Lecture 2: Matrix Algebra

Thomas Chadefaux

Contents

1 Matrices and Vectors

- A matrix is a rectangular array of numbers. For example, $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$.
  Its dimensions are (rows) $\times$ (columns). E.g., the dimensions of $\mathbf{A}$ are 2×3.

$$A_{m,n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

- a vector is a matrix with only one row or column. For example, $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

- Matrices are typically represented with bold capital letters (e.g., $\mathbf{A}$). Vectors are denoted by bold lower case letters (e.g., $\mathbf{x}$).

- An element of a matrix is denoted by a_{ij}, where i refers to the row and j to the column of the matrix.

In-class exercise: in R, Create a 4×4 matrix, and populate it with values drawn from a standard normal distribution. Save this matrix as $\mathbf{A}$.

2 Matrix operations

2.1 Addition and Subtraction

$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$; $\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}$

Both matrices must have the same dimensions. Here, both $\mathbf{A}$ and $\mathbf{B}$ are 2×3, so

$\mathbf{A} + \mathbf{B} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} \end{pmatrix}$

For example:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix} \tag{1}$$

Note that $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$

In-class exercise: in R, calculate $A + A$ (where A is the matrix you created in the previous exercise.

2.2 Multiplication

$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$; $\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{pmatrix}$

Inner dimensions must match: $(r \times c) \times (r \times c)$

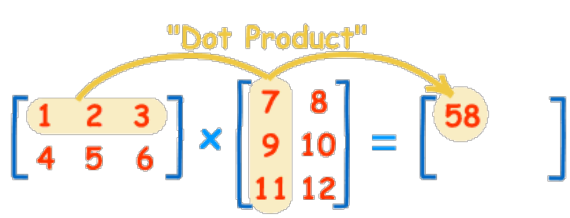


Figure 1: Matrix multiplication

$\mathbf{A}$ is 2×3 and $\mathbf{C}$ is 3×2, so the inner dimensions match, and we can calculate $\mathbf{D} = \mathbf{A} \times \mathbf{C}$ as:

$$\mathbf{D} = \begin{pmatrix} a_{11}c_{11} + a_{12}c_{21} + a_{13}c_{31} & a_{11}c_{12} + a_{12}c_{22} + a_{13}c_{32} \\ a_{21}c_{11} + a_{22}c_{21} + a_{23}c_{31} & a_{21}c_{12} + a_{22}c_{22} + a_{23}c_{32} \end{pmatrix}$$

The dimensions of the resulting matrix is 2×2, i.e., the outer dimensions: $(r \times c) \times (r \times c)$

CAREFUL: $\mathbf{A} \times \mathbf{C} \neq \mathbf{C} \times \mathbf{A}$

2.3 Order of operations

This is important and hence worth repeating: matrix multiplication is non-commutative. i.e., the order matters: $\mathbf{AB} \neq \mathbf{BA}$. However, $(AB)C = A(BC)$. I.e., matrix multiplication is associative. Also, $A(B+C) = AB + AC$ and $(A+B)C = AC + BC$.

In-class exercise: in R, create another 4×4matrix B, also with values drawn from a standard normal distribution. Calculate AB and BA.

For example, say we want to solve

$$\begin{aligned} \mathbf{Ax} &= \mathbf{c} \\ \mathbf{A}^{-1}\mathbf{Ax} &= \mathbf{A}^{-1}\mathbf{c} \quad \text{First pre-multiply by} \mathbf{A} \\ \mathbf{Ix} &= \mathbf{A}^{-1}\mathbf{c} \\ \mathbf{x} &= \mathbf{A}^{-1}\mathbf{c} \end{aligned}$$

2.4 Transpose

The transpose of a matrix $\mathbf{A}$ is denoted $\mathbf{A}^T$ or $\mathbf{A}'$. To transpose, switch the row and column positions of the matrix, such that : a_{ij} becomes a_{ji}.

In-class exercise: Find the transpose of the following matrices: $x_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$; $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{pmatrix}$; $B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}$.

In-class exercise: Let $x_2 = c(2,4,1,7)$. Using R, find $x'x$ and xx'. Hint: $\% * \%$ is the matrix multiplication operator, and t() is the function to transpose a matrix.

$$\text{If } \mathbf{A} = \begin{bmatrix} 1 & 5 & 9 \\ 2 & 3 & 8 \\ 6 & 8 & 7 \end{bmatrix}, \text{ then } \mathbf{A}^T = \begin{bmatrix} 1 & 2 & 6 \\ 5 & 3 & 8 \\ 9 & 8 & 7 \end{bmatrix}$$

Figure 2: Transpose of a matrix

This will be useful throughout. For example the sum of squares can be written:

$$\mathbf{x}'\mathbf{x} = \sum_{i=1}^{n} x_i^2.$$

Note that this is very different from $\mathbf{xx}'$. For example:

$$x_1'x_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = 30$$

But

$$x_1x_1' = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \\ 4 & 8 & 12 & 16 \end{pmatrix}.$$

3 Identity matrix

This is an important kind of a matrix. It is simply a square matrix with ones on the main diagonal and 0 everywhere else.

