Zum Logarithmus einer Matrix

Von Hans Richter in Haltingen/Baden

1. Unter $\ln z$ für komplexes z verstehe man im folgenden stets den Wert auf der längs der negativ-reellen Achse aufgeschnittenen z-Ebene mit $\ln 1 = 0$. $\mathfrak{B}$ sei der Bereich aller quadratischen Matrizen A des Grades n mit komplexen Koeffizienten, deren Eigenwerte nicht Null oder negativ reell sind. Hat A lauter verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_n$, so schreibe man mit geeignetem C:

$$A = C \cdot \begin{pmatrix} \lambda_1 . & & 0 \\ & \cdot & \\ 0 & & \cdot \lambda_n \end{pmatrix} \cdot C^{-1}$$

und setze

$$\ln A = C \cdot \begin{pmatrix} \ln \lambda_1 . & & 0 \\ & \cdot & \\ 0 & & \cdot \ln \lambda_n \end{pmatrix} \cdot C^{-1}. \tag{1.1}$$

Hat A teilweise zusammenfallende Eigenwerte, so werde $\ln A$ als $\lim_{t \to 0} \ln(A(t))$ definiert mit geeigneten $A(t)$, die getrennte Eigenwerte haben, und für die $\lim_{t \to 0} A(t) = A$ gilt. Aus der allgemeinen Theorie[1]) der Matrixfunktionen folgen die Eigenschaften:

Für alle

$$A \in \mathfrak{B} \text{ ist } \ln A \tag{1.a}$$

eindeutig definiert: *Hauptzweig* von $\ln A$.

$$e^{\ln A} = \sum_{\nu=0}^{\infty} \frac{1}{\nu!} (\ln A)^{\nu} = A\,. \tag{1.b}$$

Für reelles

$$A \text{ ist auch } \ln A \text{ reell.} \tag{1.c}$$

$\ln A$ ist die einzige für alle reellen $A \in \mathfrak{B}$ stetige und reelle Lösung von $e^X = A$.

Ist das Differential dA mit A vertauschbar, so ist

$$\ln(A + dA) = \ln A + A^{-1} \cdot dA + O((dA)^2)\,. \tag{1.d}$$

---

[1]) H. Richter, Über Matrixfunktionen. Math. Ann. **122** (1950) S. 16—34.

Die Verwendung von $\ln A$ hat sich in der allgemeinen Theorie der endlichen Verformungen als zweckmäßig erwiesen; bei speziellen Aufgaben dagegen macht es sich außerordentlich störend bemerkbar, daß man im Gegensatz zu e^A nicht auch $\ln A$ durch einen einfachen Ausdruck darstellen kann, sondern zu seiner Bildung erst die Eigenwerte von A zu berechnen hat. Hier sagt nun die allgemeine Theorie, daß es eine Darstellung der folgenden Gestalt gibt:

$$\ln A = \sum_{\nu=0}^{n-1} c_\nu A^\nu, \tag{1.2}$$

wo die c_ν affine Invarianten von A sind. Während nun im allgemeinen die c_ν nur als symmetrische Funktionen der neuen Eigenwerte bekannt sind, wollen wir in dieser Arbeit zeigen, wie man speziell bei $\ln A$ zu geschlossenen Ausdrücken für die c_ν kommen kann, die nur von den Koeffizienten der charakteristischen Gleichung von A abhängen.

2. Es sei nun $A \in \mathfrak{B}$. Wir bilden dann die von dem Parameter t abhängige Matrix

$$A(t) = E + t\cdot(A - E)\,. \tag{2.1}$$

$A(t)$ hat die Eigenwerte $1 + t\cdot(\lambda_\nu - 1)$. Es ist daher für $0 \leqq t \leqq 1$ auch $A(t) \in \mathfrak{B}$. Somit existiert

$$X(t) = \ln A(t) \text{ mit } X(0) = 0 \text{ und } X(1) = \ln A\,. \tag{2.2}$$

Es ist nach (1.d) weiter:

$$X(t + dt) - X(t) = A^{-1}(t)\cdot(A - E)\cdot dt + O((dt)^2)\,,$$

woraus folgt:

$$X'(t) = A^{-1}(t)\cdot(A - E)\,. \tag{2.3}$$

Wegen (2.2) folgt hieraus sofort

$$\ln A = \int_0^1 \frac{A - E}{A(t)}\,dt\,. \tag{2.4}$$

Es kommt also nur darauf an, $A^{-1}(t)$ in geeigneter Weise darzustellen.

