Advanced Econometrics

Lecture -1: Matrix Algebra (Hansen Appendix A)

Instructor: Ma, Jun

Renmin University of China

Fall 2018

Notation

标量 a.

向量 $\boldsymbol{a}$.

- A **scalar** a is a single number.
- A **vector** $\boldsymbol{a}$ is $k \times 1$ list of numbers, typically arranged in a column. We write this as

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{pmatrix}.$$

列向量.

最好把向量都理解为列向量.

- Equivalently, a vector $\boldsymbol{a}$ is an element of $\mathbb{R}^k$.

$\vec{a} \in \mathbb{R}^k$

- A **matrix** $\boldsymbol{A}$ is a $k \times r$ rectangular array of numbers, written as

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kr} \end{bmatrix}.$$

A是 $k \times r$ 的一个矩阵

a_{ij} 是一般的一个元素

$A = (a_{ij})$

By convention, a_{ij} refers to the element in the i^{th} row and j^{th} column of $\boldsymbol{A}$. Sometimes a matrix $\boldsymbol{A}$ is denoted by the symbol (a_{ij}).

Notation

- A matrix can be written as a set of column vectors or as a set of row vectors. That is,

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_r \end{bmatrix} = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \vdots \\ \boldsymbol{\alpha}_k \end{bmatrix}$$

  where

$$\boldsymbol{a}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ki} \end{bmatrix}$$

are column vectors and

$$\boldsymbol{\alpha}_j = \begin{bmatrix} a_{j1} & a_{j2} & \cdots & a_{jr} \end{bmatrix}$$

are row vectors.

Notation

A转秩 → A'

当A=A'时，方块
阵是对称的.

- The **transpose** of a matrix $\boldsymbol{A}$, denoted as $\boldsymbol{A}'$, is obtained by flipping the matrix on its diagonal:

$$\boldsymbol{A}' = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \vdots & \vdots & & \vdots \\ a_{1r} & a_{2r} & \cdots & a_{kr} \end{bmatrix}.$$

- If $\boldsymbol{A}$ is $k \times r$, then $\boldsymbol{A}'$ is $r \times k$. If $\boldsymbol{a}$ is a $k \times 1$ vector, then $\boldsymbol{a}'$ is an $1 \times k$ row vector.
- A matrix is **square** if $k = r$. A matrix is **symmetric** if $\boldsymbol{A} = \boldsymbol{A}'$. A square matrix is **diagonal** if the off-diagonal elements are all zero, so that $a_{ij} = a_{ji}$. A square matrix is **upper (lower) diagonal** if all elements below (above) the diagonal equal zero.

Notation

- An important diagonal matrix is the **identity** matrix, which has ones on the diagonal. The $k \times k$ identity matrix is denoted as

$$I_k = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

单位阵.

$IA = A$

- A **partitioned matrix** takes the form

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ A_{21} & A_{22} & \cdots & A_{2r} \\ \vdots & \vdots & & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{kr} \end{bmatrix}.$$

Matrix Addition

矩阵相加：如果两个矩阵维度一样，即每个元素相加。满足交换律，结合律

- If the matrices $\boldsymbol{A} = \left(a_{ij}\right)$ and $\boldsymbol{B} = \left(b_{ij}\right)$ are of the same order, we define the sum

$$\boldsymbol{A} + \boldsymbol{B} = \left(a_{ij} + b_{ij}\right).$$

- Matrix addition follows the commutative and associative laws:

$$\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}; \; \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C}.$$

Matrix Multiplication

- If $\boldsymbol{A}$ is $k \times r$ and c is scalar, we define the product as

$$\boldsymbol{A}c = c\boldsymbol{A} = \left(a_{ij}c\right).$$

- If $\boldsymbol{a}$ and $\boldsymbol{b}$ are both $k \times 1$, then their **inner product** is

内积：所有元素积的和.

$$\boldsymbol{a}'\boldsymbol{b} = a_1b_1 + a_2b_2 + \cdots a_kb_k = \sum_{j=1}^{k} a_jb_j.$$

$\vec{a}'\cdot\vec{b}=0 \Rightarrow$ 正交.

- We say that the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are **orthogonal** if $\boldsymbol{a}'\boldsymbol{b} = 0$.

Matrix Multiplication

- If $\boldsymbol{A}$ is $k \times r$ and $\boldsymbol{B}$ is $r \times s$ so that the number of columns of $\boldsymbol{A}$ equals the number of rows of $\boldsymbol{B}$ we say that $\boldsymbol{A}$ and $\boldsymbol{B}$ are conformable. In this event, the **matrix product** $\boldsymbol{AB}$ is defined.
- Writing $\boldsymbol{A}$ as a set of row vectors and $\boldsymbol{B}$ as a set of column vectors (each of length r), then the matrix product is defined as

A的列数=B的行数

才可乘. $\Rightarrow A_{k\times r}\ B_{r\times s}$

$$
\boldsymbol{AB} = \begin{bmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \\ \vdots \\ \boldsymbol{a}_k' \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_s \end{bmatrix}
$$

$$
= \begin{bmatrix} \boldsymbol{a}_1'\boldsymbol{b}_1 & \boldsymbol{a}_1'\boldsymbol{b}_2 & \cdots & \boldsymbol{a}_1'\boldsymbol{b}_s \\ \boldsymbol{a}_2'\boldsymbol{b}_1 & \boldsymbol{a}_2'\boldsymbol{b}_2 & \cdots & \boldsymbol{a}_2'\boldsymbol{b}_s \\ \vdots & \vdots & & \vdots \\ \boldsymbol{a}_k'\boldsymbol{b}_1 & \boldsymbol{a}_k'\boldsymbol{b}_2 & \cdots & \boldsymbol{a}_k'\boldsymbol{b}_s \end{bmatrix}.
$$

i行j列的元素=

A的第i行和B的第j列元素

的内积.

