Linear Algebra

In this section, we present some linear algebra results that are useful in the context of this course. For more information, you can refer to the MAT-1200 course, Deisenroth, Faisal, and Ong (2020) (in English), and Grifone (2024) (in French).

1 Some matrix properties

Let $M_{n,m}(\mathbb{R})$ be the set of matrices with n rows and m columns whose entries belong to $\mathbb{R}$. Let $M_n(\mathbb{R})$ be the set of square matrices of size n, i.e., with n rows and n columns whose entries belong to $\mathbb{R}$. Let M, N, and P be matrices in $M_{n,m}(\mathbb{R})$. Let A and B be matrices in $M_n(\mathbb{R})$. Let I_n be the identity matrix of size n, i.e., containing 1s on the diagonal and 0s on the elements outside the diagonal. Let u and v in $\mathbb{R}^n$, i.e., column vectors of size n.

Properties of the inverse of matrices

Suppose that the matrices A and B are invertible. Then the matrix product AB is invertible and is given by:

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Proof

Let $C = AB$ and $D = B^{-1}A^{-1}$. Then

$$\begin{aligned} CD &= ABB^{-1}A^{-1} \\ &= AA^{-1} \\ &= I_n \end{aligned}$$

Similarly, we find that $DC = I_n$. Thus, AB is invertible and its inverse is given by $B^{-1}A^{-1}$.

Properties of the determinant of matrices

Considering the matrices defined at the beginning of the section, we have:

1. $\det(A^\top) = \det(A)$,
2. $\det(AB) = \det(A)\det(B)$,
3. $\det(A^{-1}) = 1/\det(A)$.

Proof

The proofs of properties 1 and 2 are technical and are omitted, but can be found, for example, here. As for the third property, by definition, we have $AA^{-1} = I_n$. The determinant of I_n is equal to 1 (product of the elements on the diagonal). Therefore, $\det(AA^{-1}) = 1$. However, according to the second property, $\det(AA^{-1}) = \det(A)\det(A^{-1})$. We therefore have $\det(A^{-1}) = 1/\det(A)$.

Properties of the trace of matrices

Considering the matrices defined at the beginning of the section, we have:

1. $\mathrm{tr}(A) = \mathrm{tr}(A^\top)$,
2. $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$,
3. $\mathrm{tr}(MN^\top) = \mathrm{tr}(N^\top M)$.

Proof

For a square matrix A, let a_{ij} be the element of matrix A in row i and column j. The trace of A is given by the sum of the diagonal elements, i.e. $\mathrm{tr}(A) = \sum_{i=1}^n a_{ii}$.

1. Since transposition does not change the diagonal elements, the result is straightforward.
2. Let $C = A + B$. Since A and B are square matrices, C is a square matrix. We have $c_{ij} = a_{ij} + b_{ij}$ for all $i, j = 1, \dots, n$. Therefore

$$\mathrm{tr}(A+B) = \mathrm{tr}(C) = \sum_{i=1}^n c_{ii} = \sum_{i=1}^n a_{ii} + b_{ii} = \sum_{i=1}^n a_{ii} + \sum_{i=1}^n b_{ii} = \mathrm{tr}(A) + \mathrm{tr}(B).$$

3. The matrices $MN^\top$ and $N^\top M$ are square, with dimensions $n \times n$ and $m \times m$

respectively, so we can calculate their traces. Let $C = MN^\top$ and $D = N^\top M$.

$$\operatorname{tr}(MN^\top) = \operatorname{tr}(C) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{m} m_{ij} n_{ji} = \sum_{j=1}^{m}\sum_{i=1}^{n} n_{ji} m_{ij} = \sum_{j=1}^{m} d_{jj} = \operatorname{tr}(D) = \operatorname{tr}(N^\top M).$$

Definition

1. Let A be a symmetric matrix in $M_n(\mathbb{R})$. A is **positive definite** if $u^\top A u > 0$ for all $u \in \mathbb{R}^n$ such that $u \neq 0$.
2. Let $A \in M_n(\mathbb{R})$. A is **orthogonal** if $A^\top A = AA^\top = I_n$.

2 Eigenvalues and eigenvectors

Definition

Let $A \in M_n(\mathbb{R})$. We say that $\lambda \in \mathbb{R}$ is an **eigenvalue** of A if there exists a nonzero vector $u \in \mathbb{R}^n$ such that

$$Au = \lambda u. \tag{1}$$

The vector u is called an **eigenvector** of A corresponding to the eigenvalue λ.
The set of real numbers λ satisfying Equation 1 is called the **spectrum** of the matrix A and is denoted by $\operatorname{sp}(A)$.

Property of eigenvectors

1. If u is an eigenvector of A corresponding to an eigenvalue λ, then the vector cu, $c \in \mathbb{R}^\star$ is also an eigenvector of A corresponding to λ.
2. If A is symmetric and u_1 and u_2 are eigenvectors corresponding to different eigenvalues of A, then u_1 and u_2 are orthogonal, i.e. $u_1^\top u_2 = 0$.

Proof

1. Let $c \in \mathbb{R}^\star$ and u be an eigenvector of A associated with the eigenvalue λ. We have:

   $$A(cu) = cAu = c\lambda u = \lambda(cu).$$

   Therefore, the vector cu is also an eigenvector of A associated with the eigenvalue λ.
2. Let λ_1 and λ_2 be the eigenvalues associated with u_1 and u_2, such that $\lambda_1 \neq \lambda_2$. We

have $Au_1 = \lambda_1 u_1$ and $Au_2 = \lambda_2 u_2$. Then

$$\lambda_1 u_1^\top u_2 = u_1^\top A u_2 = \lambda_2 u_1^\top u_2.$$

This implies that $(\lambda_1 - \lambda_2) u_1^\top u_2 = 0$. However, $\lambda_1 \neq \lambda_2$. Therefore, necessarily, $u_1^\top u_2 = 0$.

This second property will be useful when we look at dimension reduction and, in particular, principal component analysis.

Characterization of matrices with their eigenvalues

1. If A is symmetric, then **all** its eigenvalues are real.
2. If A is positive definite, then **all** its eigenvalues are strictly positive.

Proof

1. Consider the more general case where A is a Hermitian matrix. The matrix A is equal to the transpose of its conjugate, denoted A^*. Let λ be an eigenvalue associated with an eigenvector u, which may be complex. We have:

$$\overline{u}^\top A u = \overline{u}^\top \lambda u = \lambda \overline{u}^\top u, \tag{2}$$

$$\overline{u}^\top A u = \overline{u}^\top A^* u = \overline{Au}^\top u = \overline{\lambda} \overline{u}^\top u. \tag{3}$$

This implies that $(\lambda - \overline{\lambda})\overline{u}^\top u = 0$. Since $u \neq 0$, we have $\lambda = \overline{\lambda}$. Therefore, $\lambda \in \mathbb{R}$.

2. Consider u, an eigenvector of A associated with the eigenvalue λ. We have that $u^\top A u = \lambda u^\top u$. However, since $u \neq 0$, $u^\top u \neq 0$. Therefore,

$$\lambda = \frac{u^\top A u}{u^\top u}.$$

Since A is positive definite, $u^\top A u > 0$ for all nonzero vectors u. We can therefore deduce that $\lambda > 0$.

3 Diagonalization of matrices

Definition

Let $A \in M_n(\mathbb{R})$. We say that A is **diagonalizable** if there exists a non-singular matrix $P \in M_n(\mathbb{R})$ and a diagonal matrix $D \in M_n(\mathbb{R})$ such that

$$P^{-1}AP = D \iff A = PDP^{-1}.$$

Spectral decomposition theorem

Let A be a symmetric matrix in $M_n(\mathbb{R})$ and $\lambda_1, \dots, \lambda_n$ its n eigenvalues. Then there exists an orthogonal matrix P in $M_n(\mathbb{R})$ such that

$$A = P\Lambda P^\top, \quad \text{where} \quad \Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n).$$

If A has n distinct positive eigenvalues, then we can take P to be the matrix whose kth column is the normalized eigenvector corresponding to the kth eigenvalue λ_k of A.

Given two symmetric matrices, A and B, how can we determine the vector u such that $u^\top Au$ is maximal, knowing that $u^\top Bu = 1$? We simply take u as the eigenvector of $B^{-1}A$ associated with λ, the maximal eigenvalue of $B^{-1}A$. We thus obtain

$$u^\top Au = u^\top \lambda Mu = \lambda U^\top Mu = \lambda.$$

Characterization of the determinant and trace of matrices with their eigenvalues

If A has eigenvalues (real, but not necessarily distinct) $\lambda_1, \dots, \lambda_n$, then

1. $\det(A) = \prod_{i=1}^n \lambda_i$
2. $\mathrm{tr}(A) = \sum_{i=1}^n \lambda_i$.

Proof

Using the spectral decomposition theorem, there exists an invertible matrix P such that $A = P\Lambda P^{-1}$, where Λ is a diagonal matrix containing the eigenvalues. We therefore have, for the determinant,

$$\det(A) = \det(P\Lambda P^{-1}) = \det(P)\det(\Lambda)\det(P^{-1}) = \det(P)\det(\Lambda)\det(P)^{-1} = \det(\lambda) = \prod_{i=1}^n \lambda_i,$$

and, for the trace,

$$\operatorname{tr}(A) = \operatorname{tr}(P\Lambda P^{-1}) = \operatorname{tr}(P^{-1}P\Lambda) = \operatorname{tr}(\Lambda) = \sum_{i=1}^{n} \lambda_i.$$

References

Deisenroth, Marc Peter, A. Aldo Faisal, and Cheng Soon Ong. 2020. *Mathematics for Machine Learning*. 1st ed. Cambridge University Press. https://doi.org/10.1017/9781108679930.

Grifone, Joseph. 2024. *Algèbre Linéaire*. 7e edition. Toulouse: CEPADUES.