If the size of I is not specified, it is assumed to be conformable, i.e., as big as needed. Note that any matrix multiplied by I returns itself: $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$

$$I = I_r = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

4 Determinant

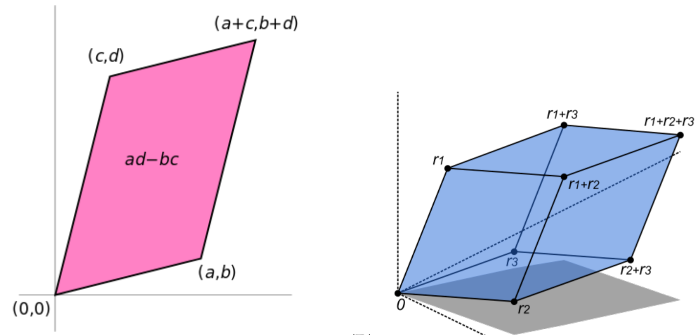
The determinant of a matrix $\mathbf{A}$, denoted $|\mathbf{A}|$ or $\det(\mathbf{A})$, summarises the matrix in a way that is useful for systems of linear equations. It also helps us find the inverse of a matrix. The intuition for the determinant is that it is the area (for a 2x2 matrix) or the volume (for a 3x3 matrix).

The determinant of a 2×2 matrix is defined by

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

The determinant of a 3×3 matrix is defined by

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a(ei - fh) - b(di - fg) + c(dh - eg)$$



(A) (B)
Figure 3: (A) Interpretation of determinant of a 2×2 matrix (A) and a 3×3 matrix (B).

In-class exercise: Calculate

$$\begin{vmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{vmatrix}$$

It is not important that you remember how to calculate the determinant. You will never need to do it. But you should understand the intuition. Why should you care? Among other reasons:

- If the determinant of a matrix is 0, it has no inverse. This will be a problem when we need to calculate $(X'X)^{-1}X'y$, since $(X'X)^{-1}$ will be undefined.
- If the rows or columns of $\mathbf{A}$ are linearly dependent, then $|\mathbf{A}| = 0$. This will matter because we will assume that X, our matrix of independent variables, has full rank, i.e., its rows and columns are linearly independent. If they are not, the determinant will be 0 and the inverse will not exist.

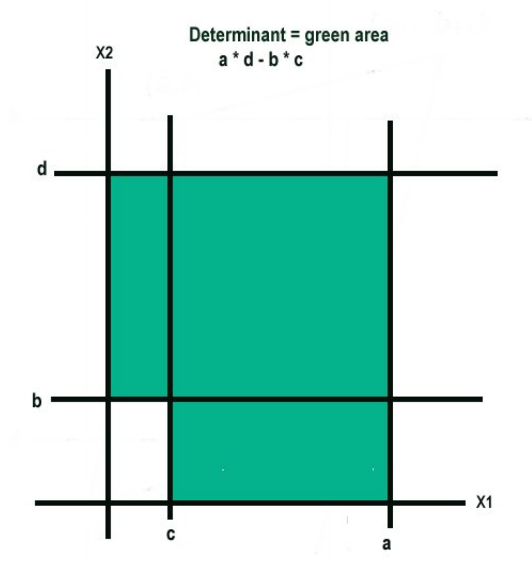


Figure 4: Calculating the determinant.

5 Inverse

The inverse of a matrix is (loosely) the equivalent of division in matrix-world. We say that a matrix is invertible or nonsingular if its inverse exists, i.e. if A^{-1} exists. Note that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. Note that only square matrices have an inverse.

Finding the inverse of a matrix is not easy, but softwares do it for you. Inverses will be needed when we want to calculate $(X'X)^{-1}X'y$

6 Rank

The rank of a matrix **A** is the maximal number of linearly independent rows or columns of **A**. To check independence, you can try to calculate the determinant (if 0, then independence is violated). The rank is the number of independent rows or columns (whichever is smaller).

When could this matter?

Example:

$$\begin{pmatrix} v_1 & v_2 & v_3 & v_4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 2 & 4 & 5 & 1 \\ 3 & 6 & 7 & 0 \\ 4 & 8 & 2 & 2 \end{pmatrix}$$

v_1, v_2 and v_3 are independent, but $v_4 = v_1 - v_2 - v_3$

For example, try this is R:

```
#--- v1 and v2 linearly dependent:
v1 = c(1,2,3,4)
v2 = c(2,4,6,8)
v3 = rnorm(4)
v4 = rnorm(4)
A <- matrix(c(v1,v2,v3,v4), nrow=4)
solve(A)

#--- dummy variable version:
v1 = c(1,1,1,1)
v2 = c(1,0,1,0)
v3 = c(0,1,0,1)
A <- matrix(c(v1,v2,v3), nrow=4)
#now try to calculate (A'A) inverse
solve(t(A)%*%A)
```

Listing 1: singularity.R

- Suppose that two of your variables are perfectly collinear. For example, you are using temperature in celsius and temperature in Fahrenheit. But C=aF, where a is a constant, and your software will complain (because it cannot calculate $X'X^{-1}$).
- Dummy variable trap. Suppose you want to measure the effect of being female on some variable y. So you write

$$\begin{aligned} y &= b_0 + b_1 \text{male} + b_2 \text{female} + e \\ &= \mathbf{Xb} + \mathbf{e} \qquad (2) \\ &= \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ \vdots & \ddots & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix} + e \qquad (3) \end{aligned}$$

But then quickly realize that $v_1 = v_2 + v_3$.