Matrix Multiplication

- Matrix multiplication is not commutative: in general $\boldsymbol{AB} \neq \boldsymbol{BA}$. However, it is associative and distributive:

$$\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C}; \; \boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}.$$

- An alternative way to write the matrix product is to use matrix partitions:

$A_{k\times 1} \cdot B_{1\times s}$

$$\begin{aligned} \boldsymbol{AB} &= \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_r \end{bmatrix} \begin{bmatrix} \boldsymbol{B}_1 \\ \boldsymbol{B}_2 \\ \vdots \\ \boldsymbol{B}_r \end{bmatrix} \\ &= \boldsymbol{A}_1\boldsymbol{B}_1 + \boldsymbol{A}_2\boldsymbol{B}_2 + \cdots + \boldsymbol{A}_r\boldsymbol{B}_r \\ &= \sum_{j=1}^{r} \boldsymbol{A}_j\boldsymbol{B}_j. \end{aligned}$$

Matrix Multiplication

- An important property of the identity matrix is that if $\boldsymbol{A}$ is $k \times r$, then $\boldsymbol{A}\boldsymbol{I}_r = \boldsymbol{A}$ and $\boldsymbol{I}_k\boldsymbol{A} = \boldsymbol{A}$.
- We say two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are **orthogonal** if $\boldsymbol{A}'\boldsymbol{B} = 0$. This means that all columns of $\boldsymbol{A}$ are orthogonal with all columns of $\boldsymbol{B}$.
- The $k \times r$ matrix $\boldsymbol{H}$, $r \leq k$, is called **orthonormal** if $\boldsymbol{H}'\boldsymbol{H} = \boldsymbol{I}_r$. This means that the columns of $\boldsymbol{H}$ are mutually orthogonal, and each column is normalized to have unit length.

A的转置乘B =0

即A的所有列与B的所有列正交.

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

两矩阵正交.

Trace

矩阵的迹.

- The **trace** of a $k \times k$ square matrix $\boldsymbol{A}$ is the sum of its diagonal elements

$$\operatorname{tr}(\boldsymbol{A}) = \sum_{i=1}^{k} a_{ii}.$$

$k \times k$ 矩阵对角线元素之和

称为矩阵的迹.

- Some straightforward properties for square matrices$\boldsymbol{A}$ and $\boldsymbol{B}$ and scalar c are

$$\operatorname{tr}(c\boldsymbol{A}) = c\operatorname{tr}(\boldsymbol{A});\ \operatorname{tr}(\boldsymbol{A}') = \operatorname{tr}(\boldsymbol{A});\ \operatorname{tr}(\boldsymbol{A}+\boldsymbol{B}) = \operatorname{tr}(\boldsymbol{A})+\operatorname{tr}(\boldsymbol{B});\ \operatorname{tr}(\boldsymbol{I}_k) = k.$$

Trace

- For $k \times r$ $\boldsymbol{A}$ and $r \times k$ $\boldsymbol{B}$ we have
$$\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}(\boldsymbol{BA})$$
since
$$\begin{aligned}\operatorname{tr}(\boldsymbol{AB}) &= \operatorname{tr}\begin{bmatrix} \boldsymbol{a}_1'\boldsymbol{b}_1 & \boldsymbol{a}_1'\boldsymbol{b}_2 & \cdots & \boldsymbol{a}_1'\boldsymbol{b}_k \\ \boldsymbol{a}_2'\boldsymbol{b}_1 & \boldsymbol{a}_2'\boldsymbol{b}_2 & \cdots & \boldsymbol{a}_2'\boldsymbol{b}_k \\ \vdots & \vdots & & \vdots \\ \boldsymbol{a}_k'\boldsymbol{b}_1 & \boldsymbol{a}_k'\boldsymbol{b}_2 & \cdots & \boldsymbol{a}_k'\boldsymbol{b}_k \end{bmatrix} \\ &= \sum_{i=1}^{k} \boldsymbol{a}_i'\boldsymbol{b}_i \\ &= \sum_{i=1}^{k}\sum_{j=1}^{r} a_{ij}b_{ji} \\ &= \sum_{j=1}^{r}\sum_{i=1}^{k} b_{ji}a_{ij} \\ &= \operatorname{tr}(\boldsymbol{BA}).\end{aligned}$$

A×B的迹 = B×A的迹

Rank and Inverse

列的个数

A的秩数等于r的时候
称A是满秩的.

- The **rank** of the $k \times r$ matrix $(r \le k)$

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_r \end{bmatrix}$$

  is the number of linearly independent columns, written as $\operatorname{rank}(\boldsymbol{A})$. We say that $\boldsymbol{A}$ has full rank if $\operatorname{rank}(\boldsymbol{A}) = r$.

方块阵是满秩的是非奇异的矩阵.

- A $k \times k$ square matrix $\boldsymbol{A}$ is said to be **nonsingular** if it is has full rank, e.g. $\operatorname{rank}(\boldsymbol{A}) = k$. This means that there is no $k \times 1$ $\boldsymbol{c} \neq \boldsymbol{0}$ such that $\boldsymbol{A}\boldsymbol{c} = \boldsymbol{0}$.
- If $k \times k$ square matrix $\boldsymbol{A}$ is nonsingular then there exists a unique $k \times k$ matrix $\boldsymbol{A}^{-1}$ called the **inverse** of $\boldsymbol{A}$ which satisfies

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}_k.$$

非奇异的矩阵
满足性质.

- For non-singular $\boldsymbol{A}$ and $\boldsymbol{C}$, some important properties include

记住 *

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}_k$$
$$(\boldsymbol{A}^{-1})' = (\boldsymbol{A}')^{-1}$$
$$(\boldsymbol{A}\boldsymbol{C})^{-1} = \boldsymbol{C}^{-1}\boldsymbol{A}^{-1}$$
$$(\boldsymbol{A} + \boldsymbol{C})^{-1} = \boldsymbol{A}^{-1}(\boldsymbol{A}^{-1} + \boldsymbol{C}^{-1})^{-1}\boldsymbol{C}^{-1}$$
$$\boldsymbol{A}^{-1} - (\boldsymbol{A} + \boldsymbol{C})^{-1} = \boldsymbol{A}^{-1}(\boldsymbol{A}^{-1} + \boldsymbol{C}^{-1})^{-1}\boldsymbol{A}^{-1}.$$

