## **Explicit Solution of the Matrix Equation** $AXB - CXD = E$


Vicente Hernández
*Departamento de Sistemas Informáticos y Computación*
*Universidad Politécnica de Valencia*
*Apdo. 22012*
*46071 Valencia, Spain*

and

Maite Gassó
*Departamento de Matemática Aplicada*



Submitted by Graciano de Oliveira



ABSTRACT

The unique solution of the matrix equation $AXB - CXD = E$ is obtained in explicit form by means of the inversion of an $n \times n$ or $m \times m$ matrix from the coefficients of the Laurent expansions of $(\lambda C - A)^{-1}$ and $(\lambda B - D)^{-1}$ and the relative characteristic polynomial of $\lambda C - A$ or $\lambda B - D$ respectively. The case of singular pencils is reduced to the regular one.


## 1. INTRODUCTION

We consider the linear matrix equation

$$AXB - CXD = E \tag{1.1}$$

where $A, C \in \mathbb{K}^{m\times m}$, $B, D \in \mathbb{K}^{n\times n}$, $E \in \mathbb{K}^{m\times n}$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$). This equation appears in the study of perturbations of the generalized eigenvalue problem [1], in the numerical solution of implicit ordinary differential equations [2], and in stability problems for descriptor systems [3].

The most general method for solving this equation is based on the Kronecker product [4]. In this approach, Equation (1.1) can be written as

$$(A \otimes B^{\mathsf{T}} - C \otimes D^{\mathsf{T}}) v(X) = v(E),$$





655 Avenue of the Americas, New York, NY 10010 0024-3795/89/$3.50

where

$$v(X) = \left[x_1^{\mathrm{T}}, \dots, x_m^{\mathrm{T}}\right]^{\mathrm{T}}$$

with $x_i^{\mathrm{T}}$ the $i$th row of $X$. The unique solution of (1.1) can then be obtained by means of the inversion of an $mn \times mn$ matrix. Another general method [5, 6] is based on the reduction of the pencils $\lambda C - A$ and $\lambda B - D$ to the Kronecker canonical form, which permits us to obtain the solution of (1.1) in a finite series form [7].

In another approach [2, 8], efficient numerical algorithms have been proposed for the solution of (1.1), based on the reduction of the pencils $(A, C), (D, B)$ to the Hessenberg form or the Schur form. In [8] it has also been proved that the matrix equation (1.1) has a unique solution if and only if (i) the pencils $\lambda C - A$ and $\lambda B - D$ are regular, and (ii) the spectra of these pencils have an empty intersection.

In Section 2, we present a theoretical approach to the solution of (1.1) based on the relative Cayley-Hamilton theorem [9]. Under hypotheses (i) and (ii) we obtain the unique solution of (1.1) by means of the inversion of an $n \times n$ or $m \times m$ matrix. Equation (1.1) is a generalization of the Sylvester equation

$$AX - XD = E \tag{1.2}$$

discussed by Jameson [10], and this paper is strongly influenced by his work. Note that Equation (1.1) can be transformed to the Sylvester equation when $B$ and $C$ is nonsingular. In Section 2 we suppose that only one of these matrices is nonsingular.

In Section 3, we discuss the case of singular pencils. We prove that if (iii) the pencils $\lambda C - A$ and $\lambda B - D$ are right and left invertible respectively, and (iv) they have disjoint spectra, then Equation (1.1) is consistent. The proof is based on the reduction to the case of regular pencils by using the same technique applied in [11] to the matrix equation in $Y$ and $Z$

$$(\lambda C - A)Y - Z(\lambda B - D) = \lambda F - E. \tag{1.3}$$

Note that Equation (1.1) is equivalent to

$$(\lambda C - A)XB - CX(\lambda B - D) = -E, \tag{1.4}$$

but it is not in general equivalent to (1.3).

