

Linear Algebra

August 2018

Intro

Broad topic:

- ▶ Deals with general linear equations and systems of linear equations
- ▶ In social sciences, we often have to manipulate the representations of these equations in the form of vectors and matrices

In stats: variables and data.

Notation - Matrices and regressions

We can represent a *DV*:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

And several *IV*:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}$$

- ▶ We can represent data through vectors and matrices and then use them to compute coefficients of linear regressions.

Keep an eye on notation – It can be tricky.

- ▶ In academic books, scalars are written as small x ; vectors are typically bolded – \mathbf{x} ; matrices are written as X .
- ▶ Oftentimes, vectors are written as \vec{x} .

Scalars

Definition: A scalar is any single component or single element.

- Typically, we assume that a scalar is just a real number:

$$x_1 \in \mathbf{R}.$$

Vectors

Definition: A vector is a collection of scalars that has certain properties.

Two types:

1. Row: $(x_1 \dots x_2 \dots x_3 \dots x_n)$.

2. column: $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$

- ▶ The dimension of the vector is the N of components in that vector.

Graphical representation:

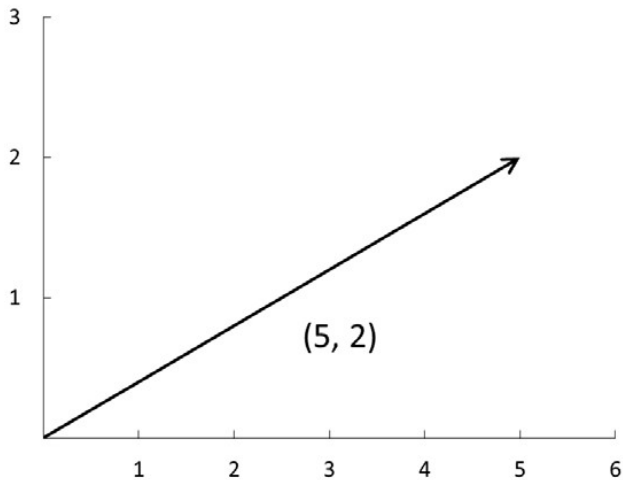


Figure 1: Vector $(5; 2)$ (*Source: M&S:277*)

How to manipulate vectors:

Vector addition:

$$\vec{a} + \vec{b} = \begin{pmatrix} 5 + 1 \\ 2 + 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}$$

$$\text{If } \vec{a} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \text{ and } \vec{b} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

- We must have the same N of dimensional elements in both vectors (in this case, 2 and 2).

Vector addition:

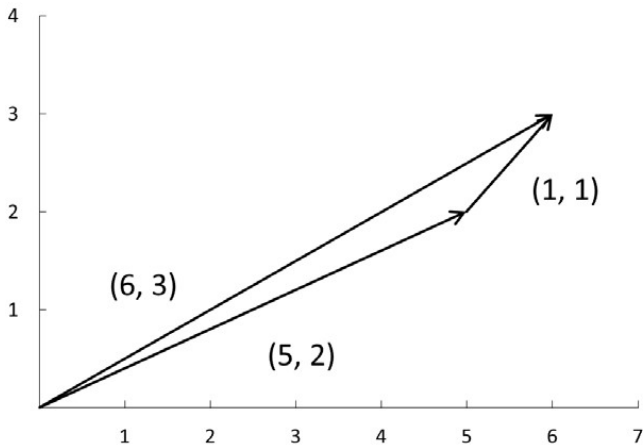


Figure 2: Vector Addition: $(5; 2) + (1; 1)$ (Source: M&S:279)

Scalar multiplication:

We can multiply scalars by vectors:

$$c * \vec{a} = 5 * \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 * 2 \\ 5 * 1 \end{pmatrix} = \begin{pmatrix} 10 \\ 5 \end{pmatrix}$$

$$\text{If } \vec{c} = 5 \text{ and } \vec{a} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Example:

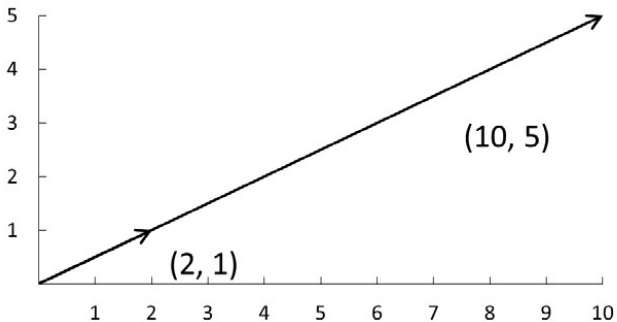


Figure 3: Scalar Multiplication: $5 \cdot \mathbf{a}$, where $\mathbf{a} = (2; 1)$ (Source: M&S:279)

Vector multiplication:

The **scalar product** is a way of multiplying vectors that results in a scalar.

- ▶ Also known as **dot product**: $\vec{a} \cdot \vec{b}$.

If a and b are both n -dimensional vectors, then:

$$\vec{a} \cdot \vec{b} = a_1b_1 + a_2b_2 + \dots + a_nb_n.$$

In other words, $\vec{a} \cdot \vec{b} = \sum_{i=1}^n a_ib_i$.

An example: $(3; 1) \cdot (2; 3) = 6 + 3 = 9$.

- ▶ The two vectors must have equal dimensions.

The dot product helps us to understand independence (perpendicularity) and collinearity.

