SOLUTION OF EQUATION $AX + XB = C$ BY INVERSION OF AN $M \times M$ OR $N \times N$ MATRIX *

ANTONY JAMESON†

SIAM J. Appl. Math.
Vol. 16, No. 5, September 1968

It is often of interest to solve the equation

$$AX + XB = C \tag{1}$$

for X, where X and C are $M \times N$ real matrices, A is an $M \times M$ real matrix, and B is an $N \times N$ real matrix. A familiar example occurs in the Lyapunov theory of stability [1], [2], [3] with

$$B = A^T.$$

Is also arises in the theory of structures [4].

Using the notation $P \times Q$ to denote the Kronecker product $(P_{ij}Q)$ (see [5]) in which each element of P is multipled by Q, we find that the equation written out in full for the MN unknowns $x_{11}, x_{21}, \ldots, x_{12}, \ldots$ in terms of $c_{11}, c_{21}, \ldots, c_{12}, \ldots$ becomes

$$\left[(I_N \times A) + \left(B^T \times I_M\right)\right] x = c, \tag{2}$$

where I_M and I_N are the $M \times M$ and $N \times N$ identity matrices. If u is a characteristic vector of A with characteristic value λ, and v is a characteristic vector of B^T with characteristic value μ, then

$$Auv^T + uv^T B = (\lambda + \mu)\, uv^T.$$

Thus $\lambda + \mu$ is a characteristic value of the system (2), which can therefore be solved if and only if

$$\lambda_i + \mu_j \neq 0 \tag{3}$$

for all i, j.

When A and B can both be reduced to diagonal form by similarity transformations

$$U^{-1}AU = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_M \end{bmatrix}$$

*Received by the editors March 7, 1968.
†Grumman Aircraft Engineering Corporation, Bethpage, Long Island, New York 11714.

and

$$V^{-1}BV = \begin{bmatrix} \mu_1 & & & \\ & \mu_2 & & \\ & & \ddots & \\ & & & \mu_N \end{bmatrix},$$

the solution of (1) is easily obtained as

$$X = U\tilde{X}V^{-1},$$

where

$$\tilde{x}_{ij} = \frac{\tilde{c}_{ij}}{\lambda_i + \mu_j}$$

and

$$\tilde{C} = U^{-1}CV.$$

In fact the matrix of the expanded system (2) is reduced to a diagonal form by the transformation

$$\left[\left(V^{-1}\right)^T \times U\right] \left[(I_N \times A) + \left(B^T \times I_M\right)\right] \left[V^T \times U^{-1}\right]$$

$$= \begin{bmatrix} \lambda_1 + \mu_1 & & & & & & \\ & \lambda_2 + \mu_1 & & & & & \\ & & \lambda_3 + \mu_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_1 + \mu_2 & & \\ & & & & & \ddots & \\ & & & & & & \lambda_M + \mu_N \end{bmatrix}.$$

It is, however, possible to obtain the solution by the inversion of an $M \times M$ or $N \times N$ matrix without recourse to diagnalization. From the expression

$$\begin{aligned}
&C_0 = 0, & & \\
&C_1 = C & &= AX + XB, \\
&C_2 = AC_1 - C_1B + AC_0B & &= A^2X - XB^2, \\
&C_3 = AC_2 - C_2B + AC_1B & &= A^3X + XB^3, \\
& & &\cdots \\
&C_k = AC_{k-1} - C_{k-1}B + AC_{k-2}B & &= A^kX - (-1)^k XB^k, \\
& & &\cdots
\end{aligned}$$

Let the characteristic equations of A and B be

$$|\lambda I - A| = \lambda^M + a_1\lambda^{M-1} + \ \cdots \ + a_M = (\lambda - \lambda_1)(\lambda - \lambda_2) \ \cdots \ (\lambda - \lambda_M) = 0$$

and

$$|\lambda I - B| = \mu^N + b_1\mu^{N-1} + \cdots + b_N = (\mu - \mu_1)(\mu - \mu_2) \cdots (\mu - \mu_N) = 0.$$

