Chapter 13

Kronecker Products

13.1 Definition and Examples

Definition 13.1. *Let $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{p \times q}$. Then the* **Kronecker product** *(or tensor product) of A and B is defined as the matrix*

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \in \mathbb{R}^{mp \times nq}. \tag{13.1}$$

Obviously, the same definition holds if A and B are complex-valued matrices. We restrict our attention in this chapter primarily to real-valued matrices, pointing out the extension to the complex case only where it is not obvious.

Example 13.2.

1. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}$. Then

$$A \otimes B = \begin{bmatrix} B & 2B & 3B \\ 3B & 2B & B \end{bmatrix} = \begin{bmatrix} 2 & 1 & 4 & 2 & 6 & 3 \\ 2 & 3 & 4 & 6 & 6 & 9 \\ 6 & 3 & 4 & 2 & 2 & 1 \\ 6 & 9 & 4 & 6 & 2 & 3 \end{bmatrix}.$$

   Note that $B \otimes A \neq A \otimes B$.

2. For any $B \in \mathbb{R}^{p \times q}$, $I_2 \otimes B = \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix}$.
   Replacing I_2 by I_n yields a block diagonal matrix with n copies of B along the diagonal.

3. Let B be an arbitrary 2×2 matrix. Then

$$B \otimes I_2 = \begin{bmatrix} b_{11} & 0 & b_{12} & 0 \\ 0 & b_{11} & 0 & b_{12} \\ b_{21} & 0 & b_{22} & 0 \\ 0 & b_{21} & 0 & b_{22} \end{bmatrix}.$$

The extension to arbitrary B and I_n is obvious.

4. Let $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$. Then

$$\begin{aligned} x \otimes y &= \left[x_1 y^T, \ldots, x_m y^T\right]^T \\ &= [x_1 y_1, \ldots, x_1 y_n, x_2 y_1, \ldots, x_m y_n]^T \in \mathbb{R}^{mn}. \end{aligned}$$

5. Let $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$. Then

$$\begin{aligned} x \otimes y^T &= [x_1 y, \ldots, x_m y]^T \\ &= \begin{bmatrix} x_1 y_1 & \ldots & x_1 y_n \\ \vdots & \ddots & \vdots \\ x_m y_1 & \ldots & x_m y_n \end{bmatrix} \\ &= x y^T \in \mathbb{R}^{m \times n}. \end{aligned}$$

13.2 Properties of the Kronecker Product

Theorem 13.3. *Let $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{r \times s}$, $C \in \mathbb{R}^{n \times p}$, and $D \in \mathbb{R}^{s \times t}$. Then*

$$(A \otimes B)(C \otimes D) = AC \otimes BD \ \ (\in \mathbb{R}^{mr \times pt}). \tag{13.2}$$

Proof: Simply verify that

$$\begin{aligned} (A \otimes B)(C \otimes D) &= \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \begin{bmatrix} c_{11}D & \cdots & c_{1p}D \\ \vdots & \ddots & \vdots \\ c_{n1}D & \cdots & c_{np}D \end{bmatrix} \\ &= \begin{bmatrix} \sum_{k=1}^n a_{1k}c_{k1}BD & \cdots & \sum_{k=1}^n a_{1k}c_{kp}BD \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^n a_{mk}c_{k1}BD & \cdots & \sum_{k=1}^n a_{mk}c_{kp}BD \end{bmatrix} \\ &= AC \otimes BD. \qquad \square \end{aligned}$$

Theorem 13.4. *For all A and B, $(A \otimes B)^T = A^T \otimes B^T$.*

Proof: For the proof, simply verify using the definitions of transpose and Kronecker product. □

Corollary 13.5. *If $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{m \times m}$ are symmetric, then $A \otimes B$ is symmetric.*

Theorem 13.6. *If A and B are nonsingular, $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.*

Proof: Using Theorem 13.3, simply note that $(A \otimes B)(A^{-1} \otimes B^{-1}) = I \otimes I = I$. □

