Chapter 20
The Kronecker Product and Linear Matrix Equations

Many applications, in particular the stability analysis of differential equations, lead to linear matrix equations, such as $AX + XB = C$. Here the matrices A, B, C are given and the goal is to determine a matrix X that solves the equation (we will give a formal definition below). In the description of the solutions of such equations, the Kronecker product,[1] another product of matrices, is useful. In this chapter we develop the most important properties of this products and we study its application in the context of linear matrix equations. Many more results on this topic can be found in the books [HorJ91, LanT85].

Definition 20.1 If K is a field, $A = [a_{ij}] \in K^{m,m}$ and $B \in K^{n,n}$, then

$$A \otimes B := [a_{ij} B] = \begin{bmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{mm}B \end{bmatrix},$$

is called the *Kronecker product* of A and B.

The Kronecker product is sometimes called the *tensor product* of matrices. This product defines a map from $K^{m,m} \times K^{n,n}$ to $K^{mn,mn}$. The definition can be extended to non-square matrices, but for simplicity we consider here only the case of square matrices. The following lemma presents the basic computational rules of the Kronecker product.

Lemma 20.2 *For all square matrices A, B, C over K, the following computational rules hold:*

(1) $A \otimes (B \otimes C) = (A \otimes B) \otimes C$.

[1]Leopold Kronecker (1832–1891) is said to have used this product in his lectures in Berlin in the 1880s. It was defined formally for the first time in 1858 by Johann Georg Zehfuss (1832–1901).

J. Liesen and V. Mehrmann, *Linear Algebra*, Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-24346-7_20

(2) $(\mu A) \otimes B = A \otimes (\mu B) = \mu(A \otimes B)$ *for all* $\mu \in K$.
(3) $(A + B) \otimes C = (A \otimes C) + (B \otimes C)$*, whenever* $A + B$ *is defined.*
(4) $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$*, whenever* $B + C$ *is defined.*
(5) $(A \otimes B)^T = A^T \otimes B^T$*, and therefore the Kronecker product of two symmetric matrices is symmetric.*

Proof Exercise. □

In particular, in contrast to the standard matrix multiplication, the order of the factors in the Kronecker product does not change under transposition. The following result describes the matrix multiplication of two Kronecker products.

Lemma 20.3 *For* $A, C \in K^{m,m}$ *and* $B, D \in K^{n,n}$ *we have*

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD).$$

Hence, in particular,

(1) $A \otimes B = (A \otimes I_n)(I_m \otimes B) = (I_m \otimes B)(A \otimes I_n)$,
(2) $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$*, if* A *and* B *are invertible.*

Proof Since $A \otimes B = [a_{ij}B]$ and $C \otimes D = [c_{ij}D]$, the block $F_{ij} \in K^{n,n}$ in the block matrix $[F_{ij}] = (A \otimes B)(C \otimes D)$ is given by

$$F_{ij} = \sum_{k=1}^{m} (a_{ik}B)(c_{kj}D) = \sum_{k=1}^{m} a_{ik}c_{kj}\, BD = \Big(\sum_{k=1}^{m} a_{ik}c_{kj}\Big) BD.$$

For the block matrix $[G_{ij}] = (AC) \otimes (BD)$ with $G_{ij} \in K^{n,n}$ we obtain

$$G_{ij} = g_{ij}BD, \quad \text{where} \quad g_{ij} = \sum_{k=1}^{m} a_{ik}c_{kj},$$

which shows $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$. Now (1) and (2) easily follow from this equation. □

In general the Kronecker product is non-commutative (cp. Exercise 20.2), but we have the following relationship between $A \otimes B$ and $B \otimes A$.

Lemma 20.4 *For* $A \in K^{m,m}$ *and* $B \in K^{n,n}$ *there exists a permutation matrix* $P \in K^{mn,mn}$ *with*

$$P^T(A \otimes B)P = B \otimes A.$$

Proof Exercise. □

For the computation of the determinant, trace and rank of a Kronecker product there exist simple formulas.