According to the Cayley-Hamilton theorem these are satisfied by A and B themselves. Therefore

$$\begin{aligned} C_N - b_1 C_{N-1} + \cdots + (-1)^{N-1} b_{N-1} C_1 & \\ = A^N X - b_1 A^{N-1} X + \cdots + (-1)^{N-1} b_{N-1} A X + (-1)^N b_N X \end{aligned}$$

and

$$\begin{aligned} C_M - a_1 C_{M-1} + \cdots + a_{M-1} C_1 & \\ = -a_M X - (-1)^M \left[X B^M - a_1 X B^{M-1} + \cdots + (-1)^{M-1} a_{M-1} X B\right]. \end{aligned}$$

Thus

$$\begin{aligned} X &= G^{-1}\left[C_N - b_1 C_{N-1} + \cdots + (-1)^{N-1} b_{N-1} C_1\right] \\ &= (-1)^{M-1}\left[C_M + a_1 C_{M-1} + \cdots + a_{M-1} C_1\right] H^{-1}, \end{aligned}$$

where

$$\begin{aligned} G &= A^N - b_1 A^{N-1} + \cdots + (-1)^N b_N I \\ &= (A + \mu_1 I)(A + \mu_2 I) \cdots (A + \mu_N I) \end{aligned}$$

and

$$\begin{aligned} H &= B^M - a_1 B^{M-1} + \cdots + (-1)^M a_M I \\ &= (B + \lambda_1 I)(B + \lambda_2 I) \cdots (B + \lambda_M I). \end{aligned}$$

Since the determinant of a product of matrices is the product of their determinants, it is evident that G is not invertible if for any $i, -\mu_i$ is a charecteristic value of A, and similarly H is not invertible if for any $i, -\lambda_i$ is a characteristic value of B; that is, G and H are not invertible if condition (3) does not hold.

The coefficients $a_1, a_2, \cdots, a_M$ and $b_1, b_2, \cdots, b_N$ can be determined using Bocher's identities [6]:

$$\begin{aligned} a_1 &= -\mathrm{tr}(A), \\ a_2 &= -\frac{1}{2}\left[a_1 \mathrm{tr}(A) + \mathrm{tr}\left(A^2\right)\right], \\ a_3 &= -\frac{1}{3}\left[a_2 \mathrm{tr}(A) + a_1 \mathrm{tr}\left(A^2\right) + \mathrm{tr}\left(A^3\right)\right], \\ &\cdots \\ a_M &= -\frac{1}{M}\left[a_{M-1} \mathrm{tr}(A) + a_{M-2} \mathrm{tr}\left(A^2\right) + \cdots + \mathrm{tr}\left(A^M\right)\right] = (-1)^M |A|. \end{aligned}$$

In the case of the Lyapunov equation

$$AX + XA^T = C,$$

the coefficients a_i and b_i coincide and could be determined in the course of taking the powers of A to form the G matrix. By adding or subtracting the characteristic equation for A, the G matrix can also be expressed in the terms of even or odd powers of A only:

$$\begin{aligned} G &= 2\,(-1)^N \left[a_N I + a_{N-2} A^2 + \ \cdots\right] \\ &= 2\,(-1)^{N-1} \left[a_{N-1} A + a_{N-3} A^3 + \ \cdots\right]. \end{aligned}$$

An explicit solution can be obtained by successive elimination of powers of A between the two expressions for G. For example, the solution of the Lyapunov equation for 2×2 matrices can be written as

$$X = \frac{1}{2\,\mathrm{tr}\,(A)} \left[C + |A|\, A^{-1} C \left(A^{-1}\right)^T\right].$$

The solution for 3×3 matrices can be written as

$$X = \frac{1}{2\,(a_1 a_2 - a_3)} \left[A^2 C - ACA^T + CA^{T^2} + \left(a_2 - a_1^2\right) C - a_1 a_3 A^{-1} C \left(A^{-1}\right)^T\right].$$

References

[1] S. Lefschetz and J. P. LaSalle, *Stability by Lyapunov's Direct Method*, Academic Press, New York, 1961.

[2] R. A. Smith, *Matrix Calculations for Liapunov quadratic forms*, J. Differential Equations, 2 (1966), pp. 208-217.

[3] S. Barnett and C. Storey, *Analysis and synthesis of stability matrices*, Ibid., 3 (1967), pp. 414-422.

[4] Er-Chieh Ma, *A finite series solution of the matrix equation $AX - XB = C$*, this Journal, 14 (1966), pp. 490-495.

[5] Richard Bellman, *Introduction to Matrix Analysis*, McGraw-Hill, New York, 1960.

[6] M. Bocher, *Introduction to Higher Algebra*, Macmillan, New York, 1947.