Eine Matrixmethode zur Lösung linearer Differenzengleichungen

Von Á. PETHŐ (Budapest)

In einer früheren Arbeit [1] haben wir über eine Anwendung des Matrizenkalküls zur Lösung der Anfangswertaufgaben (im allgemeinen partieller) linearer Differenzengleichungen in dem Falle berichtet, wo die Koeffizienten von *einer* Veränderlichen unabhängig sind. Diese Methode ließ — was ihre Ausführung anbetrifft — die wesentlichen Züge derjenigen der erzeugenden Funktionen [2] erkennen; es bestand jedoch ein prinzipieller Unterschied der letzteren gegenüber: bei unserem Kalkül traten keine Schwierigkeiten der Konvergenz auf. Derselbe ist wohl — wie es in der vorliegenden (und einer nachfolgenden) Arbeit dargetan wird — als eine Matrixdarstellung der Operatorenrechnung von J. MIKUSIŃSKY [3] anzusehen, wenn diese auf die Lösung von Differenzengleichungen angewendet wird. Diesmal wollen wir die Matrixdarstellung zur Lösung der gewöhnlichen linearen Differenzengleichungen mit konstanten Koeffizienten darlegen, wobei es sich herausstellen wird, daß unser Verfahren sowohl mit der klassischen Methode der erzeugenden Funktionen, als auch mit der von MIKUSIŃSKI *formal* vollständig übereinstimmt; die folgenden Darlegungen können mithin als eine *elementare* Begründung der Theorie der Operatorenmethode zur Lösung gewöhnlicher linearer Differenzengleichungen gelten.

§ 1. Über den Ring der Zahlenfolgen in Matrixdarstellung

1. Wir betrachten die unendlichen Matrizen (mit komplexen Elementen)

$$\Lambda = [\delta_{|i-j|,\, i-j}\lambda_{i-j}]; \qquad i, j = 0, 1, \ldots, \tag{1.1}$$

wo $\delta_{\mu,\nu}$ das Kroneckersche Symbol bedeutet, das heißt

$$\delta_{|i-j|,\, i-j}\,\lambda_{i-j} = \begin{cases} \lambda_{i-j} & \text{für } i-j \geqq 0 \\ 0 & \text{für } i-j < 0, \end{cases} \tag{1.2}$$

also sind die Matrizen (1. 1) unendliche Matrizen folgender Form (im folgenden Λ-Matrizen genannt):

$$\Lambda = \begin{bmatrix} \lambda_0 & 0 & 0 & \dots \\ \lambda_1 & \lambda_0 & 0 & \dots \\ \lambda_2 & \lambda_1 & \lambda_0 & \dots \\ \cdot & \cdot & \cdot & \dots \\ \cdot & \cdot & \cdot & \dots \end{bmatrix}.$$

Grundlegend ist folgende Bemerkung: wird in (1. 1) $\lambda_0 = \lambda_2 = \lambda_3 = \cdots \cdots = 0$, $\lambda_1 = 1$ gewählt, d. h. die Λ-Matrix

(1. 3) $$Z = [\delta_{i-j,1}]; \qquad i, j = 0, 1, \dots$$

definiert, so können die Λ-Matrizen als Potenzreihen mit numerischen Koeffizienten von Z dargestellt werden (definitionsgemäß gilt $Z^0 \equiv E = [\delta_{i-j,0}]$):

(1. 4) $$\Lambda = \sum_{k=0}^{\infty} \lambda_k Z^k.$$

Es gilt nämlich für die Potenzen von Z (Beweis siehe weiter unten)

(1. 5) $$Z^k = [\delta_{i-j,k}]; \qquad k = 0, 1, \dots;$$

die rechte Seite von (1. 4) liefert mithin

$$\sum_{k=0}^{\infty} \lambda_k Z^k = \sum_{k=0}^{\infty} \lambda_k [\delta_{i-j,k}] = \left[\sum_{k=0}^{\infty} \lambda_k \delta_{i-j,k}\right],$$

hierbei ist nach der Definition des Kroneckerschen Symbols

$$\sum_{k=0}^{\infty} \lambda_k \delta_{i-j,k} = \begin{cases} \lambda_{i-j}, & \text{wenn} \quad i-j \geqq 0 \\ 0 & \text{wenn} \quad i-j < 0, \end{cases}$$

wodurch — unter Berücksichtigung von (1. 1) und (1. 2)—(1. 4) bestätigt worden ist.

