MAT 653: Statistical Simulation

Instructor: Dr. Wei Li

2023-08-27

Back **Updated:** 2023-08-27 Statistical Simulation, Wei Li

Some basics of matrix algebra

Notations

Convention (unless otherwise noted):

All vector are taken as column vectors by default, for example, if $x \in \mathbb{R}^n$, then we can write

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Generic capital letter will often denote a matrix, e.g. A, a $m \times n$ matrix,

$$A = \left[\begin{array}{c} a_1, a_2 \cdots a_n \end{array} \right]$$

with each a_i belonging to $\mathbb{R}^m$.

Suppose $e_i, i = 1, \dots, n$ is $n \times 1$ unit vector, with 1 in the ith position and zeros elsewhere, i.e., the identity $n \times n$ matrix can be written as

$$I_n = \left[\begin{array}{c} e_1, e_2 \cdots e_n \end{array} \right]$$

For A an $m \times n$ matrix, Then the ith column of A can be expresed as $A\mathbf{e}_i$, for $i = 1, \dots, n$.

Elementary definitions and results

Matrix-vector multiplication:

$$Ax = \begin{bmatrix} a_1, a_2 \cdots a_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = b$$

Ax is a linear combination of columns vectors of A. The coefficients of the linear combinations are stored in x.

Matrix-matrix multiplication: Let

$$A = \left[\; a_1, a_2 \cdots a_n \; \right]$$

a $\ell \times m$ matrix, and

$$C = \left[\; c_1, c_2 \cdots c_n \; \right]$$

a $m \times n$ matrix, then

$$AC = B = [b_1, \cdots, b_n] : \quad (\ell \times n)$$

where $b_j = Ac_j$, $j = 1, \dots, n$.

Range space (column space): The column space col(A) of A is the span of vectors $a_1, \dots, a_n$, i.e., $\mathrm{span}\{\mathrm{a_1, \dots, a_n}\}$ the set of all linear combinations of $a_1, \dots, a_n$.

$$\mathrm{Col}(A) = \{b : b = Ax \text{ for some } x \in \mathbb{R}^n\}$$

The vector $b \in \mathbb{R}^m$ belongs to Col(A) iff $\exists$ scalars $x_1, \dots, x_n$ such that $b = x_1 a_1 + \cdots + x_n a_n$.

Kernel space (null space): The kernel space of an $m \times n$ matrix A, written as ker(A), is the set of all solutions of the homogeneous equation $Ax = 0$. In set notation,

$$\ker(A) = \{x : x \in \mathbb{R}^n \text{ and } Ax = 0\}$$

Linear independence of a set of vectors: suppose we have p vectors in $\mathbb{R}^n$, say, $v_1, \dots, v_p$

- we call this collection $V = \{v_1, \dots, v_p\}$ to be linearly independent if whenever $x_1 v_1 + \cdots + x_p v_p = 0$, we have $x_1 = \cdots = x_p = 0$
- $V = \{v_1, \dots, v_p\}$ is said to be linearly dependent if $\exists$ some vector in V that is a linear combination of the other vectors in V.

Rank of a matrix: The **row rank** of a matrix is the maximum number of rows, thought of as vectors, which are linearly independent. Similarly, the **column rank** is the maximum number of columns which are linearly independent. A basic fact is show that the row and column ranks of a matrix are equal to each other. Thus one simply speaks of the **rank** of a matrix. For a $m \times n$ matrix A, we call it **full row rank** if it rows are linearly independent; **full column rank** if its columns are linearly independent.

Theorem (rank-nullity theorem): Let A be an $m \times n$ matrix, Then

$$n = \dim(Col(A)) + \dim(\ker(A)) = \mathrm{rank} + \dim(\ker(A))$$

where the dimension dim is the maximal number of linearly independent elements that span the space.

Determinant of a matrix: Determinant is a number associated with any square matrix; we'll write it as $det(A)$ or $|A|$. There are several equivalent ways to define determinant. The determinant encodes a lot of information about the matrix. A square matrix is invertible exactly when the determinant is non-zero. Example:

$$\left| \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right| = ad - bc$$

The determinant of a square matrix is equal to the product of its eigenvalues (see later for definition).