## 2. THE CASE OF REGULAR PENCILS

We consider the matrix equation (1.4). If we suppose that $\lambda C - A$ and $\lambda B - D$ are $m \times m$ and $n \times n$ regular pencils respectively, then (1.4) can be written as [12]

$$XB(\lambda B - D)^{-1} - (\lambda C - A)^{-1}CX = -(\lambda C - A)^{-1}E(\lambda B - D)^{-1} \quad (2.1)$$

for some $\lambda \in \mathbb{K}$. The regularity of the pencils $\lambda C - A$ and $\lambda B - D$ implies the existence of the Laurent expansions

$$(\lambda C - A)^{-1} = \lambda^{-1} \sum_{k=-\mu}^{\infty} U_k \lambda^{-k}, \qquad U_k \in \mathbb{K}^{m \times m}, \quad (2.2)$$

$$(\lambda B - D)^{-1} = \lambda^{-1} \sum_{k=-\tau}^{\infty} V_k \lambda^{-k}, \qquad V_k \in \mathbb{K}^{n \times n}, \quad (2.3)$$

in a deleted neighborhood of zero. We can suppose, without loss of generality, that the indices of nilpotency $\mu, \tau$ are such that $\mu \geqslant \tau$. Then expansion (2.3) can be written as

$$(\lambda B - D)^{-1} = \lambda^{-1} \sum_{k=-\mu}^{\infty} V_k \lambda^{-k}, \qquad V_k \in \mathbb{K}^{n \times n}, \quad (2.4)$$

where

$$V_k = 0, \qquad -\mu \leqslant k < -\tau.$$

By substituting (2.2) and (2.4) in (2.1) we obtain that

$$\lambda^{-1} \sum_{k=-\mu}^{\infty} (XBV_k - U_k CX)\lambda^{-k} = \lambda^{-1} \sum_{k=-2\mu+1}^{\infty} W_k \lambda^{-k} \quad (2.5)$$

with

$$W_k = -\sum_{j=-\mu}^{k+\mu-1} U_j E V_{k-j-1}, \qquad k \geqslant -2\mu + 1.$$

Comparison of the coefficients in the two members of (2.5) allows us to obtain

$$\begin{aligned} W_k &= 0, \qquad k = -2\mu+1,\ -2\mu+2,\dots,\ -\mu-1, \\ XBV_k - U_kCX &= W_k, \qquad k \geqslant -\mu. \end{aligned} \tag{2.6}$$

Now we consider the relative characteristic polynomial of $\lambda C - A$,

$$\Delta_1(\lambda) = \det(\lambda C - A) = p_m\lambda^m + p_{m-1}\lambda^{m-1} + \cdots + p_0.$$

By the relative Cayley-Hamilton theorem [9],

$$\Delta_1(U_k) = p_mU_k + p_{m-1}U_{k-1} + \cdots + p_0U_{k-m} = 0 \tag{2.7}$$

for $k \geqslant m$ or $k \leqslant -1$.

From (2.6) we obtain that

$$\begin{aligned} XBV_0 - U_0CX &= W_0, \\ XBV_1 - U_1CX &= W_1, \\ &\vdots \\ XBV_m - U_mCX &= W_m. \end{aligned} \tag{2.8}$$

If we multiply the $i$th equation of (2.8) by $p_i$, for every $i = 0, 1, \dots, m$, and add the resultant expressions, we verify that

$$XB\Delta_1(V_m) - \Delta_1(U_m)CX = \Delta_1(W_m),$$

where

$$\Delta_1(V_m) = p_mV_m + p_{m-1}V_{m-1} + \cdots + p_0V_0, \tag{2.9}$$

$$\Delta_1(W_m) = p_mW_m + p_{m-1}W_{m-1} + \cdots + p_0W_0 \tag{2.10}$$

Taking into account the expression (2.7), we deduce that

$$XB\Delta_1(V_m) = \Delta_1(W_m).$$

Thus we can prove the following result.

THEOREM 1. *We consider the linear matrix equation*

$$AXB - CXD = E \tag{2.11}$$

*with* $A, C \in \mathbb{K}^{m\times m}$, $B, D \in \mathbb{K}^{n\times n}$, $E \in \mathbb{K}^{m\times n}$, *where the pencils* $\lambda C - A$, $\lambda B - D$ *are regular; and we suppose that their indices of nilpotency* $\mu, \tau$ *are such that* $\mu \geqslant \tau$. *We also consider the Laurent expansions*

$$\begin{aligned}(\lambda C - A)^{-1} &= \lambda^{-1} \sum_{k=-\mu}^{\infty} U_k \lambda^{-k},\\ (\lambda B - D)^{-1} &= \lambda^{-1} \sum_{k=-\tau}^{\infty} V_k \lambda^{-k},\end{aligned} \tag{2.12}$$

