C

Matrix Algebra: Determinants, Inverses, Eigenvalues

This Chapter discusses more specialized properties of matrices, such as determinants, eigenvalues and rank. These apply only to *square* matrices unless extension to rectangular matrices is explicitly stated.

§C.1 DETERMINANTS

The *determinant* of a *square* matrix $\mathbf{A} = [a_{ij}]$ is a number denoted by $|\mathbf{A}|$ or $\det(\mathbf{A})$, through which important properties such as singularity can be briefly characterized. This number is defined as the following function of the matrix elements:

$$|\mathbf{A}| = \pm \prod a_{1j_1}\, a_{2j_2}\, \ldots\, a_{nj_n}, \tag{C.1}$$

where the column indices $j_1, j_2, \ldots j_n$ are taken from the set $1, 2, \ldots, n$ with no repetitions allowed. The plus (minus) sign is taken if the permutation $(j_1\ j_2\ \ldots\ j_n)$ is even (odd).

EXAMPLE C.1

For a 2×2 matrix,

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \tag{C.2}$$

EXAMPLE C.2

For a 3×3 matrix,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}. \tag{C.3}$$

REMARK C.1

The concept of determinant is not applicable to rectangular matrices or to vectors. Thus the notation $|\mathbf{x}|$ for a vector $\mathbf{x}$ can be reserved for its magnitude (as in Appendix A) without risk of confusion.

REMARK C.2

Inasmuch as the product (C.1) contains $n!$ terms, the calculation of $|\mathbf{A}|$ from the definition is impractical for general matrices whose order exceeds 3 or 4. For example, if $n = 10$, the product (C.1) contains $10! = 3,628,800$ terms each involving 9 multiplications, so over 30 million floating-point operations would be required to evaluate $|\mathbf{A}|$ according to that definition. A more practical method based on matrix decomposition is described in Remark C.3.

§C.1.1 Some Properties of Determinants

Some useful rules associated with the calculus of determinants are listed next.

I. Rows and columns can be interchanged without affecting the value of a determinant. That is

$$|\mathbf{A}| = |\mathbf{A}^T|. \tag{C.4}$$

II. If two rows (or columns) are interchanged the sign of the determinant is changed. For example:

$$\begin{vmatrix} 3 & 4 \\ 1 & -2 \end{vmatrix} = - \begin{vmatrix} 1 & -2 \\ 3 & 4 \end{vmatrix}. \tag{C.5}$$

III. If a row (or column) is changed by adding to or subtracting from its elements the corresponding elements of any other row (or column) the determinant remains unaltered. For example:

$$\begin{vmatrix} 3 & 4 \\ 1 & -2 \end{vmatrix} = \begin{vmatrix} 3+1 & 4-2 \\ 1 & -2 \end{vmatrix} = \begin{vmatrix} 4 & 2 \\ 1 & -2 \end{vmatrix} = -10. \tag{C.6}$$

IV. If the elements in any row (or column) have a common factor α then the determinant equals the determinant of the corresponding matrix in which $\alpha = 1$, multiplied by α. For example:

$$\begin{vmatrix} 6 & 8 \\ 1 & -2 \end{vmatrix} = 2 \begin{vmatrix} 3 & 4 \\ 1 & -2 \end{vmatrix} = 2 \times (-10) = -20. \tag{C.7}$$

V. When at least one row (or column) of a matrix is a linear combination of the other rows (or columns) the determinant is zero. Conversely, if the determinant is zero, then at least one row and one column are linearly dependent on the other rows and columns, respectively. For example, consider

$$\begin{vmatrix} 3 & 2 & 1 \\ 1 & 2 & -1 \\ 2 & -1 & 3 \end{vmatrix}. \tag{C.8}$$

This determinant is zero because the first column is a linear combination of the second and third columns:

$$\text{column } 1 = \text{column } 2 + \text{column } 3 \tag{C.9}$$

Similarly there is a linear dependence between the rows which is given by the relation

$$\text{row } 1 = \tfrac{7}{8} \text{ row } 2 + \tfrac{4}{5} \text{ row } 3 \tag{C.10}$$

VI. The determinant of an upper triangular or lower triangular matrix is the product of the main diagonal entries. For example,

$$\begin{vmatrix} 3 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 4 \end{vmatrix} = 3 \times 2 \times 4 = 24. \tag{C.11}$$

This rule is easily verified from the definition (C.1) because all terms vanish except $j_1 = 1$, $j_2 = 2, \ldots\ j_n = n$, which is the product of the main diagonal entries. Diagonal matrices are a particular case of this rule.

VII. The determinant of the product of two square matrices is the product of the individual determinants:

$$|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|. \tag{C.12}$$

This rule can be generalized to any number of factors. One immediate application is to matrix powers: $|\mathbf{A}^2| = |\mathbf{A}||\mathbf{A}| = |\mathbf{A}|^2$, and more generally $|\mathbf{A}^n| = |\mathbf{A}|^n$ for integer n.

VIII. The determinant of the transpose of a matrix is the same as that of the original matrix:

$$|\mathbf{A}^T| = |\mathbf{A}|. \tag{C.13}$$

This rule can be directly verified from the definition of determinant.