Rank and Inverse

- If a $k \times k$ matrix $\boldsymbol{H}$ is orthonormal (so that $\boldsymbol{H}'\boldsymbol{H} = \boldsymbol{I}_k$), then $\boldsymbol{H}$ is non-singular and $\boldsymbol{H}^{-1} = \boldsymbol{H}'$. Furthermore, $\boldsymbol{H}\boldsymbol{H}' = \boldsymbol{I}_k$ and $\boldsymbol{H}'^{-1} = \boldsymbol{H}$.
- Another useful result for non-singular $\boldsymbol{A}$ is known as the **Woodbury matrix identity**

$$(\boldsymbol{A} + \boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{C}(\boldsymbol{C} + \boldsymbol{C}\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{C})^{-1}\boldsymbol{C}\boldsymbol{D}\boldsymbol{A}^{-1}.$$

- In particular, for $\boldsymbol{C} = -1$, $\boldsymbol{B} = \boldsymbol{b}$ and $\boldsymbol{D} = \boldsymbol{b}'$ for vector $\boldsymbol{b}$ we find what is known as the **Sherman-Morrison formula**

$$(\boldsymbol{A} - \boldsymbol{b}\boldsymbol{b}')^{-1} = \boldsymbol{A}^{-1} + (1 - \boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b})^{-1}\boldsymbol{A}^{-1}\boldsymbol{b}\boldsymbol{b}'\boldsymbol{A}^{-1}.$$

Rank and Inverse

- The following fact about inverting partitioned matrices is quite useful.

$$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{A}^{11} & \boldsymbol{A}^{12} \\ \boldsymbol{A}^{21} & \boldsymbol{A}^{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_{11\cdot 2}^{-1} & -\boldsymbol{A}_{11\cdot 2}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1} \\ -\boldsymbol{A}_{22\cdot 1}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & \boldsymbol{A}_{22\cdot 1}^{-1} \end{bmatrix},$$

  where $\boldsymbol{A}_{11\cdot 2} = \boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}$ and $\boldsymbol{A}_{22\cdot 1} = \boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}$.

- There are alternative algebraic representations for the components. For example, using the Woodbury matrix identity you can show the following alternative expressions

$$\begin{aligned} \boldsymbol{A}^{11} &= \boldsymbol{A}_{11}^{-1} + \boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22\cdot 1}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} \\ \boldsymbol{A}^{22} &= \boldsymbol{A}_{22}^{-1} + \boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11\cdot 2}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1} \\ \boldsymbol{A}^{12} &= -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22\cdot 1}^{-1} \\ \boldsymbol{A}^{21} &= -\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11\cdot 2}^{-1} \end{aligned}$$

Rank and Inverse

广义逆矩阵.

- Even if a matrix $\boldsymbol{A}$ does not possess an inverse, we can still define the **Moore-Penrose generalized inverse** $\boldsymbol{A}^-$ as the matrix which satisfies

$$①\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A} = \boldsymbol{A};\ ②\boldsymbol{A}^-\boldsymbol{A}\boldsymbol{A}^- = \boldsymbol{A}^-;\ ③\boldsymbol{A}\boldsymbol{A}^- \text{ is symmetric; } ④\boldsymbol{A}^-\boldsymbol{A} \text{ is symmetric.}$$

A^- 是存在且唯一的.

A可以是方矩阵、可以不是方矩阵. 是方矩阵时也可以是奇异的.

- For any (possibly non-square) matrix $\boldsymbol{A}$, the Moore-Penrose generalized inverse $\boldsymbol{A}^-$ exists and unique.
- ~~Another useful result for non-singular $\boldsymbol{A}$ is known as the **Woodbury matrix identity**~~
- An example:

$$\text{if } \boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}, \text{ and } \boldsymbol{A}_{11}^{-1} \text{ exists then } \boldsymbol{A}^- = \begin{bmatrix} \boldsymbol{A}_{11}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}.$$

A的列向量的所有线性组合的集合构成A的列空间.

- For any $k \times r$ matrix $\boldsymbol{A}$, the linear sub-space $\{\boldsymbol{x} \in \mathbb{R}^r : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$ is called the **null space**. The linear sub-space $\{\boldsymbol{A}\boldsymbol{x} : \boldsymbol{x} \in \mathbb{R}^r\}$ is called the **column space**, i.e., the set of vectors spanned by the columns of $\boldsymbol{A}$.

$A\vec{x} = 0$, $\vec{x} \in \mathbb{R}^r$

$\Rightarrow$ A的零空间是$\vec{x}$的集合.

- Consider the linear equations $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. Suppose that $\boldsymbol{b}$ is an element of the column space of $\boldsymbol{A}$. This means the equation system has a solution. Then $\boldsymbol{A}^-\boldsymbol{b}$ also solves the linear equations, i.e. $\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{b} = \boldsymbol{b}$ and furthermore, $\|\boldsymbol{A}^-\boldsymbol{b}\| \leq \|\boldsymbol{x}\|$ for any $\boldsymbol{x}$ that solves the equations.

若 $A\vec{x}_1 = 0$. $A\vec{x}_2 = 0$

$\vec{x}_1, \vec{x}_2 \in \{\vec{x} : A\vec{x} = 0\}$ 则有 $c_1\vec{x}_1 + c_2\vec{x}_2 \in \{\cdots\}$

Rank and Inverse

X如果没有多重共线性，
则 X 是满秩的.
⇒ rank(X) = n − 零空间的维数.

- Suppose we have a $n \times k$ matrix $\boldsymbol{X}$ with $n \geq k$. We have the following result

$$\text{rank}\,(\boldsymbol{X}\boldsymbol{X}') = \text{rank}\,(\boldsymbol{X}'\boldsymbol{X}) = \text{rank}\,(\boldsymbol{X}) \leq k.$$

$n \times n$ 的阵　　$k \times k$ 的阵

- rank $(\boldsymbol{X})$ is equal to the difference between k and the dimension of its null space. The null spaces of $\boldsymbol{X}$ and $\boldsymbol{X}'\boldsymbol{X}$ are the same: if $\boldsymbol{X}\boldsymbol{\alpha} = \mathbf{0}$, then $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\alpha} = \mathbf{0}$; if $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\alpha} = \mathbf{0}$, then $\boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\alpha} = \|\boldsymbol{X}\boldsymbol{\alpha}\|^2 = 0$ and therefore $\boldsymbol{X}\boldsymbol{\alpha} = \mathbf{0}$. Therefore, rank $(\boldsymbol{X}'\boldsymbol{X}) =$ rank $(\boldsymbol{X})$. Similarly, rank $(\boldsymbol{X}\boldsymbol{X}') =$ rank $(\boldsymbol{X}')$. Transposing a matrix does not change its rank: rank $(\boldsymbol{X}) =$ rank $(\boldsymbol{X}')$.