Theorem 13.7. *If $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times m}$ are normal, then $A \otimes B$ is normal.*

Proof:

$$\begin{aligned}(A \otimes B)^T(A \otimes B) &= (A^T \otimes B^T)(A \otimes B) \quad \text{by Theorem 13.4}\\ &= A^TA \otimes B^TB \quad \text{by Theorem 13.3}\\ &= AA^T \otimes BB^T \quad \text{since } A \text{ and } B \text{ are normal}\\ &= (A \otimes B)(A \otimes B)^T \quad \text{by Theorem 13.3.} \qquad \square\end{aligned}$$

Corollary 13.8. *If $A \in \mathbb{R}^{n\times n}$ is orthogonal and $B \in \mathbb{R}^{m\times m}$ is orthogonal, then $A \otimes B$ is orthogonal.*

Example 13.9. Let $A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$ and $B = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}$. Then it is easily seen that A is orthogonal with eigenvalues $e^{\pm j\theta}$ and B is orthogonal with eigenvalues $e^{\pm j\phi}$. The 4×4 matrix $A \otimes B$ is then also orthogonal with eigenvalues $e^{\pm j(\theta+\phi)}$ and $e^{\pm j(\theta-\phi)}$.

Theorem 13.10. *Let $A \in \mathbb{R}^{m\times n}$ have a singular value decomposition $U_A\Sigma_AV_A^T$ and let $B \in \mathbb{R}^{p\times q}$ have a singular value decomposition $U_B\Sigma_BV_B^T$. Then*

$$(U_A \otimes U_B)(\Sigma_A \otimes \Sigma_B)(V_A^T \otimes V_B^T)$$

yields a singular value decomposition of $A \otimes B$ (after a simple reordering of the diagonal elements of $\Sigma_A \otimes \Sigma_B$ and the corresponding right and left singular vectors).

Corollary 13.11. *Let $A \in \mathbb{R}_r^{m\times n}$ have singular values $\sigma_1 \geq \cdots \geq \sigma_r > 0$ and let $B \in \mathbb{R}_s^{p\times q}$ have singular values $\tau_1 \geq \cdots \geq \tau_s > 0$. Then $A \otimes B$ (or $B \otimes A$) has rs singular values $\sigma_1\tau_1 \geq \cdots \geq \sigma_r\tau_s > 0$ and*

$$\operatorname{rank}(A \otimes B) = (\operatorname{rank}A)(\operatorname{rank}B) = \operatorname{rank}(B \otimes A)\,.$$

Theorem 13.12. *Let $A \in \mathbb{R}^{n\times n}$ have eigenvalues $\lambda_i, i \in \underline{n}$, and let $B \in \mathbb{R}^{m\times m}$ have eigenvalues μ_j, $j \in \underline{m}$. Then the mn eigenvalues of $A \otimes B$ are*

$$\lambda_1\mu_1, \ldots, \lambda_1\mu_m, \lambda_2\mu_1, \ldots, \lambda_2\mu_m, \ldots, \lambda_n\mu_m.$$

Moreover, if $x_1, \ldots, x_p$ are linearly independent right eigenvectors of A corresponding to $\lambda_1, \ldots, \lambda_p$ ($p \leq n$), and $z_1, \ldots, z_q$ are linearly independent right eigenvectors of B corresponding to $\mu_1, \ldots, \mu_q$ ($q \leq m$), then $x_i \otimes z_j \in \mathbb{R}^{mn}$ are linearly independent right eigenvectors of $A \otimes B$ corresponding to $\lambda_i\mu_j$, $i \in \underline{p}$, $j \in \underline{q}$.