How can vectors be useful?

Data are stored in vectors.

- ▶ Each variable has an associated vector, and each component of this vector is an observation or data point.
- ▶ Parameters are also stored in vectors.

Example: logistic regression returns a vector of coefficients → The effect of each of the IVs on the DV.

Examples (*Source*: mathsisfun.com)

Scalar

24

Vector

$\begin{bmatrix} 2 & -8 & 7 \end{bmatrix}$

row

or

column **$\begin{bmatrix} 2 \\ -8 \\ 7 \end{bmatrix}$**

Matrix

$\begin{bmatrix} 6 & 4 & 24 \\ 1 & -9 & 8 \end{bmatrix}$

row(s) × column

Figure 4: Scalar, vector and matrix

Matrices - introductory notions

A **matrix** is a way to represent a set of numbers. For example, the following set of numbers can be represented in matrix terms:

2, 3, 7, -7, 0, 4, 5, -2, 11, 12, -6, 5.

$$\begin{pmatrix} 2 \\ 3 \\ 7 \\ -7 \\ 0 \\ 4 \\ 5 \\ -2 \\ 11 \\ 12 \\ -6 \\ 5 \end{pmatrix}$$

$$\text{or as } \begin{pmatrix} 2 & 3 & 7 \\ -7 & 0 & 4 \\ 5 & -2 & 11 \\ 12 & -6 & 5 \end{pmatrix}$$

$$\text{or as } (2 \ 3 \ 7 \ -7 \ 0 \ 4 \ 5 \ -2 \ 11 \ 12)$$

A matrix is a way to write a group of scalars in an ordered fashion. A **matrix** is shaped like a rectangle and has a certain numbers of **rows** and a certain number of **columns**. The number of rows and columns of a matrix constitute the **dimensions** (sometimes called **order**) of the matrix. The dimensions of a matrix are denoted as $r \times c$, read “r by c”. The first number (“r here”) is the number of rows and the second number (“c” here) is the number of columns in the matrix. So, the first matrix is a 12×1 matrix, the second one is 4×3 , and the last one is 1×10 .

A 1×1 (“one by one”) matrix is a special case. We think of a 1×1 matrix as a simple scalar - just a number. $[3]$ is simply 3, for example.

We could write a $p \times q$ matrix \mathbf{B} , with elements b (notice there are p times $q = pq$ elements) as

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} & \dots & b_{1q} \\ b_{21} & b_{22} & b_{23} & \dots & b_{2q} \\ b_{31} & b_{32} & b_{33} & \dots & b_{3q} \\ \dots & \dots & \dots & \dots & \dots \\ b_{p1} & b_{p2} & b_{p3} & \dots & b_{pq} \end{pmatrix}$$

where, for example, b_{pq} is the element in the p^{th} row, q^{th} column of the matrix.

Two matrices are **equal** if they are of the same dimensions and each corresponding element is the same number. Mathematically, we say $\mathbf{P}=\mathbf{Q}$ if and only if \mathbf{P} and \mathbf{Q} have the same dimensions ($r \times c$) and $p_{ij} = q_{ij}$ for all $i=1,2,3,\dots,r$ and for all $j=1,2,3,\dots,c$.

A **row matrix** is a matrix with exactly one row. Sometimes a row matrix is called a **row vector**.

A **column matrix** is a matrix exactly one column. Sometimes a column matrix is called a **column vector**.

A **zero matrix** is simply a matrix with all elements equal to zero. A zero matrix can be of any dimension. Often a zero matrix will be noted by a bold-faced zero. An example:

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The **transpose of a matrix** is the original matrix with the columns and rows interchanged. We write the transpose of matrix **J** as either J' (pronounced “J prime”) or J^T (“J transpose”).

$$J = \begin{bmatrix} 2 & 3 & 7 \\ -7 & 0 & 4 \\ 5 & -2 & 11 \end{bmatrix} \quad J^T = \begin{bmatrix} 2 & -7 & 5 \\ 3 & 0 & -2 \\ 7 & 4 & 11 \end{bmatrix}$$

A **square matrix** is a matrix with the same number of rows and columns. An example (3 by 3 matrix):

$$B = \begin{bmatrix} 2 & 3 & 7 \\ -7 & 0 & 4 \\ 5 & -2 & 11 \end{bmatrix}$$

The **diagonal of a square matrix** is a set of numbers consisting of the elements on the line from the upper-left-hand to the lower-right-hand corner of the matrix. Only a square matrix has a diagonal. Diagonal of $B = \text{diag}(\mathbf{B}) = 2, 0, 11$.

The **trace of a matrix** is simply the sum of the diagonal elements of the matrix. So, then, a matrix must be a square matrix in order for it to have a Trace. Trace of $B = \text{trace}(\mathbf{B}) = \text{tr}(\mathbf{B}) = 2 + 0 + 11 = 13$.