*and*

$$W_k = -\sum_{j=-\mu}^{k+\mu-1} U_j E V_{k-j-1}, \qquad k \geqslant -2\mu + 1, \tag{2.13}$$

*where*

$$V_k = 0, \qquad -\mu \leqslant k < -\tau.$$

*Then it follows that:*

(a) *If* (2.11) *is consistent, then*

$$W_k = 0, \qquad -2\mu + 1 \leqslant k \leqslant -\mu - 1.$$

(b) *If* $X$ *is a solution of* (2.11), *then*

$$XB\Delta_1(V_m) = \Delta_1(W_m),$$

*where* $\Delta_1(\lambda)$ *is the relative characteristic polynomial of* $\lambda C - A$, *and the matrices* $\Delta_1(V_m)$, $\Delta_1(W_m)$ *are defined by* (2.9), (2.10) *respectively.*

(c) *If the spectra of the pencils satisfy*

$$\sigma(C, A) \cap \sigma(B, D) = \varnothing, \tag{2.14}$$

*then the unique solution of* (2.11) *is a solution of*

$$XB = \Delta_1(W_m)[\Delta_1(V_m)]^{-1}.$$

(d) *If, in addition to* (2.14), *the pencil* $\lambda B - D$ *has no eigenvalues at* $\infty$, *then the unique solution of* (2.11) *is given by*

$$X = \Delta_1(W_m)[B\Delta_1(V_m)]^{-1}. \tag{2.15}$$

*Proof.* The assertions (a) and (b) have been previously proved, and (d) is an immediate consequence of (c). From (b) we only need to prove that (2.14) implies that matrix $\Delta_1(V_m)$ is nonsingular. We suppose that the pencil $\lambda C - A$ has eigenvalues at $\infty$ and finite eigenvalues $\alpha_1, \dots, \alpha_r$. If the spectra of the pencils $\lambda C - A$ and $\lambda B - D$ verify (2.14), then $\lambda B - D$ only has finite eigenvalues $\beta_1, \dots, \beta_s$ such that

$$\Delta_1(\beta_j) \neq 0, \qquad j = 1, \dots, s. \tag{2.16}$$

From [4], there exist $n \times n$ invertible matrices $P, Q$ such that

$$P(\lambda B - D)Q = [\lambda I_{n_1} - J_{n_1}(\beta_1)] \oplus \cdots \oplus [\lambda I_{n_s} - J_{n_s}(\beta_s)],$$

where $I_k$ denotes the $k \times k$ identity matrix and $J_k(\beta)$ the $k \times k$ Jordan block with eigenvalue $\beta$, $\Sigma_{j=1}^{s} n_j = n$. Then we obtain that

$$(\lambda B - D)^{-1} = Q\{[\lambda I_{n_1} - J_{n_1}(\beta_1)]^{-1} \oplus \cdots \oplus [\lambda I_{n_s} - J_{n_s}(\beta_s)]^{-1}\}P, \tag{2.17}$$

where $[\lambda I_{n_j} - J_{n_j}(\beta_j)]^{-1}$ is an upper triangular matrix

$$[\lambda I_{n_j} - J_{n_j}(\beta_j)]^{-1} = \begin{bmatrix} (\lambda - \beta_j)^{-1} & \cdot & \cdot & \cdot & \cdot \\ & & \cdot & & \vdots \\ & & & \cdot & \vdots \\ & & & & (\lambda - \beta_j)^{-1} \end{bmatrix} \tag{2.18}$$

with

$$(\lambda - \beta_j)^{-1} = \lambda^{-1} \sum_{k=0}^{\infty} \beta_j^k \lambda^{-k} \tag{2.19}$$

for every $j = 1, \dots, s$.