Theorem 20.5 *For $A \in K^{m,m}$ and $B \in K^{n,n}$ the following rules hold:*

(1) $\det(A \otimes B) = (\det A)^n\,(\det B)^m = \det(B \otimes A)$.
(2) $\operatorname{trace}(A \otimes B) = \operatorname{trace}(A)\operatorname{trace}(B) = \operatorname{trace}(B \otimes A)$.
(3) $\operatorname{rank}(A \otimes B) = \operatorname{rank}(A)\operatorname{rank}(B) = \operatorname{rank}(B \otimes A)$.

Proof (1) From (1) in Lemma 20.3 and the multiplication theorem for determinants (cp. Theorem 7.15) we get

$$\det(A \otimes B) = \det\left((A \otimes I_n)\,(I_m \otimes B)\right) = \det(A \otimes I_n)\,\det(I_m \otimes B).$$

By Lemma 20.4 there exists a permutation matrix P with $A \otimes I_n = P(I_n \otimes A)P^T$. This implies that

$$\det(A \otimes I_n) = \det\left(P(I_n \otimes A)P^T\right) = \det(I_n \otimes A) = (\det A)^n.$$

Since $\det(I_m \otimes B) = (\det B)^m$, it then follows that $\det(A \otimes B) = (\det A)^n\,(\det B)^m$, and therefore also $\det(A \otimes B) = \det(B \otimes A)$.

(2) From $(A \otimes B) = [a_{ij}B]$ we obtain

$$\begin{aligned}\operatorname{trace}(A \otimes B) &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ii}b_{jj} = \Big(\sum_{i=1}^{m} a_{ii}\Big)\Big(\sum_{j=1}^{n} b_{jj}\Big) = \operatorname{trace}(A)\operatorname{trace}(B)\\ &= \operatorname{trace}(B)\operatorname{trace}(A) = \operatorname{trace}(B \otimes A).\end{aligned}$$

(3) Exercise. □

For a matrix $A = [a_1, \ldots, a_n] \in K^{m,n}$ with columns $a_j \in K^{m,1}$, $j = 1, \ldots, n$, we define

$$\operatorname{vec}(A) := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in K^{mn,1}.$$

The application of vec turns the matrix A into a "column vector" and thus "vectorizes" A.

Lemma 20.6 *The map* $\operatorname{vec} : K^{m,n} \to K^{mn,1}$ *is an isomorphism. In particular, $A_1, \ldots, A_k \in K^{m,n}$ are linearly independent if and only if* $\operatorname{vec}(A_1), \ldots, \operatorname{vec}(A_k) \in K^{mn,1}$ *are linearly independent.*

Proof Exercise. □

We now consider the relationship between the Kronecker product and the vec map.

Theorem 20.7 *For $A \in K^{m,m}$, $B \in K^{n,n}$ and $C \in K^{m,n}$ we have*

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

Hence, in particular,

(1) $\mathrm{vec}(AC) = (I_n \otimes A)\mathrm{vec}(C)$ *and* $\mathrm{vec}(CB) = (B^T \otimes I_m)\mathrm{vec}(C)$,
(2) $\mathrm{vec}(AC + CB) = \big((I_n \otimes A) + (B^T \otimes I_m)\big)\,\mathrm{vec}(C)$.

Proof For $j = 1, \ldots, n$, the jth column of ACB is given by

$$\begin{aligned}(ACB)e_j = (AC)(Be_j) &= \sum_{k=1}^{n} b_{kj}(AC)e_k = \sum_{k=1}^{n}(b_{kj}A)(Ce_k) \\ &= [\, b_{1j}A,\ b_{2j}A,\ \ldots,\ b_{nj}A \,]\,\mathrm{vec}(C),\end{aligned}$$

which implies that $\mathrm{vec}(ACB) = (B^T \otimes A)\mathrm{vec}(C)$. With $B = I_n$ resp. $A = I_m$ we obtain (1), while (1) and the linearity of vec yield (2). □

In order to study the relationship between the eigenvalues of the matrices A, B and those of the Kronecker product $A \otimes B$, we use *bivariate polynomials*, i.e., polynomials in two variables (cp. Exercise 9.10). If

$$p(t_1, t_2) = \sum_{i,j=0}^{l} \alpha_{ij} t_1^i t_2^j \ \in\ K[t_1, t_2]$$

is such a polynomial, then for $A \in K^{m,m}$ and $B \in K^{n,n}$ we define the matrix

$$p(A, B) := \sum_{i,j=0}^{l} \alpha_{ij} A^i \otimes B^j. \qquad (20.1)$$

Here we have to be careful with the order of the factors, since in general $A^i \otimes B^j \neq B^j \otimes A^i$ (cp. Exercise 20.2).