Proof: The basic idea of the proof is as follows:

$$\begin{aligned}(A \otimes B)(x \otimes z) &= Ax \otimes Bz\\ &= \lambda x \otimes \mu z\\ &= \lambda\mu(x \otimes z). \qquad \square\end{aligned}$$

If A and B are diagonalizable in Theorem 13.12, we can take $p = n$ and $q = m$ and thus get the complete eigenstructure of $A \otimes B$. In general, if A and B have Jordan form decompositions given by $P^{-1}AP = J_A$ and $Q^{-1}BQ = J_B$, respectively, then we get the following Jordan-like structure:

$$\begin{aligned}(P \otimes Q)^{-1}(A \otimes B)(P \otimes Q) &= (P^{-1} \otimes Q^{-1})(A \otimes B)(P \otimes Q)\\ &= (P^{-1}AP) \otimes (Q^{-1}BQ)\\ &= J_A \otimes J_B.\end{aligned}$$

Note that $J_A \otimes J_B$, while upper triangular, is generally not quite in Jordan form and needs further reduction (to an ultimate Jordan form that also depends on whether or not certain eigenvalues are zero or nonzero).

A Schur form for $A \otimes B$ can be derived similarly. For example, suppose P and Q are unitary matrices that reduce A and B, respectively, to Schur (triangular) form, i.e., $P^H AP = T_A$ and $Q^H BQ = T_B$ (and similarly if P and Q are orthogonal similarities reducing A and B to real Schur form). Then

$$\begin{aligned}(P \otimes Q)^H(A \otimes B)(P \otimes Q) &= (P^H \otimes Q^H)(A \otimes B)(P \otimes Q)\\ &= (P^H AP) \otimes (Q^H BQ)\\ &= T_A \otimes T_B.\end{aligned}$$

Corollary 13.13. *Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times m}$. Then*

1. $\mathrm{Tr}(A \otimes B) = (\mathrm{Tr}A)(\mathrm{Tr}B) = \mathrm{Tr}(B \otimes A)$.

2. $\det(A \otimes B) = (\det A)^m(\det B)^n = \det(B \otimes A)$.

Definition 13.14. *Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times m}$. Then the* **Kronecker sum** (*or tensor sum*) *of A and B, denoted $A \oplus B$, is the $mn \times mn$ matrix $(I_m \otimes A) + (B \otimes I_n)$. Note that, in general, $A \oplus B \neq B \oplus A$.*

Example 13.15.

1. Let

$$A = \begin{bmatrix} 1 & 2 & 3\\ 3 & 2 & 1\\ 1 & 1 & 4\end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & 1\\ 2 & 3\end{bmatrix}.$$

Then

$$A \oplus B = (I_2 \otimes A) + (B \otimes I_3) = \begin{bmatrix} 1 & 2 & 3 & 0 & 0 & 0\\ 3 & 2 & 1 & 0 & 0 & 0\\ 1 & 1 & 4 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 2 & 3\\ 0 & 0 & 0 & 3 & 2 & 1\\ 0 & 0 & 0 & 1 & 1 & 4\end{bmatrix} + \begin{bmatrix} 2 & 0 & 0 & 1 & 0 & 0\\ 0 & 2 & 0 & 0 & 1 & 0\\ 0 & 0 & 2 & 0 & 0 & 1\\ 2 & 0 & 0 & 3 & 0 & 0\\ 0 & 2 & 0 & 0 & 3 & 0\\ 0 & 0 & 2 & 0 & 0 & 3\end{bmatrix}.$$

The reader is invited to compute $B \oplus A = (I_3 \otimes B) + (A \otimes I_2)$ and note the difference with $A \oplus B$.

2. Recall the real JCF

$$J = \begin{bmatrix} M & I & 0 & \cdots & & 0 \\ 0 & M & I & 0 & & \vdots \\ \vdots & \ddots & M & \ddots & \ddots & \\ & & & \ddots & I & 0 \\ \vdots & & & \ddots & M & I \\ 0 & \cdots & & \cdots & 0 & M \end{bmatrix} \in \mathbb{R}^{2k\times 2k},$$

where $M = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$. Define

$$E_k = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & \ddots & 1 \\ 0 & \cdots & & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{k\times k}.$$

Then J can be written in the very compact form $J = (I_k \otimes M) + (E_k \otimes I_2) = M \oplus E_k$.