Um nun (1. 5) zu beweisen, sei der k-te Basisspaltenvektor

$$e_k = [\delta_{i,k}]; \qquad i = 0, 1, \dots; \ k = 0, 1, \dots$$

definiert. Offenbar gilt

$$Ze_0 = e_1, \ Ze_1 = e_2, \dots$$

folglich

$$Z^2 e_0 = e_2, \ Z^2 e_1 = e_3, \dots$$

und allgemein

$$Z^k e_0 = e_k, \ Z^k e_1 = e_{k+1}, \dots .$$

Dies bedeutet aber eben das Bestehen von (1.5). (Betrachtet man einen Hauptminor vom Grade n der Matrix (1.3), so ist dieser eine nilpotente Matrix, nämlich verschwindet ihre $(n+1)$-te Potenz [4].)

Es ist nun leicht zu sehen, daß die Λ-Matrizen einen (kommutativen) hyperkomplexen Ring über dem Ring der komplexen Zahlen [5] definieren, in dem die folgenden Rechengesetze gelten

$$\lambda \sum_{k=0}^{\infty} \lambda'_k Z^k = \sum_{k=0}^{\infty} (\lambda \lambda'_k) Z^k, \tag{1.6}$$

$$\sum_{k=0}^{\infty} \lambda'_k Z^k + \sum_{k=0}^{\infty} \lambda''_k Z^k = \sum_{k=0}^{\infty} (\lambda'_k + \lambda''_k) Z^k, \tag{1.7}$$

$$\sum_{k=0}^{\infty} \lambda'_k Z^k \cdot \sum_{k=0}^{\infty} \lambda''_k Z^k = \sum_{k=0}^{\infty} (\lambda'_k \lambda''_0 + \cdots + \lambda'_0 \lambda''_k) Z^k. \tag{1.8}$$

(1.6)—(1.7) ist trivial und (1.8) wird wie folgt bewiesen: das Element in der i-ten Zeile und j-ten Spalte des Produktes in der linken Seite ist

a) im Falle $i \geqq j$ folgendes skalare Produkt:

$$[\lambda'_i, \ldots, \lambda'_{i-j}, \ldots, \lambda'_0, 0, \ldots] \left[\begin{array}{c} 0 \\ \vdots \\ 0 \\ \lambda''_0 \\ \vdots \\ \lambda''_{i-j} \\ \vdots \end{array}\right] \begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} j \text{ Zeilen} \\ \\ \\ \\ \\ \end{array} = \lambda'_{i-j}\lambda''_0 + \cdots + \lambda'_0 \lambda''_{i-j},$$

b) im Falle $i < j$ offenbar 0.

Wir bemerken nunmehr Folgendes: ordnet man der Λ-Matrix (1.1) die Zahlenfolge $\lambda_0, \lambda_1, \lambda_2, \ldots$ bzw. ihre erzeugende Funktion $g(z) = \lambda_0 + \lambda_1 z + + \lambda_2 z^2 + \cdots$ zu, so entsprechen die Regeln (1.6)—(1.7)—(1.8) denjenigen, welche sich auf die Addition und Multiplikation der erzeugenden Funktionen beziehen. Solange es jedoch der Fall sein kann, daß $g(z)$ allein bei $z = 0$ konvergiert und so nicht mehr fähig ist, die Zahlenfolge $\lambda_0, \lambda_1, \ldots$ „darzustellen", ist die Matrixrepräsentation dieser Zahlenfolge *per definitionem* immer möglich, was zur Folge hat, daß es auch jeweils die Rede von der Summe und dem Produkt zweier Zahlenfolgen im Sinne der Matrixzuordnung sein kann. Wir können somit den kommutativen Ring der Zahlenfolgen allgemein definieren. Unser Verfahren, nach dem die Zahlenfolgen umkehrbar eindeutig auf die Λ-Matrizen abgebildet werden, kann im wesentlichen als eine Matrixdarstellung der Operatorenrechnung von MIKUSIŃSKI angesehen werden, wenn man dieselbe auf Zahlenfolgen, als spezielle Funktionen anwendet [3], [6].