一个 k×r 的矩阵 A. 左乘 or 右乘
一个非奇异的矩阵，不改变它的秩.

- Similarly, we can show the following result: let $\boldsymbol{Q}, \boldsymbol{P}$ be non-singular matrices and $\boldsymbol{A}$ be a $k \times r$ matrix with rank rank $(\boldsymbol{A})$, then

$$\text{rank}\,(\boldsymbol{P}\boldsymbol{A}) = \text{rank}\,(\boldsymbol{A}\boldsymbol{Q}) = \text{rank}\,(\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q}) = \text{rank}\,(\boldsymbol{A})\,.$$

Determinant

- Let $\boldsymbol{A} = \left(a_{ij}\right)$ be a $k \times k$ matrix. Let $\pi = (j_1, \ldots, j_k)$ denote a permutation of $(1, \ldots, k)$. There are $k!$ such permutations. There is a unique count of the number of inversions of the indices of such permutations (relative to the natural order $(1, \ldots, k)$) and let $\varepsilon_\pi = +1$ if this count is even and $\varepsilon_\pi = -1$ if the count is odd. Then the **determinant** of $\boldsymbol{A}$ is defined as

$$\det \boldsymbol{A} = \sum_\pi \varepsilon_\pi a_{1j_1} a_{2j_2} \cdots a_{kj_k}.$$

- For example, if $\boldsymbol{A}$ is 2×2 then the two permutations of $(1, 2)$ are $(1, 2)$ and $(2, 1)$ for which $\varepsilon_{(1,2)} = 1$ and $\varepsilon_{(2,1)} = -1$. Thus

$$\begin{aligned} \det \boldsymbol{A} &= \varepsilon_{(1,2)} a_{11} a_{22} + \varepsilon_{(2,1)} a_{21} a_{12} \\ &= a_{11} a_{22} - a_{12} a_{21}. \end{aligned}$$

Determinant

行列式性质.

Theorem (A.7.1, Hansen)

Let $\boldsymbol{A} = (a_{ij})$ be a $k \times k$ matrix. Properties of the determinant

1. $\det \boldsymbol{A} = \det(\boldsymbol{A}')$
2. $\det(c\boldsymbol{A}) = c^k \det \boldsymbol{A}$
3. $\det(\boldsymbol{A}\boldsymbol{B}) = \det(\boldsymbol{B}\boldsymbol{A}) = (\det \boldsymbol{A})(\det \boldsymbol{B})$
4. $\det(\boldsymbol{A}^{-1}) = (\det \boldsymbol{A})^{-1}$ $|A^{-1}| = |A|^{-1}$
5. $\det \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{D} \end{bmatrix} = \det(\boldsymbol{A})(\det \boldsymbol{D})$ and $\det \begin{bmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} = \det(\boldsymbol{A})(\det \boldsymbol{D})$
6. $\det \boldsymbol{A} \neq 0$ if and only if $\boldsymbol{A}$ is nonsingular
7. If $\boldsymbol{A}$ is triangular (upper or lower), then $\det \boldsymbol{A} = \prod_{i=1}^k a_{ii}$
8. If $\boldsymbol{A}$ is orthonormal, then $\det \boldsymbol{A} = \pm 1$.

Eigenvalues 特征值.

特征值: $|\lambda I_k - A| = 0$ 的根
即为A的特征值.

特征向量: $(\lambda I_k - A)h = 0$
or $Ah = h\lambda$, 则h及其
线性组合是A的特征向量.

- The characteristic equation of a $k \times k$ square matrix $\boldsymbol{A}$ is

$$\det(\lambda \boldsymbol{I}_k - \boldsymbol{A}) = 0.$$

  The left side is a polynomial of degree k in λ so it has exactly k roots, which are not necessarily distinct and may be real or complex. They are called the **characteristic roots** or **eigenvalues** of $\boldsymbol{A}$.
- If is an eigenvalue of $\boldsymbol{A}$ then $\lambda \boldsymbol{I}_k - \boldsymbol{A}$ is singular so there exists a non-zero vector $\boldsymbol{h}$ such that $(\lambda \boldsymbol{I}_k - \boldsymbol{A})\boldsymbol{h} = \boldsymbol{0}$ or $\boldsymbol{A}\boldsymbol{h} = \boldsymbol{h}\lambda$. The vector $\boldsymbol{h}$ is called a **characteristic vector** or **eigenvector** of $\boldsymbol{A}$ corresponding to λ. They are typically normalized so that $\boldsymbol{h}'\boldsymbol{h} = 1$ and thus $\lambda = \boldsymbol{h}'\boldsymbol{A}\boldsymbol{h}$.
- Set $\boldsymbol{H} = [\boldsymbol{h}_1 \cdots \boldsymbol{h}_k]$ and $\boldsymbol{\Lambda} = \operatorname{diag}\{\lambda_1, \ldots, \lambda_k\}$. A matrix expression is $\boldsymbol{A}\boldsymbol{H} = \boldsymbol{H}\boldsymbol{\Lambda}$.

如果矩阵是对称的,其特征值一定是实数.

Eigenvalues

Theorem (A.8.1, Hansen)

Properties of eigenvalues. Let λ_i *and* $\boldsymbol{h}_i, i = 1, \ldots, k$, *denote the* k *eigenvalues and eigenvectors of a square matrix* $\boldsymbol{A}$.