Theorem 13.16. *Let $A \in \mathbb{R}^{n\times n}$ have eigenvalues $\lambda_i, i \in \underline{n}$, and let $B \in \mathbb{R}^{m\times m}$ have eigenvalues $\mu_j, j \in \underline{m}$. Then the Kronecker sum $A \oplus B = (I_m \otimes A) + (B \otimes I_n)$ has mn eigenvalues*

$$\lambda_1 + \mu_1, \ldots, \lambda_1 + \mu_m, \lambda_2 + \mu_1, \ldots, \lambda_2 + \mu_m, \ldots, \lambda_n + \mu_m.$$

Moreover, if $x_1, \ldots, x_p$ are linearly independent right eigenvectors of A corresponding to $\lambda_1, \ldots, \lambda_p$ $(p \le n)$, and $z_1, \ldots, z_q$ are linearly independent right eigenvectors of B corresponding to $\mu_1, \ldots, \mu_q$ $(q \le m)$, then $z_j \otimes x_i \in \mathbb{R}^{mn}$ are linearly independent right eigenvectors of $A \oplus B$ corresponding to $\lambda_i + \mu_j$, $i \in \underline{p}$, $j \in \underline{q}$.

Proof: The basic idea of the proof is as follows:

$$\begin{aligned}[(I_m \otimes A) + (B \otimes I_n)](z \otimes x) &= (z \otimes Ax) + (Bz \otimes x) \\ &= (z \otimes \lambda x) + (\mu z \otimes x) \\ &= (\lambda + \mu)(z \otimes x). \qquad \square\end{aligned}$$

If A and B are diagonalizable in Theorem 13.16, we can take $p = n$ and $q = m$ and thus get the complete eigenstructure of $A \oplus B$. In general, if A and B have Jordan form decompositions given by $P^{-1}AP = J_A$ and $Q^{-1}BQ = J_B$, respectively, then

$$\begin{aligned}&[(Q \otimes I_n)(I_m \otimes P)]^{-1}[(I_m \otimes A) + (B \otimes I_n)][(Q \otimes I_n)(I_m \otimes P)] \\ &\quad = [(I_m \otimes P)^{-1}(Q \otimes I_n)^{-1}][(I_m \otimes A) + (B \otimes I_n)][(Q \otimes I_n)(I_m \otimes P)] \\ &\quad = [(I_m \otimes P^{-1})(Q^{-1} \otimes I_n)][(I_m \otimes A) + (B \otimes I_n)][(Q \otimes I_n)(I_m \otimes P)] \\ &\quad = (I_m \otimes J_A) + (J_B \otimes I_n)\end{aligned}$$

is a Jordan-like structure for $A \oplus B$.

A Schur form for $A \oplus B$ can be derived similarly. Again, suppose P and Q are unitary matrices that reduce A and B, respectively, to Schur (triangular) form, i.e., $P^H AP = T_A$ and $Q^H BQ = T_B$ (and similarly if P and Q are orthogonal similarities reducing A and B to real Schur form). Then

$$[(Q \otimes I_n)(I_m \otimes P)]^H [(I_m \otimes A) + (B \otimes I_n)][(Q \otimes I_n)(I_m \otimes P)] = (I_m \otimes T_A) + (T_B \otimes I_n),$$

where $[(Q \otimes I_n)(I_m \otimes P)] = (Q \otimes P)$ is unitary by Theorem 13.3 and Corollary 13.8.