3. Wir bilden nun die von dem Parameter u abhängige Hilfsmatrix:

$$P(u) = \sum_{k=0}^{n-1} g_k(u)\cdot A^k \text{ mit } g_k(u) = \sum_{r=0}^{n-1-k} a_{n-1-k-r}\cdot u^r\,, \tag{3.1}$$

wo die a_ν die Koeffizienten des charakteristischen Polynoms

$$\varphi(z) \equiv |\,z\,E - A\,| \equiv \sum_{\nu=0}^{n-1} a_\nu\, z^{n-\nu} \tag{3.2}$$

von A sind.

Zunächst bestimmen wir die Spur von $P(u)$, die wir durch geschweifte Klammern kennzeichnen:

$$\{P(u)\} = \sum_{k=0}^{n-1} \sum_{r=0}^{n-1-k} a_{n-1-k-r} \cdot u^r \cdot s_k ,$$

wo $s_k = \{A^k\}$ die k-te Potenzsumme der Eigenwerte von A ist. Mit Hilfe der Newtonschen Rekursionsformel erhalten wir wegen $s_0 = n$:

$$\{P(u)\} = \sum_{r=0}^{n-1} u^r \cdot \sum_{k=0}^{n-1-r} a_{n-1-r-k} \cdot s_k = \sum_{r=0}^{n-1} u^r \cdot (r+1) \cdot a_{n-1-r}$$

oder

$$\{P(u)\} = \varphi'(u) . \tag{3.3}$$

Weiter gilt wegen $\varphi(A) = 0$, und wenn $g_{-1}(u) \equiv 0$ gesetzt wird:

$$P(u) \cdot (A - u\,E) = \sum_{k=0}^{n-1} A^k \cdot [g_{k-1}(u) - u\, g_k(u) - g_{n-1}(u) \cdot a_{n-k}]$$

oder bei Einsetzen der g_k aus (3.1):

$$P(u) \cdot (A - u\,E) = - E \cdot \varphi(u) . \tag{3.4}$$

4. In (3.4) setzen wir speziell $u = \frac{t-1}{t}$ mit $0 \leqq t \leqq 1$. Dann erhalten wir wegen $A - u\,E = \frac{1}{t} \cdot A(t)$:

$$P\left(\frac{t-1}{t}\right) \cdot \frac{A(t)}{t} = - E \cdot \varphi\left(\frac{t-1}{t}\right)$$

und wegen der Invertierbarkeit von $A(t)$:

$$A^{-1}(t) = - \frac{P\left(\frac{t-1}{t}\right)}{t \cdot \varphi\left(\frac{t-1}{t}\right)} . \tag{4.1}$$

Es wird dann schließlich wegen

$$(A - E) \cdot A^{-1}(t) = \frac{E - A^{-1}(t)}{t}$$

gemäß (2.4):

$$\ln A = \int_0^1 \left[\frac{E}{t} + \frac{P\left(\frac{t-1}{t}\right)}{t^2 \cdot \varphi\left(\frac{t-1}{t}\right)}\right] dt . \tag{4.2}$$

