Lectures on matrices*, New York, 1934.

(Eingegangen am 16. November 1963.)