2. Was die Division im obigen Matrizenring anbetrifft, gilt es, daß die Λ-Matrizen im Falle $\lambda_0 \neq 0$ (im weiteren sei dann jeweils $\lambda_0 = 1$) ein zu diesem Matrizenring gehöriges inverses Element besitzen. Dies sieht man sofort aus der Gleichung (1.8), nach der die Elemente der Inverse rekurrierend berechnet werden können. Für den Fall, wo Λ ein Polynom Grades n von Z, also wie folgt darstellbar ist:

$$ {}_n\Lambda \equiv \sum_{k=0}^{n} \lambda_k Z^k \quad \text{und} \quad \lambda_0 = 1, \tag{1.9}$$

wollen wir die Inverse auch explizit darstellen.

Es ist leicht einzusehen, daß

$$ {}_n\Lambda^{-1} = \left(E + \sum_{k=1}^{n} \lambda_k Z^k\right)^{-1} = \sum_{\nu=0}^{\infty} \left(- \sum_{k=1}^{n} \lambda_k Z^k\right)^{\nu} $$

zutrifft, solange die unendliche Summe $\sum\limits_{\nu=0}^{\infty}$ existiert; davon aber kann man sich überzeugen, indem ${}_n\Lambda^{-1}$ nach den Potenzen von Z geordnet und auf diese Weise der Beiwert von Z^k (mit $\lambda_k^{(n)}$ bezeichnet) bestimmt wird. Unter Anwendung des polynomischen Lehrsatzes haben wir

$$ {}_n\Lambda^{-1} = \sum_{\nu=0}^{\infty} \sum_{k_1+k_2+\cdots+k_n=\nu} \frac{\nu!}{k_1!\,k_2!\ldots k_n!} (-\lambda_1)^{k_1}(-\lambda_2)^{k_2}\cdots(-\lambda_n)^{k_n} Z^{k_1+2k_2+\cdots+nk_n}, $$

wo $k_1, k_2, \ldots, k_n$ nichtnegative ganze Zahlen sind, und mit der Bezeichnung

$$ k = k_1 + 2k_2 + \cdots + nk_n = \nu + k_2 + \cdots + (n-1)k_n $$

schließlich

$$ \lambda_k^{(n)} = \sum_{k_1+\cdots+nk_n=k} \frac{(k-k_2-\cdots-(n-1)k_n)!}{k_1!\ldots k_n!} (-\lambda_1)^{k_1}\cdots(-\lambda_n)^{k_n}; \qquad k = 0, 1, \ldots . \tag{1.10}$$

Dadurch ist die Darstellung von ${}_n\Lambda^{-1}$ als Element des Ringes der Λ-Matrizen:

$$ {}_n\Lambda^{-1} = \sum_{k=0}^{\infty} \lambda_k^{(n)} Z^k \tag{1.11}$$

bis auf die Auflösung der Diophantischen Gleichung $k_1 + 2k_2 + \cdots + nk_n = k$ explizit durchgeführt worden.

Wir können ${}_n\Lambda^{-1}$ auch in einer anderen Weise darstellen, wenn die Nullstellen $z_1, \ldots, z_m$ des Polynoms $\sum_{k=0}^{n} \lambda_k Z^{n-k}$ mit den Ordnungszahlen ihres Vorkommens $\alpha_1, \ldots, \alpha_m$ $(\alpha_1 + \cdots + \alpha_m = n)$, und somit die Partialbruchzerlegung

$$ {}_n\Lambda^{-1} = \frac{1}{\sum_{k=0}^{n} \lambda_k Z^k} = \sum_{r=1}^{m} \sum_{s=1}^{\alpha_r} \frac{a_{rs}}{(E - z_r Z)^s} $$

bekannt sind. Nämlich gilt die Identität

$$ \frac{1}{(E - z_r Z)^s} = \sum_{k=0}^{\infty} \binom{k+s-1}{s-1} z_r^k Z^k, $$

woraus

$$ \lambda_k^{(n)} = \sum_{r=1}^{m} \sum_{s=1}^{\alpha_r} a_{rs} \binom{k+s-1}{s-1} z_r^k \tag{1.12} $$

folgt.