Example 20.8 For $A \in \mathbb{R}^{m,m}$, $B \in \mathbb{R}^{n,n}$ and $p(t_1, t_2) = 2t_1 + 3t_1t_2^2 = 2t_1^1t_2^0 + 3t_1^1t_2^2 \in \mathbb{R}[t_1, t_2]$ we get the matrix $p(A, B) = 2A \otimes I_n + 3A \otimes B^2$.

The following result is known as *Stephanos' theorem*.[2]

[2]Named after Cyparissos Stephanos (1857–1917) who in 1900 showed besides this result also the assertion of Lemma 20.3.

Theorem 20.9 *Let $A \in K^{m,m}$ and $B \in K^{n,n}$ be two matrices that have Jordan normal forms and the eigenvalues $\lambda_1, \dots, \lambda_m \in K$ and $\mu_1, \dots, \mu_n \in K$, respectively. If $p(A, B)$ is defined as in (20.1), then the following assertions hold:*

(1) The eigenvalues of $p(A, B)$ are $p(\lambda_k, \mu_\ell)$ for $k = 1, \dots, m$ and $\ell = 1, \dots, n$.
(2) The eigenvalues of $A \otimes B$ are $\lambda_k \cdot \mu_\ell$ for $k = 1, \dots, m$ and $\ell = 1, \dots, n$.
(3) The eigenvalues of $A \otimes I_n + I_m \otimes B$ are $\lambda_k + \mu_\ell$ for $k = 1, \dots, m$ and $\ell = 1, \dots, n$.

Proof Let $S \in GL_m(K)$ and $T \in GL_n(K)$ be such that $S^{-1}AS = J_A$ and $T^{-1}BT = J_B$ are in Jordan canonical form. The matrices J_A and J_B are upper triangular. Thus, for all $i, j \in \mathbb{N}_0$ the matrices J_A^i, J_B^j and $J_A^i \otimes J_B^j$ are upper triangular. The eigenvalues of J_A^i and J_B^j are $\lambda_1^i, \dots, \lambda_m^i$ and $\mu_1^j, \dots, \mu_n^j$, respectively. Thus, $p(\lambda_k, \mu_\ell)$, $k = 1, \dots, m$, $\ell = 1, \dots, n$, are the diagonal entries of the matrix $p(J_A, J_B)$. Using Lemma 20.3 we obtain

$$\begin{aligned}
p(A, B) &= \sum_{i,j=0}^{l} \alpha_{ij} (S J_A S^{-1})^i \otimes (T J_B T^{-1})^j = \sum_{i,j=0}^{l} \alpha_{ij} (S J_A^i S^{-1}) \otimes (T J_B^j T^{-1}) \\
&= \sum_{i,j=0}^{l} \alpha_{ij} \Big((S J_A^i) \otimes (T J_B^j) \Big) (S^{-1} \otimes T^{-1}) \\
&= \sum_{i,j=0}^{l} \alpha_{ij} (S \otimes T)(J_A^i \otimes J_B^j)(S \otimes T)^{-1} \\
&= (S \otimes T) \left(\sum_{i,j=0}^{l} \alpha_{ij} (J_A^i \otimes J_B^j) \right) (S \otimes T)^{-1} \\
&= (S \otimes T) p(J_A, J_B) (S \otimes T)^{-1},
\end{aligned}$$

which implies (1).

The assertions (2) and (3) follow from (1) with $p(t_1, t_2) = t_1 t_2$ and $p(t_1, t_2) = t_1 + t_2$, respectively. □

The following result on the matrix exponential function of a Kronecker product is helpful in applications that involve systems of linear differential equations.