5. Der so für $\ln A$ erhaltene Ausdruck ist noch ungünstig, da wir das Integral nicht in konvergente Summanden aufspalten können. Um dies zu erreichen, bilden wir zunächst die Spur. Mit Hilfe von (3.3) erhalten wir:

$$\{\ln A\} = \int_0^1 \left[\frac{n}{t} + \frac{\varphi'\left(\frac{t-1}{t}\right)}{t^2 \cdot \varphi\left(\frac{t-1}{t}\right)}\right] dt = \int_0^1 \frac{\frac{d}{dt}\left[(-t)^n \cdot \varphi\left(\frac{t-1}{t}\right)\right]}{(-t)^n \cdot \varphi\left(\frac{t-1}{t}\right)}\, dt ,$$

oder wegen $(-t)^n \varphi\left(\frac{t-1}{t}\right) = \sum_\nu a_\nu (1-t)^{n-\nu} \cdot (-t)^\nu$ mit den Grenzwerten 1 bei $t = 0$ und $(-1)^n \cdot a_n$ bei $t = 1$:

$$\{\ln A\} = \ln |A| + 2\pi l i, \tag{5.1}$$

eine Beziehung, die auch unmittelbar aus der Definition von $\ln A$ folgt. l ist dabei die Anzahl der Überschreitungen von $(-t)^n \varphi\left(\frac{t-1}{t}\right)$ über die negativ reelle Achse von oben nach unten, wenn t von 0 bis 1 wächst. Bei reellem A ist stets $l = 0$.

Nachdem die Spur von $\ln A$ bekannt ist, genügt es, den Deviator von $\ln A$ zu berechnen, wobei wir denselben für eine beliebige Matrix Q durch

$$\widetilde{Q} = Q - \frac{1}{n}\{Q\} \cdot E$$

bezeichnen. Wir finden dann aus (4.2)

$$\widetilde{\ln A} = \int\limits_0^1 \frac{\widetilde{P}\left(\frac{t-1}{t}\right)}{t^2 \cdot \varphi\left(\frac{t-1}{t}\right)} dt = \int\limits_{-\infty}^0 \frac{\widetilde{P}(u)}{\varphi(u)} du,$$

und damit endlich unter Verwendung von (3.1) und (5.1):

$$\left.\begin{aligned} \ln A &= \frac{\ln |A| + 2\pi l i}{n} \cdot E + \sum_{k=1}^{n-1} c_k \cdot \widetilde{A}^k \\ \text{mit} \quad c_k &= \int\limits_{-\infty}^0 \frac{g_k(u)}{\varphi(u)} du, \quad g_k(u) = \sum_{r=0}^{n-1-k} a_{n-1-k-r} \cdot u^r. \end{aligned}\right\} \tag{5.2}$$

In der Tat ist nämlich $g_k(u)$ ein Polynom vom Grade $n-1-k$, so daß die für c_k angegebenen Integrale bei $k \geqslant 1$ sicher konvergieren. Setzt man $A^* = A \cdot \lambda$ mit positivem λ, so existiert auch $\ln A^*$. Es ist dann $a_\nu^* = \lambda^\nu \cdot a_\nu$ und daher

$$g_k^*(u) = \lambda^{n-1-k} \cdot g_k\left(\frac{u}{\lambda}\right) \text{ und } \varphi^*(u) = \lambda^n \cdot \varphi\left(\frac{u}{\lambda}\right).$$

Wir erhalten dann

$$c_k^* = \lambda^{-k} \cdot c_k.$$

(5.2) zeigt dann unmittelbar:

$$\ln A^* = \ln A + E \cdot \ln \lambda.$$

Die für die Anwendungen besonders wichtigen Fälle der reellen Matrizen der Grade $n = 2$ und $n = 3$ seien besonders angemerkt:

$$\left.\begin{aligned} \ln A &= \frac{\ln |A|}{2} \cdot E + \widetilde{A} \cdot \int\limits_{-\infty}^0 \frac{du}{u^2 + a_1 u + a_2} \quad \text{bei } n = 2 \\ \ln A &= \frac{\ln |A|}{3} \cdot E + \widetilde{A} \cdot \int\limits_{-\infty}^0 \frac{(a_1 + u)\, du}{u^3 + a_1 u^2 + a_2 u + a_3} + \widetilde{A}^2 \cdot \int\limits_{-\infty}^0 \frac{du}{u^3 + a_1 u^2 + a_2 u + a_3} \\ &\quad \text{bei } n = 3. \end{aligned}\right\} \tag{5.3}$$

(Eingegangen am 3. 2. 1950)