Chapter 2. Square matrices Lecture notes for MA1111

P. Karageorgis

pete@maths.tcd.ie

Definition 2.1 – Invertible matrices

An $n \times n$ matrix A is said to be invertible, if there is a matrix B such that $AB = I_n$ and $BA = I_n$. If such a matrix exists, then it is unique.

- The matrix B is called the inverse of A. It is denoted by $B = A^{-1}$.
- A square matrix is not invertible, if it has a row/column of zeros.

Theorem 2.2 – Inverse of a product

Suppose A_1, A_2, \ldots, A_k are invertible matrices of the same size. Then their product $A_1A_2\cdots A_k$ is also invertible and its inverse is

$$(A_1 A_2 \cdots A_k)^{-1} = A_k^{-1} \cdots A_2^{-1} A_1^{-1}.$$

Definition 2.3 – Elementary matrix

An elementary matrix is a matrix obtained from the identity matrix I_n using a single elementary row operation.

Theorem 2.4 – Elementary matrices are invertible

Every elementary matrix is invertible and its inverse is elementary.

Theorem 2.5 – Elementary row operations

Every elementary row operation corresponds to left multiplication by an elementary matrix. This matrix can be found explicitly by applying the same elementary row operation to the identity matrix.

Theorem 2.6 – Reduced row echelon form of square matrices

The reduced row echelon form of an $n \times n$ matrix is either the identity matrix I_n or else a matrix whose last row is zero. Moreover, an $n \times n$ matrix is invertible if and only if its reduced row echelon form is I_n .

- To find the inverse of an $n \times n$ matrix A, one may use row reduction of the augmented matrix $[A \ I_n]$. There are two possible scenarios.
- If row reduction leads to a matrix of the form $[I_n B]$, then the given matrix A is invertible with inverse B. Otherwise, A is not invertible.

Theorem 2.7 – One-sided inverse

Knowing that $AB = I_n$, one knows that $BA = I_n$ and vice versa.

Finding the inverse: Example

• We compute the inverse of the matrix

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

.

• Merge the given matrix with the identity matrix I_3 and then use row reduction on the resulting matrix to get

$$\begin{bmatrix} 1 & 2 & 2 & | & 1 & 0 & 0 \\ 1 & 1 & 2 & | & 0 & 1 & 0 \\ 1 & 2 & 1 & | & 0 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{1} & 0 & 0 & | & -3 & 2 & 2 \\ 0 & \mathbf{1} & 0 & | & 1 & -1 & 0 \\ 0 & 0 & \mathbf{1} & | & 1 & 0 & -1 \end{bmatrix}$$

• The inverse of A is the rightmost submatrix, namely

$$A^{-1} = \begin{bmatrix} -3 & 2 & 2\\ 1 & -1 & 0\\ 1 & 0 & -1 \end{bmatrix}$$

Invertibility conditions

Theorem 2.8 – Invertibility conditions

The following statements are equivalent for each $n \times n$ matrix A.

- **1** The system $Am{x}=m{y}$ has a unique solution for every $m{y}\in\mathbb{R}^n.$
- ② The system $Am{x}=m{y}$ has a unique solution for some $m{y}\in\mathbb{R}^n.$
- $\mathbf{\mathfrak{S}}$ The reduced row echelon form of A is the identity matrix I_n .
- **4** The matrix A is the product of elementary matrices.
- **5** The matrix A is invertible.

• The inverse of a 2×2 matrix A is given by the formula

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

provided that $ad - bc \neq 0$. The matrix A is not invertible, otherwise.

Definition of the determinant

• We are hoping to define the determinant of a square matrix in such a way that the following four properties hold.

(A1) Breaking up a row into two pieces gives

$$\det \begin{bmatrix} \vdots \\ R+S \\ \vdots \end{bmatrix} = \det \begin{bmatrix} \vdots \\ R \\ \vdots \end{bmatrix} + \det \begin{bmatrix} \vdots \\ S \\ \vdots \end{bmatrix}$$

(A2) Pulling out a common factor from a row gives

$$\det \begin{bmatrix} \vdots \\ cR \\ \vdots \end{bmatrix} = c \det \begin{bmatrix} \vdots \\ R \\ \vdots \end{bmatrix}.$$

(A3) One has det A = 0 when A has two equal rows. (A4) One has det $I_n = 1$ when I_n is the identity matrix.