1. $\det(\boldsymbol{A}) = \prod_{i=1}^{k} \lambda_i$ $\quad |A| = \lambda_1 \cdot \lambda_2 \cdots \lambda_k$
2. $\operatorname{tr}(\boldsymbol{A}) = \sum_{i=1}^{k} \lambda_i$
3. $\boldsymbol{A}$ is non-singular if and only if all its eigenvalues are non-zero.
4. If $\boldsymbol{A}$ has distinct eigenvalues, there exists a nonsingular matrix $\boldsymbol{P}$, such that $\boldsymbol{A} = \boldsymbol{P}^{-1}\boldsymbol{\Lambda}\boldsymbol{P}$ and $\boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1} = \boldsymbol{\Lambda}$.
5. The non-zero eigenvalues of $\boldsymbol{AB}$ and $\boldsymbol{BA}$ are identical.
6. If $\boldsymbol{B}$ is non-singular then $\boldsymbol{A}$ and $\boldsymbol{B}^{-1}\boldsymbol{AB}$ have the same eigenvalues.
7. If $\boldsymbol{Ah} = \boldsymbol{h}\lambda$ then $(\boldsymbol{I} - \boldsymbol{A}) = \boldsymbol{h}(1 - \lambda)$. So $\boldsymbol{I} - \boldsymbol{A}$ has the eigenvalue $1 - \lambda$ and associated eigenvector $\boldsymbol{h}$.

Eigenvalues

- Most eigenvalue applications in econometrics concern the case where the matrix $\boldsymbol{A}$ is real and symmetric. In this case all eigenvalues of $\boldsymbol{A}$ are real and its eigenvectors are mutually orthogonal. Thus $\boldsymbol{H}$ is orthonormal so $\boldsymbol{H}'\boldsymbol{H} = \boldsymbol{I}_k$ and $\boldsymbol{H}\boldsymbol{H}' = \boldsymbol{I}_k$. When the eigenvalues are all real it is conventional to write them in descending order $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$.

谱分析

正交化 ⇐

- **Spectral Decomposition**. If $\boldsymbol{A}$is a $k \times k$ real symmetric matrix, then $\boldsymbol{A} = \boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}'$ where $\boldsymbol{H}$ contains the eigenvectors and $\boldsymbol{\Lambda}$ is a diagonal matrix with the eigenvalues on the diagonal. The eigenvalues are all real and the eigenvector matrix satisfies $\boldsymbol{H}'\boldsymbol{H} = \boldsymbol{I}_k$. The decomposition can be alternatively written as $\boldsymbol{H}'\boldsymbol{A}\boldsymbol{H} = \boldsymbol{\Lambda}$.

Λ是个对角矩阵，对角线上元素是特征值.

奇异矩阵也可以正交化.

- If $\boldsymbol{A}$ is real, symmetric, and invertible, then by the spectral decomposition and the properties of orthonormal matrices, $\boldsymbol{A}^{-1} = \boldsymbol{H}'^{-1}\boldsymbol{\Lambda}^{-1}\boldsymbol{H}^{-1} = \boldsymbol{H}\boldsymbol{\Lambda}^{-1}\boldsymbol{H}'$. Thus the columns of $\boldsymbol{H}$ are also the eigenvectors of $\boldsymbol{A}^{-1}$, and its eigenvalues are $\lambda_1^{-1}, \ldots, \lambda_k^{-1}$.

Positive Definite Matrices

只要看到 $A \geqslant 0 \Rightarrow A$是半正定的.

- We say that a $k \times k$ real symmetric square matrix $\boldsymbol{A}$ is positive semi-definite if for all $\boldsymbol{c} \neq \mathbf{0}$, $\boldsymbol{c}'\boldsymbol{A}\boldsymbol{c} \geq 0$. This is written as $\boxed{\boldsymbol{A} \geq 0.}$
- We say that $\boldsymbol{A}$ is positive definite if for all $\boldsymbol{c} \neq \mathbf{0}$, $\boldsymbol{c}'\boldsymbol{A}\boldsymbol{c} > 0$. This is written as $\boldsymbol{A} > 0$.

$A > 0 \Rightarrow$ 正定的

Positive Definite Matrices

Theorem (A.9.1, Hansen)

Properties of positive semi-definite matrices

1. If $\boldsymbol{A} = \boldsymbol{G}'\boldsymbol{B}\boldsymbol{G}$ with $\boldsymbol{B} \geq 0$ and some matrix $\boldsymbol{G}$, then $\boldsymbol{A}$ is positive semi-definite. (For any $\boldsymbol{c} \neq 0$, $\boldsymbol{c}'\boldsymbol{A}\boldsymbol{c} = \boldsymbol{\alpha}'\boldsymbol{B}\boldsymbol{\alpha} \geq 0$ where $\boldsymbol{\alpha} = \boldsymbol{G}\boldsymbol{c}$.
   If $\boldsymbol{G}$ has full column rank and $\boldsymbol{B} > 0$, then $\boldsymbol{A}$ is positive ~~definite~~. semi-definite
2. If $\boldsymbol{A}$is positive definite, then $\boldsymbol{A}$is non-singular. Furthermore, $\boldsymbol{A}^{-1} > 0$.
3. $\boldsymbol{A} > 0$ if and only if it is symmetric and all its eigenvalues are positive.
4. By the spectral decomposition, $\boldsymbol{A} = \boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}'$ where $\boldsymbol{H}'\boldsymbol{H} = \boldsymbol{I}_k$ and $\boldsymbol{\Lambda}$ is diagonal with non-negative diagonal elements. All diagonal elements of $\boldsymbol{\Lambda}$ are strictly positive if and only if $\boldsymbol{A} > 0$.
5. The rank of $\boldsymbol{A}$ equals the number of strictly positive eigenvalues.
6. If $\boldsymbol{A} > 0$ then $\boldsymbol{A}^{-1} = \boldsymbol{H}\boldsymbol{\Lambda}^{-1}\boldsymbol{H}'$.
7. If $\boldsymbol{A} \geq 0$ we can find a matrix $\boldsymbol{B}$ such that $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{B}'$.We call $\boldsymbol{B}$ a matrix square root of $\boldsymbol{A}$ and is typically written as $\boldsymbol{B} = \boldsymbol{A}^{1/2}$. The matrix $\boldsymbol{B}$ need not be unique. One matrix square root is obtained using the spectral decomposition $\boldsymbol{A} = \boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}'$. Then $\boldsymbol{B} = \boldsymbol{H}\boldsymbol{\Lambda}^{1/2}\boldsymbol{H}'$ is itself symmetric and positive definite and satisfies $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{B}$.