From (2.17)–(2.19) we obtain that the coefficients of the Laurent expansion (2.12) are given by

$$V_k = Q \begin{bmatrix} \beta_1^k & \cdot & \cdot & \cdot & \cdot & \cdot \\ & \ddots & & & & \cdot \\ & & \beta_1^k & & & \cdot \\ & & & \ddots & & \cdot \\ & & & & \beta_s^k & \cdot \\ & & & & & \ddots \\ & & & & & \beta_s^k \end{bmatrix} P, \qquad k \geqslant 0. \tag{2.20}$$

By substituting (2.20) in (2.9) we obtain that

$$\Delta_1(V_m) = Q \begin{bmatrix} \Delta_1(\beta_1) & \cdot & \cdot & \cdot & \cdot & \cdot \\ & \ddots & & & & \cdot \\ & & \Delta_1(\beta_1) & & & \cdot \\ & & & \ddots & & \cdot \\ & & & & \Delta_1(\beta_s) & \cdot \\ & & & & & \ddots \\ & & & & & \Delta_1(\beta_s) \end{bmatrix} P,$$

and from (2.16) we conclude that $\Delta_1(V_m)$ is nonsingular. ∎

NOTE. An alternative result can be obtained from the relative characteristic polynomial of $\lambda B - D$, by replacing conclusions (b), (c), (d) in Theorem 1 with:

(b′) If $X$ is a solution of (2.11), then

$$\Delta_2(U_n)CX = -\Delta_2(W_n),$$

where

$$\Delta_2(\lambda) = \det(\lambda B - D) = q_n\lambda^n + q_{n-1}\lambda^{n-1} + \cdots + q_0,$$

$$\Delta_2(U_n) = q_nU_n + q_{n-1}U_{n-1} + \cdots + q_0U_0,$$

$$\Delta_2(W_n) = q_nW_n + q_{n-1}W_{n-1} + \cdots + q_0W_0.$$

(c′) If the spectra of the pencils satisfy (2.14), then the unique solution of (2.11) is a solution of

$$CX = -[\Delta_2(U_n)]^{-1}\Delta_2(W_n).$$

(d′) If, in addition to (2.14), the pencil $\lambda C - A$ has no eigenvalues at $\infty$, then the unique solution of (2.11) is given by

$$X = -[\Delta_2(U_n)C]^{-1}\Delta_2(W_n). \tag{2.21}$$

Then we obtain the following result.

Corollary 1. *If* (i) *the pencils* $\lambda C - A$ *and* $\lambda B - D$ *are regular and* (ii) $\sigma(C, A) \cap \sigma(B, D) = \varnothing$, *then the unique solution of the matrix equation* (2.11) *is given by* (2.15) *or* (2.21).

*Proof.* This assertion is a consequence of the previous note and Theorem 1, because condition (ii) implies $\lambda B - D$ or $\lambda C - A$ has no eigenvalues at $\infty$. ∎

In the following we show that Jameson's result [10] can be obtained as a consequence of Theorem 1.

Corollary 2 (Jameson). *We consider the matrix equation*

$$AX - XD = E \tag{2.22}$$

*with* $A \in \mathbb{K}^{m\times m}$, $D \in \mathbb{K}^{n\times n}$, $E \in \mathbb{K}^{m\times n}$. *Then it follows that:*

(a) *If* $X$ *is a solution of* (2.22), *then*

$$X\Delta_1(D) = -\sum_{k=1}^{m}\sum_{j=0}^{k-1} p_k A^j E D^{k-j-1},$$

*where*

$$\Delta_1(\lambda) = \det(\lambda I_m - A) = p_m\lambda^m + p_{m-1}\lambda^{m-1} + \cdots + p_0.$$

(b) *If the spectra of the coefficient matrices satisfy*

$$\sigma(A) \cap \sigma(D) = \varnothing,$$

*then the unique solution of* (2.22) *is given by*

$$X = \left(-\sum_{k=1}^{m}\sum_{j=0}^{k-1} p_k A^j E D^{k-j-1}\right)[\Delta_1(D)]^{-1}.$$

*Proof.* Equation (2.22) is a particular case of (2.11) where $B = I_n$, $C = I_m$, and verifies that the regular pencils $\lambda I_m - A$, $\lambda I_n - D$ have indices of nilpotency $\mu = \tau = 0$. The corresponding Laurent expansions are given by

$$(\lambda I_m - A)^{-1} = \lambda^{-1}\sum_{k=0}^{\infty} U_k \lambda^{-k}, \qquad U_k = A^k,$$

$$(\lambda I_n - D)^{-1} = \lambda^{-1}\sum_{k=0}^{\infty} V_k \lambda^{-k}, \qquad V_k = D^k.$$