13.3 Application to Sylvester and Lyapunov Equations

In this section we study the linear matrix equation

$$AX + XB = C, \tag{13.3}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times m}$, and $C \in \mathbb{R}^{n\times m}$. This equation is now often called a **Sylvester equation** in honor of J.J. Sylvester who studied general linear matrix equations of the form

$$\sum_{i=1}^{k} A_i X B_i = C.$$

A special case of (13.3) is the symmetric equation

$$AX + XA^T = C \tag{13.4}$$

obtained by taking $B = A^T$. When C is symmetric, the solution $X \in \mathbb{R}^{n\times n}$ is easily shown also to be symmetric and (13.4) is known as a **Lyapunov equation**. Lyapunov equations arise naturally in stability theory.

The first important question to ask regarding (13.3) is, When does a solution exist? By writing the matrices in (13.3) in terms of their columns, it is easily seen by equating the ith columns that

$$Ax_i + Xb_i = c_i = Ax_i + \sum_{j=1}^{m} b_{ji} x_j.$$

These equations can then be rewritten as the $mn \times mn$ linear system

$$\begin{bmatrix} A + b_{11}I & b_{21}I & \cdots & b_{m1}I \\ b_{12}I & A + b_{22}I & \cdots & b_{m2}I \\ \vdots & & \ddots & \vdots \\ b_{1m}I & b_{2m}I & \cdots & A + b_{mm}I \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix}. \tag{13.5}$$

The coefficient matrix in (13.5) clearly can be written as the Kronecker sum $(I_m \otimes A) + (B^T \otimes I_n)$. The following definition is very helpful in completing the writing of (13.5) as an "ordinary" linear system.

Definition 13.17. *Let $c_i \in \mathbb{R}^n$ denote the columns of $C \in \mathbb{R}^{n \times m}$ so that $C = [c_1, \ldots, c_m]$. Then* $\mathrm{vec}(C)$ *is defined to be the mn-vector formed by stacking the columns of C on top of one another, i.e.,* $\mathrm{vec}(C) = \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} \in \mathbb{R}^{mn}$.

Using Definition 13.17, the linear system (13.5) can be rewritten in the form

$$[(I_m \otimes A) + (B^T \otimes I_n)]\mathrm{vec}(X) = \mathrm{vec}(C). \tag{13.6}$$

There exists a unique solution to (13.6) if and only if $[(I_m \otimes A) + (B^T \otimes I_n)]$ is nonsingular. But $[(I_m \otimes A) + (B^T \otimes I_n)]$ is nonsingular if and only if it has no zero eigenvalues. From Theorem 13.16, the eigenvalues of $[(I_m \otimes A) + (B^T \otimes I_n)]$ are $\lambda_i + \mu_j$, where $\lambda_i \in \Lambda(A)$, $i \in \underline{n}$, and $\mu_j \in \Lambda(B)$, $j \in \underline{m}$. We thus have the following theorem.

Theorem 13.18. *Let $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{m \times m}$, and $C \in \mathbb{R}^{n \times m}$. Then the Sylvester equation*

$$AX + XB = C \tag{13.7}$$

has a unique solution if and only if A and $-B$ have no eigenvalues in common.

Sylvester equations of the form (13.3) (or symmetric Lyapunov equations of the form (13.4)) are generally not solved using the $mn \times mn$ "vec" formulation (13.6). The most commonly preferred numerical algorithm is described in [2]. First A and B are reduced to (real) Schur form. An equivalent linear system is then solved in which the triangular form of the reduced A and B can be exploited to solve successively for the columns of a suitably transformed solution matrix X. Assuming that, say, $n \geq m$, this algorithm takes only $O(n^3)$ operations rather than the $O(n^6)$ that would be required by solving (13.6) directly with Gaussian elimination. A further enhancement to this algorithm is available in [6] whereby the larger of A or B is initially reduced only to upper Hessenberg rather than triangular Schur form.

The next few theorems are classical. They culminate in Theorem 13.24, one of many elegant connections between matrix theory and stability theory for differential equations.