如果一个矩阵是半正定的，可以找到它的平方根。

Idempotent Matrices 幂等矩阵

- A $k \times k$ square matrix $\boldsymbol{A}$ is idempotent if $\boxed{\boldsymbol{A}\boldsymbol{A} = \boldsymbol{A}.}$ For example, the following matrix is idempotent

$$\boldsymbol{A} = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}.$$

- If $\boldsymbol{A}$ is idempotent and symmetric with rank r, then it has r eigenvalues which equal 1 and $k - r$ eigenvalues which equal 0. To see this, by the spectral decomposition we can write $\boldsymbol{A} = \boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}'$ where $\boldsymbol{H}$ is orthonormal and $\boldsymbol{\Lambda}$ contains the eigenvalues. Then

$$\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A} = \boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}'\boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}' = \boldsymbol{H}\boldsymbol{\Lambda}^2\boldsymbol{H}'.$$

eg. $\underset{n\times k}{X}(X'X)^{-1}X'$

$tr(X(X'X)^{-1}X')$

$= tr(\underbrace{(X'X)(X'X)^{-1}}_{I_k})$

$= k$

- We deduce that $\boldsymbol{\Lambda}^2 = \boldsymbol{\Lambda}$ and $\lambda_i^2 = \lambda_i$ for $i = 1, \ldots, k$. Hence λ_i must equal either 0 or 1. Since the rank of $\boldsymbol{A}$ is r, and the rank equals the number of positive eigenvalues, it follows that

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0}_{k-r} \end{bmatrix}.$$

幂等阵的秩等于它的迹.

- $\operatorname{tr}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A})$ and $\boldsymbol{A}$ is positive semi-definite.

Idempotent Matrices

- If $\boldsymbol{A}$ is idempotent and symmetric with rank $r < k$ then it does not possess an inverse, but its Moore-Penrose generalized inverse takes the simple form $\boldsymbol{A}^{-} = \boldsymbol{A}$.
- If $\boldsymbol{A}$ is idempotent then $\boldsymbol{I} - \boldsymbol{A}$ is also idempotent.

A是幂等 ⇒ I-A 也是幂等.

- One useful fact is that if $\boldsymbol{A}$ is idempotent then

$$\boldsymbol{c}'\boldsymbol{A}\boldsymbol{c} \leq \boldsymbol{c}'\boldsymbol{c} \text{ and } \boldsymbol{c}'(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{c} \leq \boldsymbol{c}'\boldsymbol{c}.$$

To see this, note that both $\boldsymbol{c}'\boldsymbol{A}\boldsymbol{c}$ and $\boldsymbol{c}'(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{c}$ are non-negative and

$$\boldsymbol{c}'\boldsymbol{c} = \boldsymbol{c}'\boldsymbol{A}\boldsymbol{c} + \boldsymbol{c}'(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{c}.$$

Matrix Calculus 矩阵微积分

- Let $\boldsymbol{x} = (x_1, ..., x_k)'$ be $k \times 1$ and $g : \mathbb{R}^k \to \mathbb{R}$. We adopt the following notational convention: the vector derivative is

$$\frac{\partial}{\partial \boldsymbol{x}} g(\boldsymbol{x}) = \begin{pmatrix} \frac{\partial}{\partial x_1} g(\boldsymbol{x}) \\ \vdots \\ \frac{\partial}{\partial x_k} g(\boldsymbol{x}) \end{pmatrix}$$

and

- $$\frac{\partial}{\partial \boldsymbol{x}'} g(\boldsymbol{x}) = \begin{pmatrix} \frac{\partial}{\partial x_1} g(\boldsymbol{x}) & \cdots & \frac{\partial}{\partial x_k} g(\boldsymbol{x}) \end{pmatrix}.$$

- Let $\boldsymbol{A} = \left(a_{ij}\right)_{m \times n}$ be a $m \times n$ matrix and $g : \mathbb{R}^{m \times n} \to \mathbb{R}$. The derivative of $g(\boldsymbol{A})$ with respect to $\boldsymbol{A}$ is (by convention)

$$\frac{\partial}{\partial \boldsymbol{A}} g(\boldsymbol{A}) = \begin{pmatrix} \frac{\partial g(\boldsymbol{A})}{\partial a_{11}} & \cdots & \frac{\partial g(\boldsymbol{A})}{\partial a_{1n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial g(\boldsymbol{A})}{\partial a_{m1}} & \cdots & \frac{\partial g(\boldsymbol{A})}{\partial a_{mn}} \end{pmatrix}.$$

为什么要用矩阵微积分?

$Q(\vec{x}) = \frac{1}{2}\vec{x}'A\vec{x} - \vec{b}\vec{x}$

对称.非奇异.

求一阶条件?

$Q: \mathbb{R}^k \to \mathbb{R}$

$\frac{\partial Q(\vec{x})}{\partial \vec{x}} = A\vec{x} - \vec{b} = 0$

$\vec{x} = A^{-1}\vec{b}$

Matrix Calculus

记下来就行.

Theorem (A.15.1 Hansen)

Properties of matrix derivatives

1. $\dfrac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{a}'\boldsymbol{x}) = \dfrac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}'\boldsymbol{a}) = \boldsymbol{a}$
2. $\dfrac{\partial}{\partial \boldsymbol{x}'}(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{A}$ and $\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}'\boldsymbol{A}') = \boldsymbol{A}'$
3. $\dfrac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}) = (\boldsymbol{A} + \boldsymbol{A}')\boldsymbol{x}$ $[f(x)\cdot g(x)]' = f(x)'g(x) + f(x)g'(x)$
4. $\dfrac{\partial^2}{\partial \boldsymbol{x}\partial \boldsymbol{x}'}(\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{A} + \boldsymbol{A}'$ (3+2)
5. $\dfrac{\partial}{\partial \boldsymbol{A}}\operatorname{tr}(\boldsymbol{B}\boldsymbol{A}) = \boldsymbol{B}'$
6. $\dfrac{\partial}{\partial \boldsymbol{A}}\log\det(\boldsymbol{A}) = (\boldsymbol{A}^{-1})'$