Lemma 20.10 *For $A \in \mathbb{C}^{m,m}$, $B \in \mathbb{C}^{n,n}$ and $C := (A \otimes I_n) + (I_m \otimes B)$ we have*

$$\exp(C) = \exp(A) \otimes \exp(B).$$

Proof From Lemma 20.3 we know that the matrices $A \otimes I_n$ and $I_m \otimes B$ commute. Using Lemma 17.6 we obtain

$$\begin{aligned}
\exp(C) &= \exp(A \otimes I_n + I_m \otimes B) = \exp(A \otimes I_n)\, \exp(I_m \otimes B) \\
&= \left(\sum_{j=0}^{\infty} \frac{1}{j!} (A \otimes I_n)^j \right) \left(\sum_{i=0}^{\infty} \frac{1}{i!} (I_m \otimes B)^i \right) \\
&= \sum_{j=0}^{\infty} \frac{1}{j!} \sum_{i=0}^{\infty} \frac{1}{i!} (A \otimes I_n)^j \, (I_m \otimes B)^i \\
&= \sum_{j=0}^{\infty} \frac{1}{j!} \sum_{i=0}^{\infty} \frac{1}{i!} (A^j \otimes B^i) \\
&= \exp(A) \otimes \exp(B),
\end{aligned}$$

where we have used the properties of the matrix exponential series (cp. Sect. 17.1). □

For given matrices $A_j \in K^{m,m}$, $B_j \in K^{n,n}$, $j = 1, \ldots, q$, and $C \in K^{m,n}$ an equation of the form

$$A_1 X B_1 + A_2 X B_2 + \ldots + A_q X B_q = C \tag{20.2}$$

is called a *linear matrix equation* for the unknown matrix $X \in K^{m,n}$.

Theorem 20.11 *A matrix $\widehat{X} \in K^{m,n}$ solves (20.2) if and only if $\widehat{x} := \mathrm{vec}(\widehat{X}) \in K^{mn,1}$ solves the linear system of equations*

$$Gx = \mathrm{vec}(C), \quad \textit{where} \quad G := \sum_{j=1}^{q} B_j^T \otimes A_j.$$

Proof Exercise. □

We now consider two special cases of (20.2).

Theorem 20.12 *For $A \in \mathbb{C}^{m,m}$, $B \in \mathbb{C}^{n,n}$ and $C \in \mathbb{C}^{m,n}$ the* Sylvester equation[3]

$$AX + XB = C \tag{20.3}$$

has a unique solution if and only if A and $-B$ have no common eigenvalue. If all eigenvalues of A and B have negative real parts, then the unique solution of (20.3) is given by

$$\widehat{X} = -\int_0^{\infty} \exp(tA) C \exp(tB) dt.$$

(As in Sect. 17.2 the integral is defined entrywise.)

[3]James Joseph Sylvester (1814–1897).

Proof Analogous to the representation in Theorem 20.11, we can write the Sylvester equation (20.3) as

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

If A and B have the eigenvalues $\lambda_1, \ldots, \lambda_m$ and $\mu_1, \ldots, \mu_n$, respectively, then $G = I_n \otimes A + B^T \otimes I_m$ by (3) in Theorem 20.9 has the eigenvalues $\lambda_k + \mu_\ell$, $k = 1, \ldots, m$, $\ell = 1, \ldots, n$. Thus, G is invertible, and the Sylvester equation is uniquely solvable, if and only if $\lambda_k + \mu_\ell \neq 0$ for all $k = 1, \ldots, m$ and $\ell = 1, \ldots, n$.