Theorem 13.19. *Let $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{m \times m}$, and $C \in \mathbb{R}^{n \times m}$. Suppose further that A and B are* **asymptotically stable** *(a matrix is asymptotically stable if all its eigenvalues have real parts in the open left half-plane). Then the (unique) solution of the Sylvester equation*

$$AX + XB = C \tag{13.8}$$

can be written as

$$X = -\int_0^{+\infty} e^{tA} C e^{tB}\, dt. \tag{13.9}$$

Proof: Since A and B are stable, $\lambda_i(A) + \lambda_j(B) \neq 0$ for all i, j so there exists a unique solution to (13.8) by Theorem 13.18. Now integrate the differential equation $\dot{X} = AX + XB$ (with $X(0) = C$) on $[0, +\infty)$:

$$\lim_{t \to +\infty} X(t) - X(0) = A \int_0^{+\infty} X(t)\, dt + \left(\int_0^{+\infty} X(t)\, dt \right) B. \tag{13.10}$$

Using the results of Section 11.1.6, it can be shown easily that $\lim_{t\to+\infty} e^{tA} = \lim_{t\to+\infty} e^{tB} = 0$. Hence, using the solution $X(t) = e^{tA}Ce^{tB}$ from Theorem 11.6, we have that $\lim_{t\to+\infty} X(t) = 0$. Substituting in (13.10) we have

$$-C = A\left(\int_0^{+\infty} e^{tA}Ce^{tB}\,dt\right) + \left(\int_0^{+\infty} e^{tA}Ce^{tB}\,dt\right)B$$

and so $X = -\int_0^{+\infty} e^{tA}Ce^{tB}\,dt$ satisfies (13.8). □

Remark 13.20. An equivalent condition for the existence of a unique solution to $AX + XB = C$ is that $\left[\begin{smallmatrix} A & C \\ 0 & -B \end{smallmatrix}\right]$ be similar to $\left[\begin{smallmatrix} A & 0 \\ 0 & -B \end{smallmatrix}\right]$ (via the similarity $\left[\begin{smallmatrix} I & X \\ 0 & -I \end{smallmatrix}\right]$).

Theorem 13.21. *Let $A, C \in \mathbb{R}^{n\times n}$. Then the Lyapunov equation*

$$AX + XA^T = C \tag{13.11}$$

has a unique solution if and only if A and $-A^T$ have no eigenvalues in common. If C is symmetric and (13.11) has a unique solution, then that solution is symmetric.

Remark 13.22. If the matrix $A \in \mathbb{R}^{n\times n}$ has eigenvalues $\lambda_1, \ldots, \lambda_n$, then $-A^T$ has eigenvalues $-\lambda_1, \ldots, -\lambda_n$. Thus, a sufficient condition that guarantees that A and $-A^T$ have no common eigenvalues is that A be asymptotically stable. Many useful results exist concerning the relationship between stability and Lyapunov equations. Two basic results due to Lyapunov are the following, the first of which follows immediately from Theorem 13.19.

Theorem 13.23. *Let $A, C \in \mathbb{R}^{n\times n}$ and suppose further that A is asymptotically stable. Then the (unique) solution of the Lyapunov equation*

$$AX + XA^T = C$$

can be written as

$$X = -\int_0^{+\infty} e^{tA}Ce^{tA^T}\,dt. \tag{13.12}$$

Theorem 13.24. *A matrix $A \in \mathbb{R}^{n\times n}$ is asymptotically stable if and only if there exists a positive definite solution to the Lyapunov equation*

$$AX + XA^T = C, \tag{13.13}$$

where $C = C^T < 0$.

Proof: Suppose A is asymptotically stable. By Theorems 13.21 and 13.23 a solution to (13.13) exists and takes the form (13.12). Now let v be an arbitrary nonzero vector in $\mathbb{R}^n$. Then

$$v^T Xv = \int_0^{+\infty} (v^T e^{tA})(-C)(v^T e^{tA})^T\,dt.$$

Since $-C > 0$ and e^{tA} is nonsingular for all t, the integrand above is positive. Hence $v^T X v > 0$ and thus X is positive definite.