Matrix Calculus

Let $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{x} \in \mathbb{R}^n$. Then

$$\begin{aligned}
\frac{\partial\left(\boldsymbol{a}'\boldsymbol{x}\right)}{\partial \boldsymbol{x}} &= \begin{pmatrix} \frac{\partial(\boldsymbol{a}'\boldsymbol{x})}{\partial x_1} \\ \vdots \\ \frac{\partial(\boldsymbol{a}'\boldsymbol{x})}{\partial x_n} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial(a_1x_1+\ldots+a_nx_n)}{\partial x_1} \\ \vdots \\ \frac{\partial(a_1x_1+\ldots+a_nx_n)}{\partial x_n} \end{pmatrix} \\
&= \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \\
&= \boldsymbol{a}.
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\left(\boldsymbol{a}'\boldsymbol{x}\right)}{\partial \boldsymbol{x}'} &= \begin{pmatrix} \frac{\partial(\boldsymbol{a}'\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial(\boldsymbol{a}'\boldsymbol{x})}{\partial x_n} \end{pmatrix} \\
&= (a_1, \ldots, a_n) \\
&= \boldsymbol{a}'.
\end{aligned}$$

Matrix Calculus

Let $\boldsymbol{A}$ be an $m \times n$ matrix,

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1' \\ \vdots \\ \boldsymbol{a}_m' \end{pmatrix},$$

where $\boldsymbol{a}_j \in \mathbb{R}^n$ for $j = 1, \ldots, m$.

$$\begin{aligned} \frac{\partial\left(\boldsymbol{A}\boldsymbol{x}\right)}{\partial \boldsymbol{x}'} &= \begin{pmatrix} \frac{\partial(\boldsymbol{a}_1'\boldsymbol{x})}{\partial \boldsymbol{x}'} \\ \vdots \\ \frac{\partial(\boldsymbol{a}_m'\boldsymbol{x})}{\partial \boldsymbol{x}'} \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{a}_1' \\ \vdots \\ \boldsymbol{a}_m' \end{pmatrix} \\ &= \boldsymbol{A}. \end{aligned}$$

Similarly, $\frac{\partial}{\partial \boldsymbol{x}}\left(\boldsymbol{x}'\boldsymbol{A}'\right) = \boldsymbol{A}'$.

Matrix Calculus

- Using "multiplication rule",

$$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}) = \frac{\partial \boldsymbol{x}'}{\partial \boldsymbol{x}}\boldsymbol{A}\boldsymbol{x} + \frac{\partial \boldsymbol{x}'\boldsymbol{A}'}{\partial \boldsymbol{x}}\boldsymbol{x} = (\boldsymbol{A} + \boldsymbol{A}')\,\boldsymbol{x}.$$

- By the definition of matrix multiplication and trace,

$$\operatorname{tr}(\boldsymbol{B}\boldsymbol{A}) = \sum_i \sum_j a_{ij} b_{ji} \text{ and } \frac{\partial}{\partial a_{ij}}\operatorname{tr}(\boldsymbol{B}\boldsymbol{A}) = b_{ji}.$$

  b_{ji} is the ij^{th} element of $\boldsymbol{B}'$.

- Let C_{ij} be the ij^{th} cofactor of $\boldsymbol{A}$. Laplace's expansion: for any i,

$$\det\boldsymbol{A} = \sum_{j=1}^{k} a_{ij} C_{ij}.$$

C_{ij}与a_{ij}无关.

Observe:

$$\frac{\partial}{\partial a_{ij}}\log\det(\boldsymbol{A}) = (\det\boldsymbol{A})^{-1}\frac{\partial}{\partial a_{ij}}\det\boldsymbol{A} = (\det\boldsymbol{A})^{-1} C_{ij}$$

$$\frac{\partial \log|A|}{\partial a_{ij}} = |A|^{-1}\frac{\partial |A|}{\partial a_{ij}} = |A|^{-1} C_{ij}$$

and $\boldsymbol{A}^{-1} = (\det\boldsymbol{A})^{-1}\boldsymbol{C}$.

Vector Norms

Given any vector space V (such as Euclidean space $\mathbb{R}^m$) a **norm** on V is a function $\rho : \mathrm{V} \longrightarrow \mathbb{R}$ with the properties

1. $\rho(c\boldsymbol{a}) = |c|\rho(\boldsymbol{a})$ for any real number c and $\boldsymbol{a} \in \mathrm{V}$
2. $\rho\,(\boldsymbol{a} + \boldsymbol{b}) \leq \rho\,(\boldsymbol{a}) + \rho\,(\boldsymbol{b})$
3. If $\rho\,(\boldsymbol{a}) = 0$ then $\boldsymbol{a} = \mathbf{0}$

Vector Norms

- The typical norm used for $\mathbb{R}^m$ is the **Euclidean norm**

$$\begin{aligned} \| \boldsymbol{a} \| &= (\boldsymbol{a}'\boldsymbol{a})^{1/2} \\ &= (\sum_{i=1}^{m} a_i^2)^{1/2}. \end{aligned}$$

- The p−norm ($p \geq 1$)

$$\| \boldsymbol{a} \|_p = (\sum_{i=1}^{m} |a_i|^p)^{1/p}.$$

Special cases are the Euclidean norm and the 1-norm:

$$\| \boldsymbol{a} \|_1 = \sum_{i=1}^{m} |a_i|.$$

- The "max-norm"

$$\| \boldsymbol{a} \|_\infty = \max(|a_1|, \ldots, |a_m|).$$

Vector Norms

- **Jensen's Inequality.** If $g(\cdot) : \mathbb{R} \longrightarrow \mathbb{R}$ is convex, then for any non-negative weights a_j such that $\sum_{j=1}^{m} a_j = 1$, and any real numbers $\boldsymbol{x}_j$

$$g(\sum_{j=1}^{m} a_j x_j) \leq \sum_{j=1}^{m} a_j g(x_j)$$

In particular, setting $a_j = 1/m$, then

$$g(\frac{1}{m}\sum_{j=1}^{m} x_j) \leq \frac{1}{m}\sum_{j=1}^{m} g(x_j)$$

If $g(\cdot) : \mathbb{R} \longrightarrow \mathbb{R}$ is concave then the inequalities are reversed.

