Notes of "Linear Mapping and Matrix Operations"

Jinxin Wang

1 Linear Mapping and Matrix

2 Matrix Multiplication

2.1 Block Matrix Multiplication

Let $X \in M_{m\times s}, Y \in M_{s\times n}$, then $Z = XY$ is valid and $Z \in M_{m\times n}$. Now we divide X and Y into blocks by vertical and horizontal lines:

$$X = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1k} \\ X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & & \vdots \\ X_{l1} & X_{l2} & \cdots & X_{lk} \end{pmatrix}$$

in which there are $l \times k$ blocks. $X_{i1}, X_{i2}, \ldots, X_{ik}$ are matrices with m_i rows ($m_1 + m_2 + \cdots + m_l = m$), and $X_{1j}, X_{2j}, \ldots, X_{lj}$ are matrices with s_j columns ($s_1 + s_2 + \cdots + s_k = s$).

$$Y = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1r} \\ Y_{21} & Y_{22} & \cdots & Y_{2r} \\ \vdots & \vdots & & \vdots \\ Y_{k1} & Y_{k2} & \cdots & Y_{kr} \end{pmatrix}$$

in which there are $k \times r$ blocks. $Y_{i1}, Y_{i2}, \ldots, Y_{ir}$ are matrices with s_i rows ($s_1 + s_2 + \cdots + s_k = s$), and $Y_{1j}, Y_{2j}, \ldots, Y_{kj}$ are matrices with n_j columns ($n_1 + n_2 + \cdots + n_r = n$).

By doing so, we can calculate the product $Z = XY$ by blocks. Block matrix multiplication follow the same rule as the regular matrix multiplication: the result $Z = XY$ has $l \times r$ blocks and the block Z_{ij} is

$$Z_{ij} = X_{i1}Y_{1j} + X_{i2}Y_{2j} + \cdots + X_{ik}Y_{kj}$$

in which the multiplication of blocks such as $X_{i\nu}Y_{\nu j}$ follows the rule of regular matrix multiplication. Since $X_{i\nu}$ is a $m_i \times s_\nu$ matrix and $Y_{\nu j}$ is a $s_\nu \times n_j$ matrix, the multiplication of blocks is valid and the result Z_{ij} is a $m_i \times n_j$ matrix. Thus $Z = XY$ has $\Sigma_{i=1}^{l} m_i = m$ rows and $\Sigma_{j=1}^{r} n_j = n$ columns, which is consistent with the result of regular matrix multiplication.

Next we check whether $Z = XY$ produced by block matrix multiplication has the same entry as the regular matrix multiplication in the corresponding position. Let $X = (x_{ij}), Y = (y_{ij}), Z = (z_{ij})$ and let us look at $z_{i_0 j_0}$.

By regular matrix multiplication,

$$z_{i_0j_0} = \Sigma_{t=1}^{s} x_{i_0t} y_{tj_0}$$

Suppose $z_{i_0j_0}$ is in the i_1-th row and the j_1-th column of block $Z_{i_2j_2}$, which is equivalent to

$$i_0 = m_1 + m_2 + \cdots + m_{i_2-1} + i_1$$

$$j_0 = n_1 + n_2 + \cdots + n_{j_2-1} + j_1$$

By block matrix multiplication, the block

$$Z_{i_2j_2} = X_{i_21}Y_{1j_2} + X_{i_22}Y_{2j_2} + \cdots + X_{i_2k}Y_{kj_2}$$

and $z_{i_0j_0}$ is the sum of the entries of $X_{i_2\nu}Y_{\nu j_2}(1 \leq \nu \leq k)$ in the position of the i_1-th row and the j_1-th, and each entry is the inner product of the i_1-th row vector of $X_{i_2\nu}$ and the j_1-th column vector of $Y_{\nu j_2}$. Since Z has the same row block division as X and the same column block division as Y, the i_1-th row vector of $X_{i_2\nu}$ is a part of the i_0-th row vector of X and the j_1-th column vector of $Y_{\nu j_2}$ is a part of the j_0-th column vector of Y. Hence,

$$z_{i_0j_0} = \Sigma_{t=1}^{s_1} x_{i_0t} y_{tj_0} + \Sigma_{t=s_1+1}^{s_1+s_2} x_{i_0t} y_{tj_0} + \cdots + \Sigma_{t=s_1+s_2+\cdots+s_{k-1}}^{s_1+s_2+\cdots+s_k} x_{i_0t} y_{tj_0} = \Sigma_{t=1}^{s} x_{i_0t} y_{tj_0} \tag{1}$$

Therefore, the block matrix multiplication and the regular matrix multiplication produce the same result.