Conversely, suppose $X = X^T > 0$ and let $\lambda \in \Lambda(A)$ with corresponding left eigenvector y. Then

$$\begin{aligned} 0 > y^H C y &= y^H A X y + y^H X A^T y \\ &= (\lambda + \bar{\lambda}) y^H X y. \end{aligned}$$

Since $y^H X y > 0$, we must have $\lambda + \bar{\lambda} = 2 \operatorname{Re} \lambda < 0$. Since λ was arbitrary, A must be asymptotically stable. □

Remark 13.25. The Lyapunov equation $AX + XA^T = C$ can also be written using the vec notation in the equivalent form

$$[(I \otimes A) + (A \otimes I)]\mathrm{vec}(X) = \mathrm{vec}(C).$$

A subtle point arises when dealing with the "dual" Lyapunov equation $A^T X + XA = C$. The equivalent "vec form" of this equation is

$$[(I \otimes A^T) + (A^T \otimes I)]\mathrm{vec}(X) = \mathrm{vec}(C).$$

However, the complex-valued equation $A^H X + XA = C$ is equivalent to

$$[(I \otimes A^H) + (A^T \otimes I)]\mathrm{vec}(X) = \mathrm{vec}(C).$$

The vec operator has many useful properties, most of which derive from one key result.

Theorem 13.26. *For any three matrices A, B, and C for which the matrix product ABC is defined,*

$$\mathrm{vec}(ABC) = (C^T \otimes A)\mathrm{vec}(B).$$

Proof: The proof follows in a fairly straightforward fashion either directly from the definitions or from the fact that $\mathrm{vec}(xy^T) = y \otimes x$. □

An immediate application is to the derivation of existence and uniqueness conditions for the solution of the simple Sylvester-like equation introduced in Theorem 6.11.

Theorem 13.27. *Let $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{p \times q}$, and $C \in \mathbb{R}^{m \times q}$. Then the equation*

$$AXB = C \tag{13.14}$$

has a solution $X \in \mathbb{R}^{n \times p}$ if and only if $AA^+CB^+B = C$, in which case the general solution is of the form

$$X = A^+CB^+ + Y - A^+AYBB^+, \tag{13.15}$$

where $Y \in \mathbb{R}^{n \times p}$ is arbitrary. The solution of (13.14) is unique if $BB^+ \otimes A^+A = I$.

Proof: Write (13.14) as

$$(B^T \otimes A)\mathrm{vec}(X) = \mathrm{vec}(C) \tag{13.16}$$

by Theorem 13.26. This "vector equation" has a solution if and only if

$$(B^T \otimes A)(B^T \otimes A)^+ \mathrm{vec}(C) = \mathrm{vec}(C).$$

It is a straightforward exercise to show that $(M \otimes N)^+ = M^+ \otimes N^+$. Thus, (13.16) has a solution if and only if

$$\begin{aligned}\mathrm{vec}(C) &= (B^T \otimes A)((B^+)^T \otimes A^+)\mathrm{vec}(C) \\ &= [(B^+B)^T \otimes AA^+]\mathrm{vec}(C) \\ &= \mathrm{vec}(AA^+CB^+B)\end{aligned}$$

and hence if and only if $AA^+CB^+B = C$.

The general solution of (13.16) is then given by

$$\mathrm{vec}(X) = (B^T \otimes A)^+ \mathrm{vec}(C) + [I - (B^T \otimes A)^+(B^T \otimes A)]\mathrm{vec}(Y),$$

where Y is arbitrary. This equation can then be rewritten in the form

$$\mathrm{vec}(X) = ((B^+)^T \otimes A^+)\mathrm{vec}(C) + [I - (BB^+)^T \otimes A^+A]\mathrm{vec}(Y)$$

or, using Theorem 13.26,

$$X = A^+CB^+ + Y - A^+AYBB^+.$$

The solution is clearly unique if $BB^+ \otimes A^+A = I$. □

EXERCISES

1. For any two matrices A and B for which the indicated matrix product is defined, show that $(\mathrm{vec}(A))^T(\mathrm{vec}(B)) = \mathrm{Tr}(A^TB)$. In particular, if $B \in \mathbb{R}^{n\times n}$, then $\mathrm{Tr}(B) = \mathrm{vec}(I_n)^T\mathrm{vec}(B)$.