§ 2. Lösung von gewöhnlichen linearen Differenzengleichungen mit konstanten Koeffizienten

Wir wollen nunmehr die im § 1 definierte Zuordnung von Zahlenfolgen und Λ-Matrizen auf die Lösung gewöhnlicher linearer Differenzengleichungen mit konstanten Koeffizienten anwenden. Dazu brauchen wir zwei triviale Bemerkungen ins Auge zu fassen:

1. Die Zuordnung ist eine homogen-lineare, d. h. wenn λ_k' und λ_k'' $(k = 0, 1, \ldots)$ zwei Zahlenfolgen, sowie c_1 und c_2 zwei Zahlen sind, so gilt

$$ c_1 \lambda_k' + c_2 \lambda_k'' \sim c_1 \Lambda' + c_2 \Lambda'' \quad \left(\Lambda' = \sum_{k=0}^{\infty} \lambda_k' Z^k,\ \Lambda'' = \sum_{k=0}^{\infty} \lambda_k'' Z^k\right), \tag{2.1} $$

wo $\sim$ das Zeichen der Zuordnung ist.

2. Es gilt die „Verschiebungsregel"

$$ Z^s \sum_{k=0}^{\infty} \lambda_{k+s} Z^k = \sum_{k=0}^{\infty} \lambda_k Z^k - \sum_{k=0}^{s-1} \lambda_k Z^k, \tag{2.2} $$

wo s eine positive ganze Zahl bezeichnet.

Wir können nun der Differenzengleichung

$$ a_n u_{k+n} + a_{n-1} u_{k+n-1} + \cdots + a_0 u_k = b_k; \qquad k = 0, 1, \ldots \quad (a_n = 1) \tag{2.3} $$

(wo b_k eine bekannte Zahlenfolge ist) mit den Anfangswerten

$$ u_\sigma = u^{(\sigma)}; \qquad \sigma = 0, 1, \ldots, n-1 \tag{2.4} $$

die folgende Matrizengleichung zuordnen, nachdem wir die beiden Seiten von (2.3) mit Z^n multipliziert und (2.1)—(2.2) berücksichtigt haben:

$$a_n\left(A-\sum_{k=0}^{n-1}u^{(k)}Z^k\right)+a_{n-1}Z\left(A-\sum_{k=0}^{n-2}u^{(k)}Z^k\right)+\cdots+a_0Z^nA=Z^n\sum_{k=0}^{\infty}b_kZ^k, \tag{2.5}$$

wo

$$A=\sum_{k=0}^{\infty}u_kZ^k$$

gesetzt worden ist.

Umgestellt läßt sich (2.5) in der Form

$$A=\frac{\sum_{k=0}^{\infty}b_kZ^{k+n}+c_{n-1}Z^{n-1}+\cdots+c_0E}{a_nE+a_{n-1}Z+\cdots+a_0Z^n} \tag{2.6}$$

schreiben, mit

$$c_\sigma=a_nu^{(\sigma)}+\cdots+a_{n-\sigma}u^{(0)};\qquad \sigma=0,1,\ldots,n-1. \tag{2.7}$$

Es ist nun nach (1.9)—(1.11) erlaubt

$$\frac{1}{a_nE+a_{n-1}Z+\cdots+a_0Z^n}=\sum_{k=0}^{\infty}\lambda_k^{(n)}Z^k$$

zu setzen (wo $\lambda_k^{(n)}$ gemäß (1.10) oder (1.12) berechnet werden kann), wodurch man nach Ausmultiplizieren aus (2.6)

$$u_k=b_{k-n}\lambda_0^{(n)}+\cdots+b_0\lambda_{k-n}^{(n)}+c_{n-1}\lambda_{k-n+1}^{(n)}+\cdots+c_0\lambda_k^{(n)};\qquad k=0,1,\ldots \tag{2.8}$$

erhält. Hierbei ist unter einem Ausdruck mit negativem Index Null zu verstehen.

Nach einer Umstellung wird (2.8)

$$u_k=b_{k-n}\lambda_0^{(n)}+\cdots+b_0\lambda_{k-n}^{(n)}+d_{n-1}u^{(n-1)}+\cdots+d_0u^{(0)};\qquad k=0,1,\ldots, \tag{2.9}$$

wo

$$d_\sigma=a_{\sigma+1}\lambda_{k-n+1}^{(n)}+\cdots+a_n\lambda_{k-\sigma}^{(n)};\qquad \sigma=0,1,\ldots,n-1. \tag{2.10}$$

