Heinrich Heine Universität Düsseldorf

MATHEMATISCHES INSTITUT
PROF. DR. FLORIAN JARRE

20.11.2026

Exkurs zu Mittelwerten von Matrizen

In Anlehnung an R. Bhatia: Positive Definite Matrices (2007)

Mittelwerte auf den reellen Zahlen.
Für $x, y \in \mathbb{R}$ mit $x, y > 0$ erfüllen das

- arithmetische Mittel $m_a(x,y) := \frac{x+y}{2}$
- geometrische Mittel $m_g(x,y) := (xy)^{1/2}$
- harmonische Mittel $m_h(x,y) := 2(x^{-1}+y^{-1})^{-1}$

die Ungleichung

$$m_h(x,y) \le m_g(x,y) \le m_a(x,y). \tag{1}$$

Beispiel.
Schaltet man zwei Leitungen mit elektrischen Widerständen r_1 und r_2 parallel, so erfüllt der Gesamtwiderstand r die Beziehung $r^{-1} = r_1^{-1} + r_2^{-1}$, d.h. $r = \frac{1}{2} m_h(r_1, r_2)$.
Das geometrische Mittel wird im weiteren Verlauf der Vorlesung noch eine Rolle spielen.
Weitere Mittelwerte sind in dem Buch von Bhatia angegeben.

Eigenschaften. Für $0 < x \le y$ erfüllen die obigen Mittelwerte "m" auch die Beziehungen

$$x \le m(x,y) = m(y,x) \le y \tag{2}$$

und insbesondere liegt $m(x,y)$ im selben "Raum" wie x und y. Ferner ist $m(x,y)$ als Funktion von x monoton wachsend und es gilt für reelles $\alpha \neq 0$ (d.h. für $\alpha^2 > 0$):

$$(*) \qquad m(\alpha^2 x, \alpha^2 y) = \alpha^2 m(x,y).$$

Mittelwerte auf den symmetrischen positiv definiten Matrizen.
Für positiv definite Diagonalmatrizen X, Y lässt sich die Definition von $m_a,\ m_g, m_h$ elementweise auf die Diagonaleinträge anwenden. Ersetzt man nun die Ordnung "$\ge$" auf $\mathbb{R}$ durch die semidefinite Halbordnung "$\succeq$" auf dem Raum $\mathcal{S}^n$ der symmetrischen $n \times n$-Matrizen (die sogenannte Löwner Halbordnung), so bleiben die obigen Eigenschaften (1) und (2) erhalten.

Die Beziehung $(*)$ soll nun im Folgenden genauer betrachtet werden: Die Zahlen α^2 (mit $\alpha \neq 0$) zeichnen sich durch die Verträglichkeit mit der Ordnung "$\ge$" auf $\mathbb{R}$ aus,

$$0 < x \le y \qquad \Longleftrightarrow \qquad 0 < \alpha^2 x \le \alpha^2 y.$$

Auf dem Raum der symmetrischen Matrizen gilt diese Ordnungserhaltung bezüglich der Löwner Halbordnung auch für Kongruenztransformationen mit invertierbaren Matrizen S:

$$0 \prec X \preceq Y \qquad \Longleftrightarrow \qquad 0 \prec S^T X S \preceq S^T Y S \tag{3}$$

(wohingegen S^TSX im Allgemeinen nicht im Raum der symmetrischen Matrizen läge und die Löwner Halbordnung daher nicht auf S^TSX anwendbar wäre). Es liegt daher nahe, für die Verallgemeinerung der Mittelwerte von Diagonalmatrizen auf beliebige symmetrische positiv definite Matrizen anstelle von $(*)$ die schärfere Forderung der Invarianz unter Kongruenztransformationen

$$(**) \qquad m(S^TXS, S^TYS) = S^Tm(X,Y)S \qquad \text{für } X, Y \succ 0 \text{ und alle invertierbaren } S$$

zu stellen.