A **symmetric matrix** is a square matrix that is equal to its transpose. In other words, a $p \times p$ matrix is symmetric if and only if $m_{jk} = m_{kj}$ for all k between 1 and p (where $1 \leq j \leq p$ and $1 \leq k \leq p$).

$$C = C^T = \begin{bmatrix} 3 & 1 & -4 \\ 1 & 9 & 0 \\ -4 & 0 & 7 \end{bmatrix}$$

An **upper triangular matrix** is a square matrix with all the elements below the diagonal equal to zero (matrix A). A **lower triangular matrix** is a square matrix with all the elements above the diagonal equal to zero (matrix B).

$$A = \begin{bmatrix} 2 & 3 & 7 \\ 0 & 5 & 4 \\ 0 & 0 & 11 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 0 & 0 \\ 3 & 4 & 0 \\ 7 & 4 & 9 \end{bmatrix}$$

A **diagonal matrix** is a square matrix that is both upper and lower triangular, that is, all the elements that are not on the diagonal are equal to zero (Matrix D).

$$D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

The **scalar matrix** is a diagonal matrix with the diagonal elements all equal to each other (M matrix). The **identity matrix** is a scalar matrix with all of the diagonal elements equal to one. We will reserve the capital letter I for the identity matrix (see below).

$$M = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

A **partitioned matrix (block matrix)** is a matrix made up of other matrices.

Matrix addition

$$\begin{bmatrix} 3 & 8 \\ 4 & 6 \end{bmatrix} + \begin{bmatrix} 4 & 0 \\ 1 & -9 \end{bmatrix} = \begin{bmatrix} 7 & 8 \\ 5 & -3 \end{bmatrix}$$

Figure 5: Matrix addition

Matrix subtraction

$$\begin{bmatrix} 3 & 8 \\ 4 & 6 \end{bmatrix} - \begin{bmatrix} 4 & 0 \\ 1 & -9 \end{bmatrix} = \begin{bmatrix} -1 & 8 \\ 3 & 15 \end{bmatrix}$$

Figure 6: Matrix subtraction

Multiplication by a constant

$$2 \times \begin{bmatrix} 4 & 0 \\ 1 & -9 \end{bmatrix} = \begin{bmatrix} 8 & 0 \\ 2 & -18 \end{bmatrix}$$

Figure 7: Multiplication by a constant

Matrix multiplication

We use the dot product by multiplying matching members:

Example (*Source* mathisfun.com):

What is $A \cdot B$? If $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ and $B = \begin{pmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{pmatrix}$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \times \begin{bmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{bmatrix} = \begin{bmatrix} 58 \\ \end{bmatrix}$$

Figure 8: 1st row and 1st column

$$(1, 2, 3) \cdot (7, 9, 11) = 1 \times 7 + 2 \times 9 + 3 \times 11 = 58$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \times \begin{bmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{bmatrix} = \begin{bmatrix} 58 & 64 \\ 139 & 154 \end{bmatrix}$$

Figure 9: 1st row and 2nd column

$$(1, 2, 3) \cdot (8, 10, 12) = 1 \times 8 + 2 \times 10 + 3 \times 12 = 64$$

We do it for the rest of rows and columns and we get:

$$(4, 5, 6) \cdot (7, 9, 11) = 4 \times 7 + 5 \times 9 + 6 \times 11 = 139$$

$$(4, 5, 6) \cdot (8, 10, 12) = 4 \times 8 + 5 \times 10 + 6 \times 12 = 154$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \times \begin{bmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{bmatrix} = \begin{bmatrix} 58 & 64 \\ 139 & 154 \end{bmatrix} \quad \checkmark$$

Figure 10: Answer

Let's look at a practical example:

Thomson House sells beef pies (\$3), chicken pies (\$4), and vegetable pies (\$2).

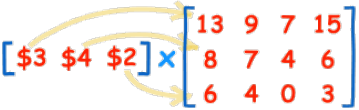
This is how many pies they sold from Monday to Thursday:

	Mon	Tue	Wed	Thu
Beef	13	9	7	15
Chicken	8	7	4	6
Vegetable	6	4	0	3

Figure 11: Pies

How much money did they make each day?

Answer:


$$\begin{bmatrix} \$3 & \$4 & \$2 \end{bmatrix} \times \begin{bmatrix} 13 & 9 & 7 & 15 \\ 8 & 7 & 4 & 6 \\ 6 & 4 & 0 & 3 \end{bmatrix} = \begin{bmatrix} \$83 & \$63 & \$37 & \$75 \end{bmatrix}$$

$\$3 \times 13 + \$4 \times 8 + \$2 \times 6$

Figure 12: Matrix solution

Matrix properties (1)

k, k_1, k_2 - Scalar (Number) **A, B, C** - Matrix **I** - Identity Matrix **0** - Zero Matrix **D** - Diagonal Matrix (Diagonal elements = d_{ii})

Addition and scalar multiplication: $A + B = B + A$.

$(A + B) + C = A + (B + C)$. $(k_1 + k_2)A = k_1A + k_2A$.

$k(A + B) = kA + kB$. $kA = Ak$. $0A = 0$. $k_1(k_2A) = (k_1k_2)A$.

$kAB = (Ak)B = a(kB) = ABk$.

Matrix multiplication: The conformability condition. In general,

$AB \neq BA$. $(AB)C = A(BC)$. $(A + B)C = AC + BC$.

$A(B + C) = AB + AC$. $IA = AI = A$. $0A = A0 = 0$. $AB \rightarrow$

pre-multiplying by A, post-multiplying by B, $A \rightarrow$ Lead, $B \rightarrow$ Lag.

Let's practice:

If $X = \begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix}$ and $Y = \begin{bmatrix} -2 & 4 \\ -3 & 1 \end{bmatrix}$, then what is XY ?

If $A = \begin{bmatrix} 3 & -1 & 2 \\ -2 & 4 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ -1 & 4 \\ -3 & 2 \end{bmatrix}$, then what is AB ?