Da nach (2.8) $\lambda_{k-n+1}^{(n)}$ die Lösung der Gleichung (2.3) bei $b_k=u^{(0)}=\cdots=u^{(n-2)}=0$, $u^{(n-1)}=1$ ist, gilt die Identität

$$a_n\lambda_{k+1}^{(n)}+a_{n-1}\lambda_k^{(n)}+\cdots+a_0\lambda_{k-n+1}^{(n)}=0;\qquad k=0,1,\ldots$$

oder, was gleichbedeutend ist:

$$a_n\lambda_{k+1-l}^{(n)}+a_{n-1}\lambda_{k-l}^{(n)}+\cdots+a_0\lambda_{k-n+1-l}^{(n)}=0;\quad l=0,1,\ldots;\ k=l,l+1,\ldots. \tag{2.11}$$

Nach (2.11) kann man für $l=\sigma+1$ in (2.9) statt d_σ auch

$$-e_\sigma \equiv -(a_0\lambda_{k-n-\sigma}^{(n)}+\cdots+a_\sigma\lambda_{k-n}^{(n)}); \qquad k=\sigma+1, \sigma+2, \ldots \tag{2.12}$$

schreiben. Das bedeutet aber, daß sich (2.9) auch folgendermaßen darstellen läßt:

$$\begin{aligned} u_k = (b_{k-n}\lambda_0^{(n)}+\cdots+b_0\lambda_{k-n}^{(n)})+(d_{n-1}u^{(n-1)}+\cdots+d_\nu u^{(\nu)})-\\ -(e_{\nu-1}u^{(\nu-1)}+\cdots+e_0u^{(0)}), \end{aligned} \tag{2.13}$$

wo

$$k=\nu, \nu+1, \ldots \quad \text{und} \quad n=\begin{cases} 2\nu \\ 2\nu+1 \end{cases}$$

je nach dem, ob n eine gerade oder ungerade Zahl ist. Der Vorteil von (2.13) gegenüber (2.9) besteht darin, daß in (2.13) der Koeffizient von $u^{(\sigma)}$ $(\sigma=0, 1, \ldots, n-1)$ ein Ausdruck von höchstens ν bzw. $\nu+1$ Glieder ist, solange in (2.9) die Anzahl der Glieder auch n sein kann.

Zum Beispiel: es sind die Fibonaccischen Zahlen [2] zu bestimmen. Zu lösen ist also die Differenzengleichung

$$u_{k+2}-u_{k+1}-u_k=0$$

mit den Anfangswerten: $u^{(0)}=0$, $u^{(1)}=1$. Nach (2.8) erhält man

$$u_k=\lambda_{k-1}^{(2)}; \qquad k=0, 1, \ldots,$$

wo jetzt

$$\frac{1}{E-Z-Z^2}=\sum_{k=0}^{\infty}\lambda_k^{(2)}Z^k$$

ist.

Wir wollen nun die Zahlen $\lambda_k^{(2)}$ durch Partialbruchzerlegung gewinnen, daß heißt:

$$\frac{1}{E-Z-Z^2}=\frac{1}{\left(E-\frac{1+\sqrt{5}}{2}Z\right)\left(E-\frac{1-\sqrt{5}}{2}Z\right)}=$$

$$=\frac{1}{\sqrt{5}}\left(\frac{\frac{1+\sqrt{5}}{2}}{E-\frac{1+\sqrt{5}}{2}Z}-\frac{\frac{1-\sqrt{5}}{2}}{E-\frac{1-\sqrt{5}}{2}Z}\right)=$$

$$=\frac{1}{\sqrt{5}}\sum_{k=0}^{\infty}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{k+1}-\left(\frac{1-\sqrt{5}}{2}\right)^{k+1}\right]Z^k,$$

woraus

$$u_k = \lambda_{k-1}^{(2)} = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^k - \left(\frac{1-\sqrt{5}}{2}\right)^k\right]; \quad k = 0, 1, \ldots$$

folgt.

Literatur

[1] Á. Pethő, Eine Matrixmethode zur Lösung von Anfangswertaufgaben linearer Differenzengleichungen, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960) 203—213.

[2] Ch. Jordan, Calculus of finite differences, *Budapest*, 1939.

[3] J. Mikusiński, Operatorenrechnung, *Berlin*, 1957.

[4] I. M. Gelfand, Előadások a lineáris algebráról, *Budapest*, 1955.

[5] L. Rédei, Algebra, 1. Teil, *Leipzig*, 1959.

[6] I. Fenyő, Eine neue Methode zur Lösung von Differenzengleichungen nebst Anwendungen, *Periodica Polytechnica (Elektrotechnik)* **3** (1959) 135—151.

(Eingegangen am 3. Dezember 1960.)