Computation of determinants: first method

Theorem 2.9 – Determinants using row reduction

If there is a definition of the determinant so that (A1)-(A4) hold, then it is unique and one may compute determinants using row reduction. The determinant remains the same, if we add a multiple of one row to another, and it changes by a minus sign, if we interchange two rows.

• As a typical example, one can use row reduction to compute

$$\det \begin{bmatrix} 3 & 6 \\ 2 & 9 \end{bmatrix} = 3 \det \begin{bmatrix} 1 & 2 \\ 2 & 9 \end{bmatrix} = 3 \det \begin{bmatrix} 1 & 2 \\ 0 & 5 \end{bmatrix}$$
$$= 15 \det \begin{bmatrix} \mathbf{1} & 2 \\ 0 & 1 \end{bmatrix} = 15 \det \begin{bmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{1} \end{bmatrix} = 15.$$

• More generally, the determinant of a 2×2 matrix is given by

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

Computation of determinants: second method

Definition 2.10 – Expansion by minors

The minor M_{ij} of a matrix A is obtained by deleting both the *i*th row and the *j*th column of A. To define the determinant of a square matrix using induction, one defines det[a] = a for 1×1 matrices and then

$$\det A = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det M_{i1}$$

for all larger matrices. This defines $\det A$ for any square matrix A.

• For instance, one can use expansion by minors to compute

$$\det \begin{bmatrix} 2 & 4 & 3 \\ 3 & 5 & 1 \\ 0 & 2 & 3 \end{bmatrix} = 2 \det \begin{bmatrix} 5 & 1 \\ 2 & 3 \end{bmatrix} - 3 \det \begin{bmatrix} 4 & 3 \\ 2 & 3 \end{bmatrix}$$
$$= 2(15 - 2) - 3(12 - 6) = 26 - 18 = 8.$$

Lower and upper triangular matrices

Theorem 2.11 – Lower and upper triangular matrices

The determinant of a lower or upper triangular matrix is equal to the product of its diagonal entries.

- A square matrix is called lower/upper triangular, if the entries that lie above/below its main diagonal are zero.
- One can prove this theorem using either expansion by minors or row reduction. For instance, expansion by minors gives

$$\det \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11} \det \begin{bmatrix} a_{22} & 0 \\ a_{32} & a_{33} \end{bmatrix} = a_{11} a_{22} a_{33}.$$

 Thus, the theorem holds in the case of 3 × 3 lower triangular matrices. To prove the general case, one uses a similar argument and induction.

Definition 2.12 – Transpose of a matrix

The transpose of an $m \times n$ matrix A is the $n \times m$ matrix A^t obtained by turning rows into columns. One can easily check that

$$(A+B)^t = A^t + B^t, \qquad (A^t)^t = A, \qquad (AB)^t = B^t A^t.$$

Theorem 2.13 – Further properties of determinants

- **1** One has det(AB) = (det A)(det B) for all $n \times n$ matrices A, B.
- **2** To say that A is invertible is to say that $det A \neq 0$.
- **3** If A is invertible, then $\det A^{-1} = 1/(\det A)$.
- 4 If A^t is the transpose of A, then $\det A^t = \det A$.

Cofactor matrix and adjoint

Definition 2.14 – Cofactor matrix and adjoint

Let A be a square matrix and let M_{ij} be its minors. The cofactor C_{ij} is defined as the signed determinant of the minor M_{ij} , namely

$$C_{ij} = (-1)^{i+j} \cdot \det M_{ij}.$$

The adjoint of A is defined as the transpose of the cofactor matrix C.

 $\bullet\,$ When it comes to 2×2 matrices, for instance, one has

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \implies C = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix} \implies \operatorname{adj} A = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

• The identity $A \cdot \operatorname{adj} A = (\det A)I_n$ holds for every square matrix A. In particular, the inverse of an invertible matrix A is given by

$$A^{-1} = \frac{1}{\det A} \cdot \operatorname{adj} A.$$

Adjoint matrix: Example, page 1

• We compute both the adjoint and the inverse of the matrix

$$A = \begin{bmatrix} 1 & x & y \\ & 1 & z \\ & & 1 \end{bmatrix}.$$

• First, we compute the cofactor entries C_{ij} . There are 9 entries to be found. The cofactor entry C_{11} is given by

$$C_{11} = +\det \begin{bmatrix} \frac{1-x-y}{0} & 1 & z\\ 0 & 0 & 1 \end{bmatrix} = 1,$$

while the cofactor entry C_{12} is given by

$$C_{12} = -\det \begin{bmatrix} \frac{1-x-y}{0} & 1 & z \\ 0 & 0 & 1 \end{bmatrix} = 0.$$

The remaining cofactor entries may be computed similarly.

Adjoint matrix: Example, page 2

• Proceeding as above, one finds that the cofactor entries are

$$C_{11} = C_{22} = C_{33} = 1, \qquad C_{21} = -x,$$

$$C_{12} = C_{13} = C_{23} = 0, \qquad C_{32} = -z, \qquad C_{31} = xz - y.$$

Once we now merge those into a matrix, we get

$$C = \begin{bmatrix} 1 & 0 & 0 \\ -x & 1 & 0 \\ xz - y & -z & 1 \end{bmatrix} \implies \operatorname{adj} A = C^{t} = \begin{bmatrix} 1 & -x & xz - y \\ 0 & 1 & -z \\ 0 & 0 & 1 \end{bmatrix}.$$

• Since A is upper triangular, its determinant is equal to the product of its diagonal entries. This gives $\det A = 1$, so

$$A^{-1} = \frac{1}{\det A} \cdot \operatorname{adj} A = \begin{bmatrix} 1 & -x & xz - y \\ 0 & 1 & -z \\ 0 & 0 & 1 \end{bmatrix}$$

Theorem 2.15 – Complete expansion of the determinant

The determinant of an $n \times n$ matrix A can be expressed in the form

$$\det A = \sum \operatorname{sign} \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} a_{1i_1} a_{2i_2} \cdots a_{ni_n}$$

with the sum taken over all permutations i_1, i_2, \ldots, i_n of $1, 2, \ldots, n$.

- A permutation is called odd/even, if one needs an odd/even number of index swaps in order to get from (i_1, i_2, \ldots, i_n) to $(1, 2, \ldots, n)$.
- Even permutations have sign +1. Odd permutations have sign -1.
- The complete expansion of the determinant contains n! terms, so it is not particularly useful for large n. We shall mainly use it to compute determinants when n = 2, 3 and use row reduction when $n \ge 4$.

Determinant of 3×3 matrices

• The determinant of a 3×3 matrix A is given by the formula

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$$

 $-a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}.$

 To find the six terms on the right hand side, one repeats the first two columns of the matrix and then forms six diagonal products as in the figure below. Note that three of the products are associated with a plus sign, while the other three have a minus sign, instead.



Permutation matrices

- A permutation matrix is a square matrix that has a single 1 in each row and each column, all other entries being zero. Left multiplication by such a matrix results in a permutation of the rows.
- As a typical example, let us consider the equation

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ x_4 & y_4 \end{bmatrix} = \begin{bmatrix} x_3 & y_3 \\ x_4 & y_4 \\ x_2 & y_2 \\ x_1 & y_1 \end{bmatrix}$$

• We note that row 1 is mapped to row 3, row 2 is mapped to row 4, and so on. The corresponding permutation of the rows is thus

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}.$$

 In general, left multiplication by a permutation matrix P maps row i to row j whenever the (i, j)th entry of P is equal to 1.

Cycles and transpositions

Definition 2.16 – Cycles and transpositions

A permutation that cyclically permutes k indices is called a k-cycle. A permutation that interchanges two indices is called a transposition.

- A typical example of a 3-cycle is the permutation 1 → 3 → 4 → 1. It is very common to denote such a cycle by (134).
- Every permutation can be expressed in terms of cycles. For instance,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 1 & 6 & 3 & 7 & 4 & 2 \end{pmatrix} = (1572)(364).$$

Since a k-cycle may be obtained using k − 1 transpositions, its sign is given by (−1)^{k−1}. As for the permutation above, its sign is given by

$$sign(1572)(364) = (-1)^3(-1)^2 = -1.$$