Inverse of a matrix: Suppose A is an $m \times m$ square matrix, the $m \times m$ matrix Z is said to be the inverse of A iff $AZ = I = ZA$. We then call A is invertible, denote its inverse by A^{-1}. A square matrix that is not invertible is called singular. When A is invertible, then its inverse is given by

$$A^{-1} = \frac{1}{|A|} adj(A)$$

where $adj(A)$ is the adjoint matrix of A.

A square matrix is invertible if and only if its determinant is not zero (i.e., its eigenvalues are all non-zero).

Example:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

if $ad - bc \neq 0$.

Theorems: A square matrix has a unique inverse, A^{-1} iff the column vector are linearly independent, i.e., no column vector of A is a linear combination of the others.

If A is invertible, then for any b vector, $Ax = b$ has a unique solution x which is $x = A^{-1}b$, the coefficients needed to represent b as a linear combination of columns of A.

Transpose of a matrix: The transpose of a matrix $A = [a_{i,j}]$ is given by $A^T = [a_{j,i}]$.

Inner prodcut: Suppose $x, y \in \mathbb{R}^n$, Euclidean inner product is $\langle x, y \rangle = x^T y = \sum_{i=1}^n x_i y_i$. The Euclidean norm of x is $\|x\| := \sqrt{\langle x, x \rangle} = \sqrt{\sum_{i=1}^n x_i^2}$. The angel α $(0 \leq \alpha \leq \pi)$ between the two vectors x, y is defined by $\cos(\alpha) = \frac{\langle x,y \rangle}{\|x\| \|y\|}$.

Trace of a matrix: The trace of a square matrix A is the sum of all the diagonal entries of A. Note: for compatiblae matrices, trace has cyclic property, namely,

$$\text{tr(ABCD)} = \text{tr(DABC)} = \text{tr(CDAB)}.$$

Orthogonal vectors: Two vectors v, w are orthogonal to each other if $\langle v, w \rangle = 0$.

orthogonal complement: The orthogonal complement of a subspace W in a vector space V is the set of all vectors in V that are orthogonal to every vector in W. It's denoted as $W^\perp$.

Column space and Kernel space: $\text{Col}(A)^\perp = \text{Ker}\left(A^T\right)$ and $\text{Row}(A)^\perp = \text{Ker}(A)$

Theorem: $S = \{v_1, \ldots, v_p\}$ be a set of nonzero orthogonal vectors in $\mathbb{R}^n$, then S is linearly independent.

Proof:

Orthogonal/orthonormal matrix: An $m \times n$ matrix U has orthonormal columns if $U^T U = I$.

An $m \times m$ square matrix P is called an orthogonal matrix if

$$PP^T = P^T P = I_m, \quad \text{or } P^{-1} = P^T.$$

Any square matrix with orthonormal columns is an orthogonal matrix, and such a matrix must have orthonormal rows too.

Fact: Multiplication by orthogonal matrices preserves inner product, hence length and angles. For this reason, orthogonal matrices are often called rotation matrices.

Check:

Positive definite matrix: A square and symmetric matrix A is called a positive definite matrix (denoted by $A > 0$), if $x^T Ax > 0$ for all $x \neq 0$. A positive definite matrix is invertible. But the converse is not true. Similarly, one can define negative definite matrix.

Positive semi-definite matrix: A square and symmetric matrix A is called a positive semi-definite matrix (denoted by $A \geq 0$), if $x^T Ax \geq 0$ for all $x \neq 0$. For a positive semi-definite matrix, it is invertible if and only if it is also positive definite. Similarly, one can define negative semi-definite matrix.

Indefinite matrix: A square and symmetric matrix A that is neither positive semi-definite nor negative semi-definite is called indefinite.

Projection of a vector on the span of another vector: the projection of a vector $y \in \mathbb{R}^n$ onto the span of another vector $x \in \mathbb{R}^n$ ($\|x\| \neq 0$), i.e., span(x) := {ax : a $\in \mathbb{R}$} is given by

$$P_x y = \frac{\langle x, y\rangle}{\langle x, x\rangle} x = \frac{xx'}{\langle x, x\rangle} y$$

which has the property that $\langle y - P_x y, P_x y\rangle = 0$

Proof: