Square matrices and linear transformations

Definition 1. (Linear transformation) *A linear function* $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^N$ *is called a "linear transformation".*

Clearly, the matrix representations of linear transformations have the same number of rows and columns. Such matrices are called "square matrices".

In this section we restrict ourselves to square matrices. That is $A \in \mathbb{R}^{N \times N}$.

Square matrices

Special square matrices

- Identity matrix.

  The most special square matrix is the "identity matrix".

  $$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = (\delta_{ij}) \tag{1}$$

  Here $\delta_{ij} = 1$ when $i = j$ and 0 otherwise. Sometimes the notation I_n is used to make explicit the size of the identity matrix.

  Lemma 2. *For all* $A \in \mathbb{R}^{N \times N}$,

  $$I\,A = A; \qquad A\,I = A \tag{2}$$

  Proof. Direct calculation. □

- Permutation matrix.

  A permutation matrix P is a matrix whose entries are only 0 and 1, and along each row or column there is exactly one 1.

  Any permutation matrix can be obtained from the identity matrix I through switching rows or columns finitely many times.

  Exercise 1. Check that $P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ is a permutation matrix. Take a vector $\boldsymbol{x} \in \mathbb{R}^3$ and calculate $P\boldsymbol{x}$, then take any 3×3 matrix A and calculate $P\,A$ and $A\,P$. What do you observe?

  Exercise 2. Let P, Q be permutation matrices. Prove that PQ is still a permutation matrix.

- Diagonal matrix.

  A "diagonal matrix" is of the form

  $$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & d_N \end{pmatrix} \tag{3}$$

  Diagonal matrices are often denoted as $D = \operatorname{diag}(d_1, \ldots, d_N)$.

Exercise 3. Let $D_1, D_2 \in \mathbb{R}^{N \times N}$ be diagonal matrices. Prove that $D_1 \pm D_2, D_1 D_2$ are still diagonal matrices.

- Upper/lower triangular matrix.

  Two still less special classes of matrices are upper triangulr and lower triangular matrices.

$$U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1N} \\ 0 & u_{22} & & u_{2N} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & u_{NN} \end{pmatrix}, \qquad L = \begin{pmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ l_{N1} & l_{N2} & \cdots & l_{NN} \end{pmatrix} \tag{4}$$

  that is $u_{ij} = 0$ whenever $i > j$ and $l_{ij} = 0$ whenever $i < j$.

Exercise 4. Prove: The transpose of an upper triangular matrix is lower triangular and vice versa.

Exercise 5. Let $U, V \in \mathbb{R}^{n \times n}$ be upper triangular. Prove that UV is still upper triangular. Find out the relation between the u_{ii}, v_{ii} and the i-i diagonal entry of UV.

Is there a similar claim for lower triangular matrices? Can you prove it?

Exercise 6. Let $U, L \in \mathbb{R}^{N \times N}$ be upper and lower triangular respectively. Show through an example that UL and LU may not be triangular.

- Orthogonal matrix.

  A matrix $O \in \mathbb{R}^{N \times N}$ is said to be orthogonal if and only if $O^T O = I$ where O^T is the transpose of O.

Exercise 7. Write

$$O = (\ \boldsymbol{a}_1 \ \cdots \ \boldsymbol{a}_N\ ) \tag{5}$$

as a row of columns. Prove that O is orthogonal if and only if

1. $\|\boldsymbol{a}_i\| = 1$ for $i = 1, 2, \ldots, N$.
2. $\boldsymbol{a}_i \perp \boldsymbol{a}_j$ for all $i \neq j$.

Exercise 8. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$. Let $O \in \mathbb{R}^{N \times N}$ be orthogonal. Then

$$(O\boldsymbol{x}) \cdot (O\boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y}. \tag{6}$$

In particular, $\|O\boldsymbol{x}\| = \|\boldsymbol{x}\|$.

Exercise 9. Let $O \in \mathbb{R}^{N \times N}$ be such that $\|O\boldsymbol{x}\| = \|\boldsymbol{x}\|$ for any $\boldsymbol{x} \in \mathbb{R}^N$. Prove that O is orthogonal.

Theorem 3. (LDU decomposition) *Let $A \in \mathbb{R}^{N \times N}$. Then*

$$A = P L D U \tag{7}$$

where P is a permutation matrix, L is a lower triangular matrix, D is a diagonal matrix, U is an upper triangular matrix. Furthermore all the main diagonal entries of L, U are 1. That is $l_{ii} = u_{ii} = 1$ for $i = 1, 2, \ldots, N$.