If $A = \begin{bmatrix} 3 & -2 & 5 \\ 0 & -1 & 6 \\ -4 & 2 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & -1 & 0 \\ 3 & -5 & 2 \\ 1 & 4 & -2 \end{bmatrix}$, then what is AB ?

1.

If $X = \begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix}$ and $Y = \begin{bmatrix} -2 & 4 \\ -3 & 1 \end{bmatrix}$, then what is XY ?

Let's work it out:

$$\begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} -2 & 4 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} 3 \times (-2) + (-1) \times (-3) & 3 \times 4 + (-1) \times 1 \\ 2 \times (-2) + 5 \times (-3) & 2 \times 4 + 5 \times 1 \end{bmatrix} = \begin{bmatrix} -3 & 11 \\ -19 & 13 \end{bmatrix}$$

2.

If $A = \begin{bmatrix} 3 & -1 & 2 \\ -2 & 4 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ -1 & 4 \\ -3 & 2 \end{bmatrix}$, then what is AB ?

$$\begin{aligned} & \begin{bmatrix} 3 & -1 & 2 \\ -2 & 4 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ -1 & 4 \\ -3 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 3 \times 2 + (-1) \times (-1) + 2 \times (-3) & 3 \times 0 + (-1) \times 4 + 2 \times 2 \\ -2 \times 2 + 4 \times (-1) + 0 \times (-3) & -2 \times 0 + 4 \times 4 + 0 \times 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ -8 & 16 \end{bmatrix} \end{aligned}$$

3.

$$\text{If } A = \begin{bmatrix} 3 & -2 & 5 \\ 0 & -1 & 6 \\ -4 & 2 & -1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & -1 & 0 \\ 3 & -5 & 2 \\ 1 & 4 & -2 \end{bmatrix}, \text{ then what is } AB?$$

This is where it gets tedious:

$$\begin{aligned} AB &= \begin{bmatrix} 3 & -2 & 5 \\ 0 & -1 & 6 \\ -4 & 2 & -1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 \\ 3 & -5 & 2 \\ 1 & 4 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 3 \times 2 + (-2) \times 3 + 5 \times 1 & 3 \times (-1) + (-2) \times (-5) + 5 \times 4 & 3 \times 0 + (-2) \times 2 + 5 \times (-2) \\ 0 \times 2 + (-1) \times 3 + 6 \times 1 & 0 \times (-1) + (-1) \times (-5) + 6 \times 4 & 0 \times 0 + (-1) \times 2 + 6 \times (-2) \\ -4 \times 2 + 2 \times 3 + (-1) \times 1 & -4 \times (-1) + 2 \times (-5) + (-1) \times 4 & -4 \times 0 + 2 \times 2 + (-1) \times (-2) \end{bmatrix} \\ &= \begin{bmatrix} 5 & 27 & -14 \\ 3 & 29 & -14 \\ -3 & -10 & 6 \end{bmatrix} \end{aligned}$$

Another exercise

Suppose a company manufactures two kinds of furniture: chairs and sofas. Each chair requires 12 feet of wood, 2 bolts of cloth, and 5 pounds of feathers. Each sofa requires 30 feet of wood, 7 bolts of cloth, and 20 pounds of feathers. Suppose we must choose one, and only one, of two potential suppliers for the materials we will need.

Supplier 1 charges \$8 per foot for wood, \$11 per bolt for cloth, and \$6 per pound of feathers. Supplier 2 charges \$4 per foot for wood, \$23 per bolt for cloth, and \$5 per pound for feathers. With the goal of minimizing the total cost charged by the supplier this month, which supplier should the company use, given that 45 chairs and 30 sofas will be produced this month? Note that the total cost charged by each supplier is equal to the following multiplicative product: how much the supplier charges for each material times how much material is required for each product times how much of each product is being produced.

Solution

To represent how much each supplier charges for each material, let

$S =$

$$\begin{bmatrix} 8 & 11 & 6 \\ 4 & 23 & 5 \end{bmatrix}$$

.

To represent how much material is required for each product, let R

$=$

$$\begin{bmatrix} 12 & 30 \\ 2 & 7 \\ 5 & 20 \end{bmatrix}$$

.

To represent how much of each product is being produced, let $P =$

$$\begin{bmatrix} 45 \\ 30 \end{bmatrix}$$

.

The total cost charged by each supplier can be calculated by multiplying these three matrices.

$$C = SRP = \begin{bmatrix} 19,770 \\ 16,785 \end{bmatrix}$$

.

So, this month the company should use Supplier 2. The total cost will be \$16,785.

Determinants and condition numbers

The **determinant of a matrix** is a number that describes some characteristics of a matrix - $\det(a)$ or $|A|$.

Every square matrix has a determinant - and only a square matrix has a determinant. We sometimes pre-multiply or post-multiply a non-square matrix by its transpose and then calculate the determinant of the resulting square matrix.

Quite often the determinant is used to consider the nature of the (non)singularity of a matrix. - $\det(A) \neq 0$ - we have no perfect linear combination of columns or of rows $\rightarrow A$ is nonsingular (A^{-1} does exist). - $\det(A) = 0$ - we have at least one perfect linear combination of columns and of rows $\rightarrow A$ is singular (A^{-1} does not exist).

- ▶ $\det(A)$ is “close” to 0 - we come close to having at least one perfect linear combination of columns and of rows. A is technically non-singular (A^{-1} does exist), but it is almost singular. The matrix is badly scaled and ill-conditioned. The inverse is unstable: a tiny change in one of the elements of the original matrix can result in huge changes in some or all of the elements of its inverse. It might even cause the inverse to not exist.

Sometimes **the condition number** of a matrix ($\text{cond}(a)$ or $\kappa(A)$) is reported by the statistical software in the regression diagnostic. The condition number describes the properties related to a matrix's determinant value and to its inverse.

If the condition number is small (close to one), the matrix is well-conditioned. Its determinant values its large, the inverse of the matrix is stable. Large (much bigger than one), then the matrix is ill-conditioned. Its determinant value is small. The inverse of the matrix is unstable. Infinite, then the matrix is singular. Its determinant value is zero. The inverse of the matrix does not exist.

Calculating the determinant of a matrix

A =

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

.

$$\det(A) = a_{11} \cdot a_{22} - a_{21} \cdot a_{12}.$$

C =

$$\begin{bmatrix} 3 & 10 \\ 4 & 6 \end{bmatrix}$$

.

$$\det(C) = 3 \cdot 6 - 4 \cdot 10 = 18 - 40 = (-22).$$

The minor of an element is the determinant of the submatrix of that element. **The cofactor of a minor** is the signed minor of an element such that its sign is the same as the minor when the row and column numbers of the element have an even sum, and its sign is the opposite of the minor when the row and column numbers of the element have an odd sum.

The determinant of an $n \times n$ matrix, where $n > 2$, is the sum of the products of each element and its cofactor for any row or column.

Example:

C =

$$\begin{bmatrix} 1 & 4 & 3 \\ 1 & 2 & 0 \\ 2 & 3 & 1 \end{bmatrix}$$

.

We will work with column 2.

$$\det(C) = 4C_{12} + 2C_{22} + 3C_{32}.$$

$$M_{12} =$$

$$\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$$

.

$$M_{12} = 1 \cdot 1 - 2 \cdot 0 = 1.$$

Similarly, $M_{22} = -5$ and $M_{32} = -3$. Given the sign of the cofactors, the determinant of C will be calculated in the following way: $\det(C) = 4 \cdot (-1) + 2 \cdot (-5) + 3 \cdot (3) = -5$.

The inverse of a matrix

A square matrix that is not invertible is called singular; it has a determinant of zero. An invertible square matrix is called nonsingular, and its determinant is non-zero.

An $n \times n$ matrix, A , is invertible (only square matrices can be inverted) if one can find a second $n \times n$ matrix, B , such that the product AB and the product BA both produce the $n \times n$ identity matrix, $I_{n \times n}$. In such a situation, B is the inverse of A .

$A^{-1} = \left(\frac{1}{\det(A)}\right) \times C^T$. (C^T is the transpose of the matrix of cofactors of A).

$A =$

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 4 & 3 \\ -6 & -2 & 2 \end{bmatrix}$$

.

$$\det(A) = 1 \cdot M_{11} - 2 \cdot M_{12} + 1 \cdot M_{13} = 1(14) - 2(18) + 1(24) = 2.$$

$$M_{11} = 14, M_{12} = 18, M_{13} = 24, \dots, M_{33} = 4.$$

$$C = \begin{bmatrix} 14 & -18 & 24 \\ -6 & 8 & -10 \\ 2 & -3 & 4 \end{bmatrix} \quad C^T = \begin{bmatrix} 14 & -6 & 2 \\ -18 & 8 & 3 \\ 24 & -10 & 4 \end{bmatrix}$$

$$A^{-1} = \left(\frac{1}{\det(A)}\right) \times C^T = \frac{1}{2} \times C^T.$$

$$A^{-1} = \begin{bmatrix} 7 & -3 & 1 \\ -9 & 4 & \frac{-3}{2} \\ 12 & -15 & 2 \end{bmatrix}$$

.

Matrix properties (2)

Transpose: $(A^T)^T = A$. $(AB)^T = B^T A^T$. $(ABC)^T = C^T B^T A^T$.
 $(A^{-1})^T = (A^T)^{-1}$. AA^T and $A^T A \rightarrow$ each is square and symmetric.
 $I^T = I$. $\det(A^T) = \det(A)$. $\text{rank}(A^T) = \text{rank}(A)$. $(kA)^T = k(A)^T$.
 $(A + B)^T = A^T + B^T$. If $A^T A = 0$, then $A = 0$.

Determinant: $\det(AB) = [\det(A)][\det(B)]$. $\det(A) = \det(A^T)$.
 $\det(A^{-1}) = \frac{1}{\det(A)}$. $\det(kA) = k^n \det(A)$. (A is a $n \times n$ matrix).

Inversion: $(A^{-1})^{-1} = A$. $(AB)^{-1} = B^{-1}A^{-1}$.
 $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$. $(A^T)^{-1} = (A^{-1})^T$. $AA^{-1} = A^{-1}A = I$.
 $I^{-1} = I$. $\det(A^{-1}) = \frac{1}{\det(A)}$. $\text{rank}(A^T) = \text{rank}(A)$.
 $(kA)^{-1} = (\frac{1}{k})A^{-1}$. $A^{-k} = (A^k)^{-1}$. $D^{-1} \rightarrow$ also a diagonal matrix.
Diagonal elements = $\frac{1}{d_{ii}}$.