- **Weighted Geometric Mean Inequality.** For any non-negative real weights a_j such that $\sum_{j=1}^{m} a_j = 1$, and any non-negative real numbers x_j

$$x_1^{a_1} x_2^{a_2} \cdots x_m^{a_m} \leq \sum_{j=1}^{m} a_j x_j.$$

Vector Norms

- **Loève's c_r Inequality.** For $r > 0$,
$$|\sum_{j=1}^{m} a_j |^r \leq c_r \sum_{j=1}^{m} | a_j |^r$$
where $c_r = 1$ when $r \leq 1$ and $c_r = m^{r-1}$ when $r \geq 1$. Special case: c_2 **Inequality.** For any $m \times 1$ vectors $\boldsymbol{a}$ and $\boldsymbol{b}$,
$$(\boldsymbol{a} + \boldsymbol{b})' (\boldsymbol{a} + \boldsymbol{b}) \leq 2\boldsymbol{a}'\boldsymbol{a} + 2\boldsymbol{b}'\boldsymbol{b}.$$

- **Hölder's Inequality.** If $p > 1, q > 1$, and $1/p + 1/q = 1$, then for any $m \times 1$ vectors $\boldsymbol{a}$ and $\boldsymbol{b}$,
$$\sum_{j=1}^{m} | a_j b_j | \leq \| \boldsymbol{a} \|_p \| \boldsymbol{b} \|_q .$$

- **Minkowski's Inequality.** For any $m \times 1$ vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, if $p \geq 1$, then
$$\| \boldsymbol{a} + \boldsymbol{b} \|_p \leq \| \boldsymbol{a} \|_p + \| \boldsymbol{b} \|_p .$$

- **Schwarz Inequality.** For any $m \times 1$ vectors $\boldsymbol{a}$ and $\boldsymbol{b}$,
$$| \boldsymbol{a}'\boldsymbol{b} | \leq \| \boldsymbol{a} \| \| \boldsymbol{b} \| .$$

Matrix Norms

- The **Frobenius norm** of an $m \times k$ matrix $\boldsymbol{A}$ is the Euclidean norm applied to its elements:

$$\begin{aligned} \| \boldsymbol{A} \|_F &= (\mathrm{tr}(\boldsymbol{A}'\boldsymbol{A}))^{1/2} \\ &= \left(\sum_{i=1}^{m} \sum_{j=1}^{k} a_{ij}^2 \right)^{1/2} . \end{aligned}$$

- When $m \times m$ $\boldsymbol{A}$ is real symmetric, then

$$\| \boldsymbol{A} \|_F = \left(\sum_{l=1}^{m} \lambda_l^2 \right)^{1/2} ,$$

where λ_l, $l = 1, \ldots, m$ are the eigenvalues of $\boldsymbol{A}$. To see this,

$$\| \boldsymbol{A} \|_F = (\mathrm{tr}\,(\boldsymbol{H\Lambda H'H\Lambda H'}))^{1/2} = (\mathrm{tr}\,(\boldsymbol{\Lambda\Lambda}))^{1/2} = \left(\sum_{l=1}^{m} \lambda_l^2 \right)^{1/2} .$$

- For any $m \times 1$ vectors $\boldsymbol{a}$ and $\boldsymbol{b}$,

$$\| \boldsymbol{ab}' \|_F = \mathrm{tr}\,(\boldsymbol{ba}'\boldsymbol{ab}')^{1/2} = (\boldsymbol{b}'\boldsymbol{ba}'\boldsymbol{a})^{1/2} = \| \boldsymbol{a} \| \| \boldsymbol{b} \|$$

and $\| \boldsymbol{aa}' \|_F = \| \boldsymbol{a} \|^2$.

Matrix Norms

- The **spectral norm** of an $m \times k$ matrix is

$$\| \boldsymbol{A} \|_2 = \left(\lambda_{max} \left(\boldsymbol{A}'\boldsymbol{A}\right)\right)^{1/2},$$

  where $\lambda_{\max}(\boldsymbol{B})$ denotes the largest eigenvalue of the symmetric matrix $\boldsymbol{B}$.
- If $\boldsymbol{A}$ is $m \times m$ and symmetric with eigenvalues λ_j then

$$\| \boldsymbol{A} \|_2 = \max_{j \le m} | \lambda_j | .$$

- Suppose $\boldsymbol{A}$ is $m \times k$ with rank r,

$$\| \boldsymbol{A} \|_2 \le \| \boldsymbol{A} \|_F \quad \text{and} \quad \| \boldsymbol{A} \|_F \le \sqrt{r} \| \boldsymbol{A} \|_2 .$$

Matrix Norms

- Given any vector norm $\|\cdot\|$, the **induced matrix norm** is

$$\| \boldsymbol{A} \| = \sup_{\boldsymbol{x}'\boldsymbol{x}=1} \| \boldsymbol{Ax} \| = \sup_{\boldsymbol{x}\neq 0} \frac{\| \boldsymbol{Ax} \|}{\| \boldsymbol{x} \|}.$$

The triangle inequality is satisfied:

$$\| \boldsymbol{A}+\boldsymbol{B} \| = \sup_{\boldsymbol{x}'\boldsymbol{x}=1} \| \boldsymbol{Ax}+\boldsymbol{Bx} \| \leq \sup_{\boldsymbol{x}'\boldsymbol{x}=1} \| \boldsymbol{Ax} \| + \sup_{\boldsymbol{x}'\boldsymbol{x}=1} \| \boldsymbol{Bx} \| = \| \boldsymbol{A} \| + \| \boldsymbol{B} \| .$$

- For any vector $\boldsymbol{x}$, $\| \boldsymbol{Ax} \| \leq \| \boldsymbol{A} \| \| \boldsymbol{x} \|$. The induced matrix norm satisfies this property which is a matrix form of the Schwarz inequality:

$$\| \boldsymbol{AB} \| = \sup_{\boldsymbol{x}'\boldsymbol{x}=1} \| \boldsymbol{ABx} \| \leq \sup_{\boldsymbol{x}'\boldsymbol{x}=1} \| \boldsymbol{A} \| \| \boldsymbol{Bx} \| = \| \boldsymbol{A} \| \| \boldsymbol{B} \| .$$

- The matrix norm induced by the Euclidean vector norm is the spectral norm

$$\sup_{\boldsymbol{x}'\boldsymbol{x}=1} \| \boldsymbol{Ax} \|^2 = \sup_{\boldsymbol{x}'\boldsymbol{x}=1} \boldsymbol{x}'\boldsymbol{A}'\boldsymbol{Ax} = \lambda_{max}\,(\boldsymbol{A}'\boldsymbol{A}) = \| \boldsymbol{A} \|_2^2$$