From (2.13) we obtain that

$$W_k = -\sum_{j=0}^{k-1} A^j E D^{k-j-1}, \qquad k \geqslant 1.$$

Then applying Theorem 1, the result is obtained. ■

## 3. THE CASE OF SINGULAR PENCILS

We consider the matrix equation (1.4) where $\lambda C - A$ and $\lambda B - D$ are $m \times n$ and $p \times q$ singular pencils respectively. We can transform these pencils to their Kronecker canonical form [4] via the strictly equivalent transformations

$$\begin{aligned} P_1(\lambda C - A)Q_1 &= \lambda C_0 - A_0, \\ P_2(\lambda B - D)Q_2 &= \lambda B_0 - D_0. \end{aligned} \tag{3.1}$$

If we suppose that $\lambda C - A$ and $\lambda B - D$ are right and left invertible respectively, then

$$\begin{aligned} \lambda C_0 - A_0 = {} & L_{r_1}(\lambda) \oplus \cdots \oplus L_{r_a}(\lambda) \\ & \oplus \left[ I_{m_1} - \lambda J_{m_1}(0) \right] \oplus \cdots \oplus \left[ I_{m_b} - \lambda J_{m_b}(0) \right] \\ & \oplus \left[ \lambda I_{n_1} - J_{n_1}(\alpha_1) \right] \oplus \cdots \oplus \left[ \lambda I_{n_c} - J_{n_c}(\alpha_c) \right], \\ \lambda B_0 - D_0 = {} & L_{l_1}^{\mathrm{T}}(\lambda) \oplus \cdots \oplus L_{l_d}^{\mathrm{T}}(\lambda) \\ & \oplus \left[ I_{p_1} - \lambda J_{p_1}(0) \right] \oplus \cdots \oplus \left[ I_{p_e} - \lambda J_{p_e}(0) \right] \\ & \oplus \left[ \lambda I_{q_1} - J_{q_1}(\beta_1) \right] \oplus \cdots \oplus \left[ \lambda I_{q_f} - J_{q_f}(\beta_f) \right], \end{aligned}$$

where $L_k(\lambda)$ denotes the $k \times (k+1)$ singular pencil given by

$$L_k(\lambda) = \begin{bmatrix} \lambda & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \lambda & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \lambda & -1 \end{bmatrix}$$

and

$$m = \sum_{i=1}^{a} r_i + \sum_{i=1}^{b} m_i + \sum_{i=1}^{c} n_i, \qquad n = m + a,$$

$$q = \sum_{i=1}^{d} l_i + \sum_{i=1}^{e} p_i + \sum_{i=1}^{f} q_i, \qquad p = q + d.$$

From (3.1), Equation (1.4) is equivalent to

$$(\lambda C_0 - A_0)X_0B_0 - C_0X_0(\lambda B_0 - D_0) = -E_0, \tag{3.2}$$

where

$$Q_1X_0P_2 = X, \qquad E_0 = P_1EQ_2.$$

If we make partitions of $X_0$ and $E_0$ conforming with the blocks on the diagonal forms $\lambda C_0 - A_0$ and $\lambda B_0 - D_0$, Equation (3.2) reduces to the set of independent equations

$$[\lambda C_0 - A_0]_i[X_0]_{ij}[B_0]_j - [C_0]_i[X_0]_{ij}[\lambda B_0 - D_0]_j = -[E_0]_{ij}. \tag{3.3}$$

When the canonical blocks $[\lambda C_0 - A_0]_i$ and $[\lambda B_0 - D_0]_j$ are regular, we are in the situation of Section 2, and (3.3) has a unique solution under the assumption that the spectra of these pencils have an empty intersection. When one or both blocks are singular, we can apply the same technique described in [11] in order to reduce the problem to a regular one. For example, when both blocks are singular,

$$[\lambda C_0 - A_0]_i = L_{r_i}(\lambda), \qquad [\lambda B_0 - D_0]_j = L_{l_j}^{\mathrm{T}}(\lambda).$$

By deleting the first column in $L_{r_i}(\lambda)$ and the last row in $L_{l_j}^{\mathrm{T}}(\lambda)$, the truncated blocks $[\overline{\lambda C_0 - A_0}]_i$ and $[\overline{\lambda B_0 - D_0}]_j$ are regular with $r_i$ eigenvalues at $\infty$ and $l_j$ eigenvalues at 0 respectively. This corresponds to taking the first row and the last column in $[X_0]_{ij}$ equal to zero and solving for the truncated matrix $[\overline{X}_0]_{ij}$. In this way Equation (3.3) is replaced by

$$\left[\overline{\lambda C_0 - A_0}\right]_i\left[\overline{X_0}\right]_{ij}\left[\overline{B_0}\right]_j - \left[\overline{C_0}\right]_i\left[\overline{X_0}\right]_{ij}\left[\overline{\lambda B_0 - D_0}\right]_j = -[E_0]_{ij}, \tag{3.4}$$

where the regular pencils have disjoint spectra. This equation has a unique solution which also yields a solution to (3.3) on adding a zero row and column, to reconstruct $[X_0]_{ij}$.

Putting all these solutions together, we thus construct a nonunique matrix $X_0$ satisfying (3.2). Thus we obtain the following result.

THEOREM 2. *We consider the linear matrix equation*

$$AXB - CXD = E \tag{3.5}$$

*where* $A, C \in \mathbb{K}^{m\times n}$, $B, D \in \mathbb{K}^{p\times q}$, $E \in \mathbb{K}^{m\times q}$. *If* (iii) $\lambda C - A$ *and* $aB - D$ *are right and left invertible respectively, and* (iv) $\sigma(C, A) \cap \sigma(B, D) = \varnothing$, *then Equation* (3.5) *has a solution.*

*The authors would like to express their gratitude to Dr. P. Van Dooren for his helpful suggestion concerning the case of singular pencils.*

REFERENCES

1 K. E. Chu, Exclusion theorems for the generalized eigenvalue problem, Numerical Analysis Rpt. NA/11/85, Dept of Mathematics, Univ. of Reading, 1985.
2 M. A. Epton, Methods for the solution of $AXB - CXD = E$ and its application in the numerical solution of implicit ordinary differential equations, *BIT* 20:341–345 (1980).
3 D. J. Bender, Lyapunov-like equations and reachability/observability Gramians for descriptor systems, *Linear Algebra Appl.*, to appear.
4 F. R. Gantmacher, *Matrix Theory*, Vols. 1, 2, Chelsea, New York, 1977.
5 P. Rózsa, Lineare Matrizengleichungen und Kroneckersche Produkte, Z. *Angew. Math. Mech.* 58T:395–397 (1978).
6 P. Rózsa, Linear matrix equations and Kronecker products, in *Proceedings of the Fourth Symposium on Basic Problems of Numerical Mathematics*, Plzen, 1978, pp. 153–162.
7 Er – Chieh Ma, A finite series solution of the matrix equation $AX - XB = C$, *SIAM J. Appl. Math.* 14:490–495 (1966).
8 K. E. Chu, The Solution of the Matrix Equations $AXB - CXD = E$ and $(YA - DZ, YC - BZ) = (E, F)$. Numerical Analysis Rpt. NA/10/85, Dept. of Mathematics, Univ. of Reading, 1985.
9 F. L. Lewis, Adjoint matrix, Bezout theorem, Cayley-Hamilton theorem and Fadeev's method for the matrix pencil $(sE - A)$ in *Proceedings of the 22nd IEEE Conference on Decision and Control*, 1983, pp. 1282–1288.
10 A. Jameson, Solution of the equation $AX - XB = C$ by inversion of an $M \times M$ or $N \times N$ matrix, *SIAM J. Appl. Math.* 16:1020–1023 (1968).
11 P. Van Dooren, Reducing subspaces: Definitions, properties and algorithms, in *Matrix Pencils* (A. Ruche and B. Kagstrom, Eds.), Lecture Notes in Math., Springer, New York, 1983, pp. 58–73.
12 S. K. Mitra, The matrix equation $AXB + CXD = E$, *SIAM J. Appl. Math.* 32:823–825 (1977).

*Received 4 January 1988; final manuscript accepted 1 November 1988*