Table 12.2: Matrix and Vector Transpose Properties

Inverse	$(A^T)^T = A$
Additive property	$(A + B)^T = A^T + B^T$
Multiplicative property	$(AB)^T = B^T A^T$
Scalar multiplication	$(cA)^T = cA^T$
Inverse transpose	$(A^{-1})^T = (A^T)^{-1}$
If A is symmetric	$A^T = A$

Table 12.3: Matrix Determinant Properties

Transpose property	$\det(A) = \det(A^T)$
Identity matrix	$\det(I) = 1$
Multiplicative property	$\det(AB) = \det(A) \det(B)$
Inverse property	$\det(A^{-1}) = \frac{1}{\det(A)}$
Scalar multiplication ($n \times n$)	$\det(cA) = c^n \det(A)$
If A is triangular or diagonal	$\det(A) = \prod_{i=1}^n a_{ii}$

Table 12.4: Matrix Inverse Properties

Inverse	$(A^{-1})^{-1} = A$
Multiplicative property	$(AB)^{-1} = B^{-1}A^{-1}$
Scalar multiplication ($n \times n$)	$(cA)^{-1} = c^{-1}A^{-1}$ if $c \neq 0$

Figure 13: (*Source*: M&S:298)

Exercises:

1. Suppose that

$$A = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$$

.

$$B = \begin{bmatrix} 2 & 1 & 3 \\ 0 & 4 & 1 \end{bmatrix}$$

.

$$C = \begin{bmatrix} -2 & 7 \\ 6 & 0 \\ -1 & 5 \end{bmatrix}$$

.

1. Find the product matrix for each of the following matrices that are conformable for multiplication: AB , BA , AC , CA , BC , CB .
2. Should we be surprised that BC and CB are not equal to each other? Briefly explain why.

Solution

AB - not conformable for multiplication.

$$BA = \begin{bmatrix} 10 \\ 13 \end{bmatrix}$$

.

AC - not conformable for multiplication. CA - not conformable for multiplication.

$$BC = \begin{bmatrix} -1 & 29 \\ 23 & 5 \end{bmatrix}$$

.

$$CB = \begin{bmatrix} -4 & 26 & 1 \\ 12 & 6 & 18 \\ -2 & 19 & 2 \end{bmatrix}$$

2. We should not be surprised that $BC \neq CB$. BC is a 2×2 matrix while CB is a 3×3 matrix. Also, matrix multiplication is not commutative (the order matters).

2.

1. Find the trace of the matrix:

$$D = \begin{bmatrix} 3 & 5 & 8 \\ 2 & -4 & 1 \\ -1 & 2 & 9 \end{bmatrix}$$

.

2. What is the only type of matrix that has a trace? Why?

The trace is the sum of the diagonal elements of a matrix. So the trace is: $3 + (-4) + 9 = 8$.

Only a square matrix has a diagonal, so that is the only type of matrix that can have a trace.

3. A symmetric matrix is a square matrix that is equal to its Explain why.
4. If a matrix has dimensions 4×9 , then its transpose will have dimensions
5. Complete the following matrix so that it is symmetric.

$$E = \begin{bmatrix} 3 & .. & .. \\ 4 & 2 & 8 \\ 5 & .. & 1 \end{bmatrix}$$

3. Transpose. Interchanging the rows and the columns does not matter in a symmetric matrix, since the element in the i^{th} row, j^{th} column is the same as the element in the j^{th} row, i^{th} column.
4. 9×4 . Since rows and column interchange roles when you take the transpose.
- 5.

$$E = \begin{bmatrix} 3 & 4 & 5 \\ 4 & 2 & 8 \\ 5 & 8 & 1 \end{bmatrix}$$

symmetry.

6. Recall that I_p is the notation for a $p \times p$ identity matrix. Let

$$F = \begin{bmatrix} 3 & 6 \\ 2 & 8 \\ -4 & 1 \end{bmatrix}$$

1. Find the value for p so that $I_p F = F$.
2. Find the value for p so that $F I_p = F$.

1. $p = 3$

2. $p = 2$

7. 7.1 Find AA' and also find A^T if

$$A = \begin{bmatrix} 4 & 0 & 2 \\ 3 & 1 & 5 \end{bmatrix}$$

. symmetry.

2. Is it a coincidence that both AA' and $A^T A$ are a square and symmetric matrix?

1.

$$AA^T = \begin{bmatrix} 20 & 22 \\ 22 & 35 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 25 & 3 & 23 \\ 3 & 1 & 5 \\ 23 & 5 & 29 \end{bmatrix}$$

2. Not a coincidence. Anytime you pre-multiply or post-multiply a matrix by its transpose the result always will be a square, and also a symmetric matrix.

8. Calculate the determinant of each matrix.

$$A = \begin{bmatrix} 6 & 5 \\ 3 & 2 \end{bmatrix}$$

.

$$B = \begin{bmatrix} -1 & 5 & -2 \\ 1 & 2 & 3 \\ 0 & 4 & -3 \end{bmatrix}$$

.

$$\det(A) = (6) \cdot (2) - (3) \cdot (5) = -3. \quad \det(B) = 25.$$

9. matrixcalc.org

Input the matrix

$$C = \begin{bmatrix} 6 & 3 & 5 \\ 1 & 4 & 0 \\ 2 & 1 & 1 \end{bmatrix}$$

1. Use the matrix calculator to find the determinant of C. What does the fact that the determinant is non-zero imply about the nature of C?
2. Use the matrix calculator to find the rank of C. Is C of full rank? What does that mean?
3. Use the matrix calculator to find the inverse of C.