Für m_a und m_h lässt sich diese Forderung auch direkt verifizieren:

$$S^Tm_a(X,Y)S = \tfrac{1}{2}S^T(X+Y)S = \tfrac{1}{2}(S^TXS + S^TYS) = m_a(S^TXS, S^TYS),$$

und

$$S^Tm_h(X,Y)S = 2S^T(X^{-1}+Y^{-1})^{-1}S = 2(S^{-1}(X^{-1}+Y^{-1})S^{-T})^{-1}$$
$$= 2(S^{-1}X^{-1}S^{-T} + S^{-1}Y^{-1}S^{-T})^{-1} = 2((S^TXS)^{-1} + (S^TYS)^{-1})^{-1} = m_h(S^TXS, S^TYS).$$

In diesem Zusammenhang ist auch die Gleichung

$$\begin{aligned} m_h(X,Y) &= 2(X^{-1}+Y^{-1})^{-1} = 2(X^{-1}(X+Y)Y^{-1})^{-1} = 2(Y(X+Y)^{-1}X) \\ &= 2((X+Y)(X+Y)^{-1}X - X(X+Y)^{-1}X) = 2(X - X(X+Y)^{-1}X) \end{aligned} \quad (4)$$

interessant, die wegen $m_h(X,Y) = m_h(Y,X)$ auch gilt wenn die Argumente X und Y vertauscht werden.

Bestimmung der Mittelwerte anhand $(**)$ und Diagonalmatrizen:
Es zeigt sich nun, *falls* die Forderung $(**)$ erfüllbar ist, so legt sie auch die Erweiterung der Mittelwerte von positiv definiten Diagonalmatrizen auf symmetrische positiv definite Matrizen eindeutig fest:

Seien dazu $X, Y \succ 0$ beliebig gegeben und

$$X^{-1/2}YX^{-1/2} = UDU^T$$

eine Eigenwertzerlegung von $X^{-1/2}YX^{-1/2}$ mit einer Diagonalmatrix D und einer orthogonalen Matrix U. Setzt man dann $S^T := U^TX^{-1/2}$ so folgt, wegen

$$S^TXS = U^TX^{-1/2}XX^{-1/2}U = I$$

und

$$S^TYS = U^TX^{-1/2}YX^{-1/2}U = U^T(UDU^T)U = D$$

dass die Matrix S sowohl X als auch Y kongruent auf Diagonalgestalt transformiert und mit $(**)$ ein Wert

$$m(X,Y) = m(S^{-T}IS^{-1}, S^{-T}DS^{-1}) \overbrace{=}^{(**)} S^{-T}m(I,D)S^{-1}$$

aus der Festlegung des Mittels bei Diagonalmatrizen folgt. Zu zeigen ist dann aber noch, dass der betrachtete Mittelwert m so gegeben ist, dass das Ergebnis rechts unabhängig von der Wahl der obigen Herleitung ist. (Man hätte z.B. auch Y auf die Einheitsmatrix transformieren können oder anstelle von $X^{-1/2}$ den inversen Cholesky-Faktor nutzen können.) Für m_a und m_h wurde oben schon gezeigt, dass $(**)$ immer erfüllt ist, und $m(X,Y)$ damit eindeutig gegeben ist.

Der Fall m_g:
Für Diagonalmatrizen X, Y ist $m_g(X,Y) = (XY)^{1/2}$. Für allgemeine positiv definite Matrizen ist XY ähnlich zur Matrix $X^{-1/2}XYX^{1/2} = X^{1/2}YX^{1/2} \succ 0$, hat also n positive Eigenwerte mit zugehörigen Eigenvektoren, sodass sich die Wurzel $(XY)^{1/2}$ unter Beibehaltung der Eigenvektoren von XY und Übergang zu den Wurzeln der Eigenwerte wieder bilden lässt mit $((XY)^{1/2})^2 = XY$. Allerdings ist $(XY)^{1/2}$ in der Regel nicht symmetrisch, liegt also nicht im selben Raum wie X, Y und erfüllt daher z.B. auch nicht die Eigenschaft (2) der Mittelwerte auf den reellen Zahlen.

Bei positiv definiten Diagonalmatrizen lässt sich die Beziehung $A = (XY)^{1/2}$ für das geometrische Mittel auch in der äquivalenten symmetrischen Form $AX^{-1}A = Y$ mit $A \succ 0$ schreiben. Im Gegensatz zu $A = (XY)^{1/2}$ hat diese Gleichung für gegebene symmetrische Matrizen $X, Y \succ 0$ genau eine symmetrische positiv definite Lösung A, denn es ist $AX^{-1}A = Y$ genau dann wenn

$$\begin{aligned} & X^{-1/2}AX^{-1}AX^{-1/2} = X^{-1/2}YX^{-1/2} \\ \iff\ & (X^{-1/2}AX^{-1/2})^2 = X^{-1/2}YX^{-1/2} \\ \iff\ & X^{-1/2}AX^{-1/2} = (X^{-1/2}YX^{-1/2})^{1/2} \\ \iff\ & A = X^{1/2}(X^{-1/2}YX^{-1/2})^{1/2}X^{1/2}. \end{aligned} \tag{5}$$

(Die Eindeutigkeit in der vorletzten Äquivalenz folgt weil mit $A \succ 0$ auch $X^{-1/2}AX^{-1/2} \succ 0$ gilt.) Die positiv definite Lösung ist also eindeutig bestimmt.

Invertiert man nun beide Seiten von $AX^{-1}A = Y$, so ergibt sich die äquivalente Gleichung $A^{-1}XA^{-1} = Y^{-1} \iff X = AY^{-1}A$, was dann (unter Vertauschung von X und Y in (5)) die eindeutige Lösung hat $A = Y^{1/2}(Y^{-1/2}XY^{-1/2})^{1/2}Y^{1/2}$.

Also ist die symmetrische positive definite Matrix

$$A = X^{1/2}(X^{-1/2}YX^{-1/2})^{1/2}X^{1/2} = Y^{1/2}(Y^{-1/2}XY^{-1/2})^{1/2}Y^{1/2}$$

auch symmetrisch bezüglich der Argumente X und Y und sie ist die eindeutige Lösung der charakteristischen Gleichung $AX^{-1}A = Y$.
Dies legt die Definition

$$m_g(X,Y) := X^{1/2}(X^{-1/2}YX^{-1/2})^{1/2}X^{1/2}$$

nahe.
Dass diese Definition nun auch $(**)$ erfüllt, ergibt sich aus der charakteristischen Gleichung $AX^{-1}A = Y$ beim Übergang von $A = m_g(X,Y)$ zu $\tilde{A} = m_g(S^TXS, S^TYS)$. Die charakteristische Gleichung für $\tilde{A}$ ist dann

$$\begin{aligned} & \tilde{A}(S^TXS)^{-1}\tilde{A} = S^TYS \\ \iff\ & S^{-T}\tilde{A}(S^TXS)^{-1}\tilde{A}S^{-1} = Y \\ \iff\ & (S^{-T}\tilde{A}S^{-1})X^{-1}(S^{-T}\tilde{A}S^{-1}) = Y. \end{aligned}$$

Also ist $S^{-T}\tilde{A}S^{-1}$ auch die eindeutige positiv definite Lösung von $AX^{-1}A = Y$ d.h. es gilt $S^{-T}\tilde{A}S^{-1} = m_g(X,Y)$ bzw. $\tilde{A} = S^T m_g(X,Y)S$ wie in $(**)$ gefordert.

Die Beziehung $(**)$ zeigt nun auch dass die Anordnung der Mittelwerte erhalten bleibt. Es gilt

$$m_h(X,Y) \preceq m_g(X,Y) \preceq m_a(X,Y)$$

zunächst für positiv definite Diagonalmatrizen X, Y und wegen $(**)$ und (3) dann auch für beliebige symmetrische positiv definite Matrizen X, Y.

Und falls $0 \prec X \preceq Y$, so gilt auch

$$X \preceq m_h(X,Y) \preceq m_g(X,Y) \preceq m_a(X,Y) \preceq Y.$$

Die Ungleichung ganz rechts folgt direkt aus $X \preceq Y$ und ganz links folgt aus $X+Y \succeq 2X$ für die Inversen $2(X+Y)^{-1} \preceq X^{-1}$. Wird diese Ungleichung von links und rechts mit X multipliziert, so ergibt sich $2X(X+Y)^{-1}X \preceq X$. Und dann folgt $X = 2X - X \preceq 2X - 2X(X+Y)^{-1}X = m_h(X,Y)$ wegen (4).