2. Prove that for all matrices A and B, $(A \otimes B)^+ = A^+ \otimes B^+$.

3. Show that the equation $AXB = C$ has a solution for all C if A has full row rank and B has full column rank. Also, show that a solution, if it exists, is unique if A has full column rank and B has full row rank. What is the solution in this case?

4. Show that the general linear equation

$$\sum_{i=1}^{k} A_iXB_i = C$$

can be written in the form

$$[B_1^T \otimes A_1 + \cdots + B_k^T \otimes A_k]\mathrm{vec}(X) = \mathrm{vec}(C).$$

5. Let $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$. Show that $x^T \otimes y = yx^T$.

6. Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times m}$.

   (a) Show that $\|A \otimes B\|_2 = \|A\|_2\|B\|_2$.

   (b) What is $\|A \otimes B\|_F$ in terms of the Frobenius norms of A and B? Justify your answer carefully.

   (c) What is the spectral radius of $A \otimes B$ in terms of the spectral radii of A and B? Justify your answer carefully.

7. Let $A, B \in \mathbb{R}^{n\times n}$.

   (a) Show that $(I \otimes A)^k = I \otimes A^k$ and $(B \otimes I)^k = B^k \otimes I$ for all integers k.

   (b) Show that $e^{I\otimes A} = I \otimes e^A$ and $e^{B\otimes I} = e^B \otimes I$.

   (c) Show that the matrices $I \otimes A$ and $B \otimes I$ commute.

   (d) Show that
   $$e^{A\oplus B} = e^{(I\otimes A)+(B\otimes I)} = e^B \otimes e^A .$$
   (Note: This result would look a little "nicer" had we defined our Kronecker sum the other way around. However, Definition 13.14 is conventional in the literature.)

8. Consider the Lyapunov matrix equation (13.11) with
   $$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$
   and C the symmetric matrix
   $$\begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}.$$
   Clearly
   $$X_{\rm s} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
   is a symmetric solution of the equation. Verify that
   $$X_{\rm ns} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$
   is also a solution and is nonsymmetric. Explain in light of Theorem 13.21.

9. **Block Triangularization:** Let
   $$S = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$
   where $A \in \mathbb{R}^{n\times n}$ and $D \in \mathbb{R}^{m\times m}$. It is desired to find a similarity transformation of the form
   $$T = \begin{bmatrix} I & 0 \\ X & I \end{bmatrix}$$
   such that $T^{-1}ST$ is block upper triangular.

(a) Show that S is similar to

$$\begin{bmatrix} A + BX & B \\ 0 & D - XB \end{bmatrix}$$

if X satisfies the so-called **matrix Riccati equation**

$$C - XA + DX - XBX = 0.$$

(b) Formulate a similar result for block lower triangularization of S.

10. **Block Diagonalization:** Let

$$S = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix},$$

where $A \in \mathbb{R}^{n \times n}$ and $D \in \mathbb{R}^{m \times m}$. It is desired to find a similarity transformation of the form

$$T = \begin{bmatrix} I & Y \\ 0 & I \end{bmatrix}$$

such that $T^{-1}ST$ is block diagonal.

(a) Show that S is similar to

$$\begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}$$

if Y satisfies the Sylvester equation

$$AY - YD = -B.$$

(b) Formulate a similar result for block diagonalization of

$$S = \begin{bmatrix} A & 0 \\ C & D \end{bmatrix}.$$