1. $\det(C) = -14$. C is of full rank, nonsingular, and has an inverse.
2. $\text{rank}(C) = 3$. Since $\text{rank}(C) = \min(\# \text{rows}, \# \text{columns})$, we know that C is of full rank. That means that no row (column) of the square matrix C is a perfect linear combination of the other rows (columns). Therefore, C is non-singular and it has an inverse.
- 3.

$$C^{-1} = \begin{bmatrix} \frac{-2}{7} & \frac{-1}{7} & \frac{10}{7} \\ \frac{1}{14} & \frac{2}{7} & \frac{-5}{14} \\ \frac{1}{2} & 0 & \frac{-3}{2} \end{bmatrix}$$

The Kronecker product (tensor product)

The result of the Kronecker product ($A \otimes B$) is a “block matrix” (sometimes called a “partitioned matrix”). It is one huge matrix whose elements consist of smaller matrices (as opposed to mere numbers).

One purpose of the Kronecker product is to more efficiently represent a big matrix that has a pattern as you go from block to block within that big matrix. This can be especially useful if you're trying to solve (or optimise) a function in which the unknown is a matrix as opposed to a simple single variable.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{22} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{22} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & a_{13}B \\ a_{21}B & a_{22}B & a_{22}B \\ a_{31}B & a_{32}B & a_{33}B \end{bmatrix}$$

Matrix illustration of OLS estimation

$$y = \alpha + \beta X + \epsilon.$$

OLS regression proposes that the best fit is produced by selecting the values of α and β that minimize the sum of squared errors $(\sum_i (\epsilon_i)^2)$. You will likely learn that the OLS estimator is the best linear unbiased estimator (BLUE) in your statistics coursework.

It turns out that we can calculate a matrix that contains the values of α and β , call it $\hat{\beta}$, by using the equation: $\hat{\beta} = (X^T X)^{-1} X^T y$.

Vector spaces and systems of equations

Vector spaces contain, for example, the points on a line (one-dimensional space, 1-D), or in a plane (two-dimensional space, 2-D), or in three-dimensional (3-D) space.

A linear combination of any number of vectors in general looks like this: $a_1 x_1 + \dots + a_n x_n$, where the x_i are all vectors in some vector space and the a_i are all scalars. In words, a linear combination means that we add (or subtract) scalar multiples of vectors to get some new vector.

The columns (rows) are linearly independent if each column (row) is not equal to a linear combination of the others.

The rank of a matrix is the maximum number of linearly independent rows or columns; these are always the same number.

For square matrices, the rank is also useful in determining linear independence of vectors. If the rank of a square matrix is equal to its number of rows or columns, then the matrix is nonsingular. Since nonsingular matrices have non-zero determinants, one way of checking whether a set of n n -dimensional is linearly independent is to create a matrix out of the set and then take the determinant of this matrix. If the determinant is non-zero, the vectors are linearly independent; otherwise, the vectors are linearly dependent.

If the rank of a matrix is equal to the number of rows in the matrix or the number of columns in the matrix, then the matrix is of **full rank**. If the rank of a matrix is less than both the number of rows in the matrix and the number of columns in the matrix, then the matrix is not of full rank (**rank deficient**).

A is of full rank $\rightarrow A^T A$ is of full rank, $(A^T A)^{-1}$ does exist, $A^T A$ is nonsingular, $\det(A^T A) \neq 0$, all of the eigenvalue of $A^T A \neq 0$.

B is not of full rank $\rightarrow B^T B$ is not of full rank, $(B^T B)^{-1}$ does not exist, $B^T B$ is singular, $\det(B^T B) = 0$, At least one eigenvalue of $B^T B = 0$.

Solving systems of equations

$$\begin{aligned} 2x - y + 3z &= 9 \\ x + 4y - 5z &= -6 \\ x - y + z &= 2 \end{aligned}$$

1. Substitution

- ▶ Systems in which all the equations are different (i.e., not multiples of each other or linear combinations of other equations) and there is the same number of equations as unknowns often have unique solutions. We say the system in this case is **uniquely determined**.
- ▶ Systems in which there are more unknowns than equations often have an infinite number of solutions, as do systems in which one of the equations can be formed from some linear combination of the others, but the equations are not contradictory. We say the system in this case is **underdetermined**.

- ▶ There are cases when there are more equations than unknowns, or any time in which some of the equations are contradictory, which is more likely to happen the more equations there are. We say the system in this case is **overdetermined**.

$x = 2 + y - z$. Plug this x into equations one and two. In the end, $x = 1, y = 2, z = 3$.

2. Elimination

First, (-2) times equation two plus equation one produces an equation with no x , since $2x - y + 3z - 2(x + 4y - 5z) = 9 - 2(-6)$, which reduces to $-9y + 13z = 21$.

Second, we can do the same thing with (-2) times equation three plus equation one, since $2x - y + 3z - 2(x - y + z) = 9 - 2(2)$, which reduces to $y + z = 5$.

Thus our equations are now:

$$2x - y + 3z = 9, \quad -9y + 13z = 21, \quad y + z = 5.$$

Next we want to eliminate y from the last equation, which we can do by multiplying 9 times the last equation and adding it to the second equation, since $-9y + 13z + 9(y + z) = 21 + 9(5)$, which reduces to $22z = 66$. Now our equations are

$2x - y + 3z = 9, \quad -9y + 13z = 21, \quad 22z = 66$. In the end, we get the same solutions.