Proof. See any numerical linear algebra or matrix theory textbook. If you want to prove it yourself, review the process of Gaussian elimination and try to write each step as a matrix multiplication. $\square$

Theorem 4. (Polar decomposition) *Let $A \in \mathbb{R}^{N \times N}$. Then there are orthogonal matrices O_1, O_2 and diagonal matrix D with non-negative diagonal entries, such that $A = O_1 D O_2$.*

Inverse

Definition 5. **(Inverse matrix)** *Let $A \in \mathbb{R}^{N \times N}$. If there is a matrix $B \in \mathbb{R}^{N \times N}$ such that $A\,B = I$ and $B\,A = I$, then B is called the "inverse" of A and denoted A^{-1}.*

Those matrices that have inverses are called "invertible" or "non-singular". The rest are called "singular".

Exercise 10. Let $A \in \mathbb{R}^{1 \times 1}$. When is A invertible? Find its inverse.

Exercise 11. Prove that the identity matrix I is invertible. Find its inverse.

Exercise 12. Let D be diagonal. When is it invertible? Find its inverse.

Exercise 13. Find sufficient and necessary condition for an upper triangular matrix to be invertible. Prove that the inverse of an upper triangular matrix is lower triangular.

Exercise 14. Let $O \in \mathbb{R}^{N \times N}$ be an orthongal matrix. Prove that its inverse O^{-1} is also orthogonal.

Lemma 6. *Let $A, B \in \mathbb{R}^{N \times N}$ be invertible. Then $A\,B$ is also invertible and $(A\,B)^{-1} = B^{-1}\,A^{-1}$.*

Proof. Check

$$(A\,B)\,(B^{-1}\,A^{-1}) = A\,(B\,B^{-1})\,A^{-1} = A\,I\,A^{-1} = A\,A^{-1} = I; \tag{8}$$

$$(B^{-1}\,A^{-1})\,(A\,B) = B^{-1}\,(A^{-1}\,A)\,B = B^{-1}\,I\,B = B^{-1}\,B = I. \tag{9}$$

Therefore the matrix $B^{-1}\,A^{-1}$ is the inverse of the matrix $A\,B$. □

Exercise 15. Let $A \in \mathbb{R}^{N \times N}$. Prove that A is invertible if and only if A^T is invertible. Furthermore,

$$(A^T)^{-1} = (A^{-1})^T. \tag{10}$$

Notation 7. *In the following we will use the simpler notation A^{-T} to denote the above matrix $(A^T)^{-1} = (A^{-1})^T$.*

Exercise 16. Let $A \in \mathbb{R}^{N \times N}$. Assume there is $B \in \mathbb{R}^{N \times N}$ such that $A\,B = I$. Prove that $B\,A = I$.

Theorem 8. *A matrix $A \in \mathbb{R}^{N \times N}$ is singular if and only if there is a column vector $\boldsymbol{x}$ (or a row vector $\boldsymbol{y}^T$) such that*

$$A\,\boldsymbol{x} = 0\;(or\;\boldsymbol{y}^T A = 0) \tag{11}$$

Determinant

Theorem 9. *For each $n \in \mathbb{N}$, there is a unique function $\theta \colon \mathbb{R}^{N \times N} \mapsto \mathbb{R}$ such that*

a) θ is linear in each column of A;

b) $\theta(\tilde{A}) = -\theta(A)$ if $\tilde{A}$ is obtained from A by interchanging two columns.

c) $\theta(I) = 1$.

Proof. Omitted. But see the following for the main idea of the proof. □

Exercise 17. Let $A=(a)\in\mathbb{R}^{1\times 1}$. Prove that $\theta(A)=a$.

Example 10. Let $A=\left(\begin{smallmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{smallmatrix}\right)$. Then $\theta(A)=a_{11}\,a_{22}-a_{12}\,a_{21}$..

Proof. By linearity of θ with respect to each column, we have

$$\begin{aligned}\theta\left(\left(\begin{array}{cc} a_{11} & a_{12}\\ a_{21} & a_{22}\end{array}\right)\right) &= \theta\left(\left(\begin{array}{cc} a_{11} & a_{12}\\ 0 & a_{22}\end{array}\right)\right)+\theta\left(\left(\begin{array}{cc} 0 & a_{12}\\ a_{21} & a_{22}\end{array}\right)\right)\\ &= a_{11}\,\theta\left(\left(\begin{array}{cc} 1 & a_{12}\\ 0 & a_{22}\end{array}\right)\right)+a_{21}\,\theta\left(\left(\begin{array}{cc} 0 & a_{12}\\ 1 & a_{22}\end{array}\right)\right)\\ &= a_{11}a_{12}\,\theta\left(\left(\begin{array}{cc} 1 & 1\\ 0 & 0\end{array}\right)\right)+a_{11}\,a_{22}\,\theta\left(\left(\begin{array}{cc} 1 & 0\\ 0 & 1\end{array}\right)\right)\\ &\quad +a_{21}\,a_{12}\,\theta\left(\left(\begin{array}{cc} 0 & 1\\ 1 & 0\end{array}\right)\right)+a_{21}\,a_{22}\,\theta\left(\left(\begin{array}{cc} 0 & 0\\ 1 & 1\end{array}\right)\right). \end{aligned} \tag{12}$$

By property b) we see that the first and fourth terms are 0, and

$$\theta\left(\left(\begin{array}{cc} 0 & 1\\ 1 & 0\end{array}\right)\right)=-\theta\left(\left(\begin{array}{cc} 1 & 0\\ 0 & 1\end{array}\right)\right). \tag{13}$$

Finally by property c) we have $\theta(A)=a_{11}\,a_{22}-a_{12}\,a_{21}$ as desired. □

Exercise 18. Let $A=\left(\begin{smallmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{smallmatrix}\right)$. Then $\theta(A)=a_{11}\,a_{22}\,a_{33}+a_{12}\,a_{23}\,a_{31}+a_{13}\,a_{32}\,a_{21}-a_{13}\,a_{22}\,a_{31}-a_{12}\,a_{21}\,a_{33}-a_{23}\,a_{32}\,a_{11}$.

Definition 11. *The unique function in Theorem 9 is called the* determinant*. The traditional notation is* $\det(A)$.

Exercise 19. Let $A\in\mathbb{R}^{N\times N}$. Denote by $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_N$ the n columns. Then if there are i,j such that $\boldsymbol{a}_i=\boldsymbol{a}_j$, then $\det(A)=0$.

Exercise 20. Let $A\in\mathbb{R}^{N\times N}$. Denote by $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_N$ the N columns. If there are $c_1,\ldots,c_N\in\mathbb{R}$ such that

$$c_1\,\boldsymbol{a}_1+\cdots+c_N\,\boldsymbol{a}_N=\mathbf{0}\in\mathbb{R}^{N\times N}, \tag{14}$$

then $\det(A)=0$.

Exercise 21. Use the above exercise to prove that $\det A=\det\tilde{A}$ where $\tilde{A}$ is obtained from A by replace column j with the sum of column j and column i, but keep other columns intact.

Exercise 22. Prove that if $D=\mathrm{diag}(d_1,\ldots,d_N)$ is diagonal, then $\det(D)=d_1\,d_2\cdots d_N$.

Exercise 23. Prove that if P is a permutation matrix, then $|\det(P)|=1$.

Exercise 24. Prove that for upper triangular matrices, $\det(U)=u_{11}\cdots u_{NN}$.

Theorem 12. *Let* $A,B\in\mathbb{R}^{N\times N}$*. Prove that* $\det(A\,B)=(\det(A))\,(\det(B))$.

Proof. Consider the function $\theta\colon\mathbb{R}^{N\times N}\mapsto\mathbb{R}$ by

$$\theta(X):=\det(A\,X). \tag{15}$$

Then θ satisfies a),b) in Theorem 9 and with c) replaced by $\theta(I) = r = \det(A)$. The same method proving Theorem 9 gives $\theta(X) = r\det(X) = (\det(A))(\det(X))$ which ends the proof. $\square$

Remark 13. The above proof is from *Introduction to Differential Equations* by Michael E. Taylor, AMS 2011.

Exercise 25. Let $A \in \mathbb{R}^{N \times N}$. Prove $\det(A^T) = \det(A)$.

Exercise 26. Use the above exercise to prove $\det(O) = 1$ for any orthogonal matrix O.

Theorem 14. *Let $A \in \mathbb{R}^{N \times N}$. Then A is invertible if and only if $\det(A) \neq 0$.*

Theorem 15. *Let $A \in \mathbb{R}^{N \times N}$ and let $P L D U$ be its LDU decomposition. Then*

$$|\det(A)| = |\det(D)|. \tag{16}$$

Linear transformations

Lemma 16. *The linear transformation with matrix representation I is the identity transformation $\boldsymbol{i}(\boldsymbol{x}) = \boldsymbol{x}$.*

Lemma 17. *Let $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^N$ be a linear transformation that is invertible. Then its inverse $\boldsymbol{g}$ is also a linear transformation.*

Exercise 27. Prove the above lemma.

Theorem 18. *Let $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^N$ be a linear transformation. Then $\boldsymbol{f}$ is invertible if and only if its matrix representation A is invertible. Furthermore the inverse transformation $\boldsymbol{g}$ has representation A^{-1}.*

Proof.

- "If". If A is invertible, then we define the linear transformation $\boldsymbol{g}$ through

  $$\boldsymbol{g}(\boldsymbol{x}) := A^{-1} \boldsymbol{x}. \tag{17}$$

  Now it is easy to check that

  $$(\boldsymbol{f} \circ \boldsymbol{g})(\boldsymbol{x}) = \boldsymbol{x}, \qquad (\boldsymbol{g} \circ f)(\boldsymbol{y}) = \boldsymbol{y} \tag{18}$$

  for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$.

- "Only if". Assume $\boldsymbol{f}$ is invertible. Let $\boldsymbol{g}$ be its inverse. Since $\boldsymbol{g}$ is a linear transformation, it has a matrix representation B. Thus we have

  $$A B = B A = I \tag{19}$$

  which means A is invertible and its inverse is B. $\square$