Let A and B be matrices with eigenvalues that have negative real parts. Then A and $-B$ have no common eigenvalues and (20.3) has a unique solution. Let $J_A = S^{-1}AS$ and $J_B = T^{-1}BT$ be Jordan canonical forms of A and B. We consider the linear differential equation

$$\frac{dZ}{dt} = AZ + ZB, \quad Z(0) = C, \tag{20.4}$$

that is solved by the function

$$Z : [0, \infty) \to \mathbb{C}^{m,n}, \quad Z(t) := \exp(tA)C\exp(tB)$$

(cp. Exercise 20.10). This function satisfies

$$\begin{aligned}\lim_{t\to\infty} Z(t) &= \lim_{t\to\infty} \exp(tA)C\exp(tB)\\ &= \lim_{t\to\infty} S\underbrace{\exp(tJ_A)}_{\to 0}\underbrace{S^{-1}CT}_{\text{constant}}\underbrace{\exp(tJ_B)}_{\to 0}T^{-1} = 0.\end{aligned}$$

Integration of equation (20.4) from $t = 0$ to $t = \infty$ yields

$$-C = -Z(0) = \lim_{t\to\infty}(Z(t) - Z(0)) = A\int_0^\infty Z(t)dt + \left(\int_0^\infty Z(t)dt\right)B.$$

(Here we use without proof the existence of the infinite integrals.) This implies that

$$\widehat{X} := -\int_0^\infty Z(t)dt = -\int_0^\infty \exp(tA)C\exp(tB)dt$$

is the unique solution of (20.3). □

Theorem 20.12 also gives the solution of another important matrix equation.

Corollary 20.13 *For $A, C \in \mathbb{C}^{n,n}$ the* Lyapunov equation[4]

$$AX + XA^H = -C \tag{20.5}$$

[4] Alexandr Mikhailovich Lyapunov (also Ljapunov or Liapunov; 1857–1918).

has a unique solution $\widehat{X} \in \mathbb{C}^{n,n}$ if the eigenvalues of A have negative real parts. If, furthermore, C is Hermitian positive definite, then also $\widehat{X}$ is Hermitian positive definite.

Proof Since by assumption A and $-A^H$ have no common eigenvalues, the unique solvability of (20.5) follows from Theorem 20.12, and the solution is given by the matrix

$$\widehat{X} = -\int_0^\infty \exp(tA)(-C)\exp\left(tA^H\right)dt = \int_0^\infty \exp(tA)C\exp\left(tA^H\right)dt.$$

If C is Hermitian positive definite, then $\widehat{X}$ is Hermitian and for $x \in \mathbb{C}^{n,1} \setminus \{0\}$ we have

$$x^H\widehat{X}x = x^H\left(\int_0^\infty \exp(tA)C\exp\left(tA^H\right)dt\right)x = \int_0^\infty \underbrace{x^H\exp(tA)C\exp\left(tA^H\right)x}_{>0}\,dt > 0.$$

The last inequality follows from the monotonicity of the integral and the fact that for $x \neq 0$ also $\exp(tA^H)x \neq 0$, since $\exp\left(tA^H\right)$ is invertible for every real t. □

Exercises

20.1 Prove Lemma 20.2.
20.2 Construct two square matrices A, B with $A \otimes B \neq B \otimes A$.
20.3 Prove Lemma 20.4.
20.4 Prove Theorem 20.5 (3).
20.5 Prove Lemma 20.6.
20.6 Show that $A \otimes B$ is normal if $A \in \mathbb{C}^{m,m}$ and $B \in \mathbb{C}^{n,n}$ are normal. Is it true that if $A \otimes B$ is unitary, then A and B are unitary?
20.7 Use the singular value decompositions of $A = V_A\Sigma_A W_A^H \in \mathbb{C}^{m,m}$ and $B = V_B\Sigma_B W_B^H \in \mathbb{C}^{n,n}$ to derive the singular value decomposition of $A \otimes B$.
20.8 Show that for $A \in \mathbb{C}^{m,m}$ and $B \in \mathbb{C}^{n,n}$ and the matrix 2-norm, the equation $\|A \otimes B\|_2 = \|A\|_2\|B\|_2$ holds.
20.9 Prove Theorem 20.11.
20.10 Let $A \in \mathbb{C}^{m,m}$, $B \in \mathbb{C}^{n,n}$ and $C \in \mathbb{C}^{m,n}$. Show that $Z(t) = \exp(tA)C\exp(tB)$ is the solution of the matrix differential equation $\frac{dZ}{dt} = AZ + ZB$ with the initial condition $Z(0) = C$.