3. Matrix inversion

Rewrite the system of equation:

The matrix of coefficients:

$$A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 4 & -5 \\ 1 & -1 & 1 \end{bmatrix}$$

.

and the vector of solutions:

$$b = \begin{bmatrix} 9 \\ -6 \\ 2 \end{bmatrix}$$

.

$$Ax = B. \quad A^{-1}Ax = A^{-1}b. \quad \text{but } A^{-1}A = I. \quad x = A^{-1}b.$$

Remember: $A^{-1} = (\frac{1}{\det(A)}) \times C^T$.

$\det(A) = -11$. Then we compute nine minors and take into account the sign.

$$A^{-1} = (\frac{1}{-11}) \times \begin{bmatrix} 1 & -2 & -7 \\ -6 & -1 & 13 \\ -5 & 1 & 9 \end{bmatrix}$$

.

We multiply by the solution vector and we obtain the same results.

Exercises

1. Solve the following systems of equations using substitution or elimination, or both: $x - 3y + 6z = -1$. $2x - 5y + 10z = 1$. $3x - 8y + 17z = 1$.
2. Solve the following systems of equations using substitution or elimination, or both: $4x + 2y - 3z = 1$. $6x - 3y - 5z = 0$. $x = y + 2z = 9$.

Let

$$D = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

.

$$E = \begin{bmatrix} 3 & 1 \\ 6 & 2 \end{bmatrix}$$

.

$$g = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$$

.

$$g = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

.

1. Solve $Dx = g$ for x using matrix inversion.
2. Solve $xE = h$ for x using matrix inversion.

1. It is not possible as stated, since g has dimension three and D is of rank two.
2. Since E has determinant 0 it is singular, and does not have an inverse. Thus we can't use matrix inversion. As one of the rows is a multiple of the other we only have one unique equation for two variables, indicating an infinite number of solutions to the system of equations represented by the matrix.

Eigenvalues and eigenvectors

We define the **eigenvalues** of the square matrix A as a set of λ values that satisfy the expression $AX = \lambda X$, (X is a column vector, λ is a scalar). The equation is equivalent with $(A - \lambda)X = 0$. Given that λ is a scalar, we introduce the identity matrix and re-write the equation as $(A - \lambda I_n)X = 0$.

The characteristic equation of A involves a determinant; more specifically, $\det(A - \lambda I_n) = 0$. The values (up to "n" of them) of λ that satisfy the equation are the eigenvalues for the A matrix.

Then substituting each eigenvalue (i.e. each λ) back into the $(A - \lambda I_n)X = 0$ equation and solving for X vector gives us the **eigenvectors** for the A matrix.

Let

$$A = \begin{bmatrix} 3 & 10 \\ 4 & 6 \end{bmatrix}$$

.

$n = 2$.

$$A - \lambda I_2 = \begin{bmatrix} 3 & 10 \\ 4 & 6 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 10 \\ 4 & 6 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

The final result is

$$\begin{bmatrix} 3 - \lambda & 10 \\ 4 & 6 - \lambda \end{bmatrix}$$

$$\det(A - \lambda I_2) = (\lambda - 11)(\lambda + 2).$$

Given that $\det(A - \lambda I_2) = 0$. $\rightarrow \lambda = 11$ and $\lambda = -2$.

So the eigenvalues are $\lambda = 11$ and $\lambda = -2$.

Now we will plug each of these eigenvalues into the $(A - \lambda I_n)X = 0$ equation to find the eigenvectors of A. We obtain, for $\lambda = 11$,

$$X = \begin{bmatrix} (5/4)k \\ k \end{bmatrix}$$

, where $k = x_2$ (x_1 x_2 are the two solutions we are looking for).
For $\lambda = -2$,

$$X = \begin{bmatrix} -2k \\ k \end{bmatrix}$$

.

Spectral decomposition, Eigenanalysis, Eigendecomposition.

Eigenvalues properties

- ▶ The sum of the eigenvalues of a matrix is equal to the trace of that matrix.
- ▶ The product of the eigenvalues of a matrix is equal to the determinant of that matrix.
- ▶ The rank of a matrix is the same as the number of nonzero eigenvalues of that matrix.
- ▶ If a matrix has at least one eigenvalue equal to zero, then it is singular (i.e. has no inverse). Similarly, a matrix is nonsingular (has an inverse) if and only if it has no eigenvalues equal to zero.

- ▶ A matrix with at least one eigenvalue close to zero is called ill-conditioned (ill-behaved matrix). Otherwise, the matrix is said to be a well-conditioned (well-behaved matrix). The **condition number** is the absolute value of the largest divided by the smallest eigenvalue. The **condition index** is simply the square root of the condition number. The larger the condition number, the more ill-conditioned the matrix is. An ill-conditioned matrix can be problematic, since a trivially small change in that matrix can result in substantive changes in its inverse (and in the related calculations).

Exercise

Find the eigenvalues and corresponding eigenvectors for

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

Bibliography

Moore, W. H., & Siegel, D. A. (2013). A mathematics course for political and social research. Princeton University Press.

Kropko, J. (2015). Mathematics for social scientists. SAGE Publications.

ICPSR 2018 course on linear algebra (Tim McDaniel).

www.mathopolis.com

www.mathsisfun.com.