Remark 1. *Notice that to make block matrix multiplication work, the division of columns of the left matrix in a matrix multiplication must be the same as the division of rows of the right matrix. As illustrated above, X has k columns of blocks with the j-th column of blocks containing s_j columns, and Y has k rows of blocks with the i-th row of blocks containing s_i rows.*

Remark 2. *From the equation (1) we can see, the essense of block matrix multiplication is breaking the calculation of the entries of a product of two matrices into pieces (blocks).*

2.1.1 Application of Block Matrix Multiplication

$$\begin{pmatrix} E & A \\ 0 & E \end{pmatrix} \begin{pmatrix} A & 0 \\ -E & B \end{pmatrix} = \begin{pmatrix} 0 & AB \\ -E & B \end{pmatrix} \tag{2}$$

$$\begin{pmatrix} E & 0 \\ -CA^{-1} & E \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} E & -A^{-1}B \\ 0 & E \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & D - CA^{-1}B \end{pmatrix} \tag{3}$$

3 Matrix Transposition

4 Rank of a Product of Matrices

5 Square Matrix

5.1 Commuting Matrices

Definition 1 (Diagonal Matrix and Scalar Matrix)**.** *A diagonal matrix is a a matrix in which all entries outside the main diagonal are all zero.*

Definition 2 (Identity Matrix)**.**

$$E = (\delta_{ij}), \delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases} \tag{4}$$

Definition 3 (Scalar Matrix)**.** *A scalar matrix is the result of the identity matrix multiplied with a scalar.*

Definition 4 (Matrix Unit)**.** *A matrix unit is a matrix with only one nonzero entry with value 1. The matrix unit with the nonzero entry in the i-th row and j-th column is denoted as E_{ij}*

Remark 3. *A matrix unit is not necessarily a square matrix.*

Definition 5 (Commuting Matrices)**.** *Given two matrices $A, B \in M_n(\mathbb{R})$, A and B are said to commute if $AB = BA$, or equivalently their commutator $[A, B] = AB - BA = 0$.*

Remark 4. *Notice that commuting matrices must be square matrix, because if $A \in M_{s\times n}, B \in M_{n\times r}, s \neq r$, then $AB \in M_{s\times r}, BA \in M_{r\times s}$, hence it is invalid to compare them.*

Theorem 1. *If a matrix $A \in M_n(\mathbb{R})$ commutes with $\forall B \in M_n(\mathbb{R})$, then A is a scalar matrix.*

证明. Hint:

- If $AE_{12} = E_{12}A$, then $\forall k \neq 1, a_{k1} = 0$, and $\forall k \neq 2, a_{2k} = 0$, and $a_{11} = a_{22}$.
- If $AE_{ij} = E_{ij}A$, then $\forall k \neq i, a_{ki} = 0$, and $\forall k \neq j, a_{jk} = 0$, and $a_{ii} = a_{jj}$.
- Consider $\forall B \in M_n\mathbb{R}, AB = BA$.

□

Remark 5. *When proving a property applies to any matrix in $M_{m\times n}$, one method is to consider all matrix units in $M_{m\times n}$, since the set of matrix units is a basis of $M_{m\times n}$.*

5.2 Inverse Matrix

Lemma 1 (Uniqueness of Inverse Matrix)**.**

$$A' = A'E = A'(AA'') = (A'A)A'' = A''$$

Definition 6 (Inverse Matrix)**.**

Definition 7 (Non-degenerate Matrix and Degenerate Matrix)**.** *A matrix $A \in M_n(\mathbb{R})$ is non-degenerate if* $\operatorname{rank} A = n$. *$A$ is degenerate if* $\operatorname{rank} A < n$.

Remark 6. *We only talk about non-degenerate matrices and degenerate matrices when it comes to square matrices.*

Theorem 2. *$A \in M_n(\mathbb{R})$ is non-degenerate if and only if A is inversible.*

证明. Hint:
$\Leftarrow$: Use the rank of the product of two matrices.
$\Rightarrow$: Use the unqiueness of the solution of $AX = 0$, and the transpose of the product of two matrices. □

Corollary 1. *If $A \in M_n\mathbb{R}$ is inversible, then A^T is inversible, and $(A^T)^{-1} = (A^{-1})^T$.*

Corollary 2. *If $B \in M_m(\mathbb{R})$ and $C \in M_n(\mathbb{R})$ are non-degenerate, $\forall A \in M_{m\times n}(\mathbb{R})$, it holds that*

$$\operatorname{rank} BAC = \operatorname{rank} A$$

Corollary 3. *If $A, B \in M_n(\mathbb{R})$, and $AB = E \vee BA = E$, then $B = A^{-1}$.*

Corollary 4. *If $A, B, \ldots, C, D \in M_n(\mathbb{R})$ are non-degenerate, then $AB \cdots CD$ is non-degenerate, and its inverse matrix is*

$$(AB \cdots CD)^{-1} = D^{-1}C^{-1} \cdots B^{-1}A^{-1}$$

5.3 Calculation of Powers of a Matrix

Example 1 (Powers of a Scalar Matrix)**.**

Example 2.

$$A = \begin{pmatrix} a & c \\ 0 & b \end{pmatrix}$$

Example 3.

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

6 Equivalence Class of Matrices

Definition 8 (Elementary Matrix)**.** *We call a matrix which can be obtained by doing one time of elementary row operation or column operation to the identity matrix as an elementary matrix.*

There are three kinds of elementary matrices, corresponding to the three kinds of elementary operations:

$$F_{s,t} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & 1 & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \tag{5}$$

Remark 7. *This matrix can be derived by exchanging the s-th column and the t-th column of E, or equivalently exchanging the s-th row and the t-th row of E.*

$$\begin{aligned} F_{s,t}(\lambda) &= E + \lambda E_{s,t} \\ &= \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & \cdots & \lambda & \cdots \\ & & & \ddots & & \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix} \end{aligned} \tag{6}$$

Remark 8. *This matrix can be derived by adding the s-th column multiplied by λ to the t-th column, or equivalently adding the t-th row multiplied by λ to the s-th row.*

$$\begin{aligned} F_{s}(\lambda) &= E + (\lambda - 1) E_{s,s} \\ &= \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \lambda & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \end{aligned} \tag{7}$$

Remark 9. *This matrix can be derived by multiplying the s-th row by λ, or equivalently multiplying the s-th column by λ.*

Proposition 1. *Given an elementary matrix F, FA is the result of applying the corresponding elementary row operation to A, and AF is the result of applying the corresponding elementary column operation to A.*

证明. We can easily prove this proposition by carrying out the matrix multiplication with block matrix multiplication. To check the result of FA, divide A into row blocks (i.e. a block is a row of A). To check the result of AF, divide A into column blocks (i.e. a block is a column of A). □

Remark 10. *The elementary row operation to turn A into FA is the same as the elementary row operation to turn E into F. For instance, $F_{s,t}$ is obtained by exchanging the s-th row and the t-th row of E, then $F_{s,t}A$ is the result of exchanging the s-th row and the t-th row of A.*

Similarly, the elementary column operation to turn A into AF is the same as the elementary column operation to turn E into F.

Proposition 2. *For any matrix $A \in M_{s\times t}$, it can be transformed into the following form*

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} \tag{8}$$

with the three kinds of elementary row operations and column operations, and $\operatorname{rank} A = \operatorname{rank} E_r$.

Definition 9. *If $A, B \in M_{s\times t}$, we say they are equivalent, denoted by $A\ B$, if there are non-degenerate matrices P and Q such that $B = PAQ$.*

Remark 11. *Note that the prerequisite to discuss whether two matrices are equivalent or not is that they have the same dimension (size).*

Remark 12. *It is clear that the equivalence between two matrices defined above is a kind of equivalence relation, since it satisfies the three properties:*

- *reflexivity*
- *symmetry*
- *transivity*

From the above proposition, we can conclude that a matrix with its rank as r is equivalent to a matrix in a standard form

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

with the same rank. Therefore, with the symmetry and transivity of the equivalence relation between two matrices, all matrices with the same rank are equivalent. Any pair of matrices with different rank are not equivalent. The equivalence relation leads to the following theorem:

Theorem 3. *$M_{m\times n}(\mathbb{R})$ has a partition of $\min(m, n) + 1$ equivalence classes. All matrices with its rank as r including the representative element is in a equivalence class.*

Corollary 5. *Every non-degenerate matrix $A \in M_n(\mathbb{R})$ can be expressed as the product of elementary matrices.*

证明. Based on the above proposition, we can find a series of elementary matrices $P_1, P_2, \ldots, P_s$ and $Q_1, Q_2, \ldots, Q_t$ such that

$$\begin{aligned} E &= P_s \cdots P_2 P_1 A Q_1 Q_2 \cdots Q_t \\ A &= P_1^{-1} P_2^{-1} \cdots P_s^{-1} E Q_t^{-1} \cdots Q_2^{-1} Q_1^{-1} \\ &= P_1^{-1} P_2^{-1} \cdots P_s^{-1} Q_t^{-1} \cdots Q_2^{-1} Q_1^{-1} \end{aligned}$$

□

7 Calculation of Inverse Matrix

8 Elementary Block Matrix

9 Solution Space