REMARK C.3

Rules VI and VII are the key to the practical evaluation of determinants. Any square nonsingular matrix **A** (where the qualifier "nonsingular" is explained in §C.3) can be decomposed as the product of two triangular factors

$$\mathbf{A} = \mathbf{L}\mathbf{U}, \tag{C.14}$$

where **L** is unit lower triangular and **U** is upper triangular. This is called a LU triangularization, LU factorization or LU decomposition. It can be carried out in $O(n^3)$ floating point operations. According to rule VII:

$$|\mathbf{A}| = |\mathbf{L}|\,|\mathbf{U}|. \tag{C.15}$$

But according to rule VI, $|\mathbf{L}| = 1$ and $|\mathbf{U}| = u_{11}u_{22}\ldots u_{nn}$. The last operation requires only $O(n)$ operations. Thus the evaluation of $|\mathbf{A}|$ is dominated by the effort involved in computing the factorization (C.14). For $n = 10$, that effort is approximately $10^3 = 1000$ floating-point operations, compared to approximately 3×10^7 from the naive application of (C.1), as noted in Remark C.1. Thus the LU-based method is roughly 30, 000 times faster for that modest matrix order, and the ratio increases exponentially for large n.

§C.1.2 Cramer's Rule

Cramer's rule provides a recipe for solving linear algebraic equations in terms of determinants. Let the simultaneous equations be as usual denoted as

$$\mathbf{A}\mathbf{x} = \mathbf{y}, \tag{C.16}$$

where **A** is a given $n \times n$ matrix, **y** is a given $n \times 1$ vector, and **x** is the $n \times 1$ vector of unknowns. The explicit form of (C.16) is Equation (A.1) of Appendix A, with $n = m$.

The explicit solution for the components $x_1, x_2 \ldots, x_n$ of **x** in terms of determinants is

$$x_1 = \frac{\begin{vmatrix} y_1 & a_{12} & a_{13} & \ldots & a_{1n} \\ y_2 & a_{22} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \\ y_n & a_{n2} & a_{n3} & \ldots & a_{nn} \end{vmatrix}}{|\mathbf{A}|}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & y_1 & a_{13} & \ldots & a_{1n} \\ a_{21} & y_2 & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \\ a_{n1} & y_n & a_{n3} & \ldots & a_{nn} \end{vmatrix}}{|\mathbf{A}|}, \ldots \tag{C.17}$$

The rule can be remembered as follows: in the numerator of the quotient for x_j, replace the j^{th} column of **A** by the right-hand side **y**.

This method of solving simultaneous equations is known as *Cramer's rule*. Because the explicit computation of determinants is impractical for $n > 3$ as explained in Remark C.3, this rule has practical value only for $n = 2$ and $n = 3$ (it is marginal for $n = 4$). But such small-order systems arise often in finite element calculations at the *Gauss point level*; consequently implementors should be aware of this rule for such applications.

EXAMPLE C.3

Solve the 3×3 linear system

$$\begin{bmatrix} 5 & 2 & 1 \\ 3 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 5 \\ 3 \end{bmatrix}, \tag{C.18}$$

by Cramer's rule:

$$x_1 = \frac{\begin{vmatrix} 8 & 2 & 1 \\ 5 & 2 & 0 \\ 3 & 0 & 2 \end{vmatrix}}{\begin{vmatrix} 5 & 2 & 1 \\ 3 & 2 & 0 \\ 1 & 0 & 2 \end{vmatrix}} = \frac{6}{6} = 1, \quad x_2 = \frac{\begin{vmatrix} 5 & 8 & 1 \\ 3 & 5 & 0 \\ 1 & 3 & 2 \end{vmatrix}}{\begin{vmatrix} 5 & 2 & 1 \\ 3 & 2 & 0 \\ 1 & 0 & 2 \end{vmatrix}} = \frac{6}{6} = 1, \quad x_3 = \frac{\begin{vmatrix} 5 & 2 & 8 \\ 3 & 2 & 5 \\ 1 & 0 & 3 \end{vmatrix}}{\begin{vmatrix} 5 & 2 & 1 \\ 3 & 2 & 0 \\ 1 & 0 & 2 \end{vmatrix}} = \frac{6}{6} = 1. \tag{C.19}$$

EXAMPLE C.4

Solve the 2×2 linear algebraic system

$$\begin{bmatrix} 2+\beta & -\beta \\ -\beta & 1+\beta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix} \tag{C.20}$$

by Cramer's rule:

$$x_1 = \frac{\begin{vmatrix} 5 & -\beta \\ 0 & 1+\beta \end{vmatrix}}{\begin{vmatrix} 2+\beta & -\beta \\ -\beta & 1+\beta \end{vmatrix}} = \frac{5+5\beta}{2+3\beta}, \quad x_2 = \frac{\begin{vmatrix} 2+\beta & 5 \\ -\beta & 0 \end{vmatrix}}{\begin{vmatrix} 2+\beta & -\beta \\ -\beta & 1+\beta \end{vmatrix}} = \frac{5\beta}{2+3\beta}. \tag{C.21}$$

§C.1.3 Homogeneous Systems

One immediate consequence of Cramer's rule is what happens if

$$y_1 = y_2 = \ldots = y_n = 0. \tag{C.22}$$

The linear equation systems with a null right hand side

$$\mathbf{A}\mathbf{x} = \mathbf{0}, \tag{C.23}$$

is called a *homogeneous system*. From the rule (C.17) we see that if $|\mathbf{A}|$ is nonzero, all solution components are zero, and consequently the only possible solution is the trivial one $\mathbf{x} = \mathbf{0}$. The case in which $|\mathbf{A}|$ vanishes is discussed in the next section.

§C.2 SINGULAR MATRICES, RANK

If the determinant $|\mathbf{A}|$ of a $n \times n$ square matrix $\mathbf{A} \equiv \mathbf{A}_n$ is zero, then the matrix is said to be *singular*. This means that at least one row and one column are linearly dependent on the others. If this row and column are removed, we are left with another matrix, say $\mathbf{A}_{n-1}$, to which we can apply the same criterion. If the determinant $|\mathbf{A}_{n-1}|$ is zero, we can remove another row and column from it to get $\mathbf{A}_{n-2}$, and so on. Suppose that we eventually arrive at an $r \times r$ matrix $\mathbf{A}_r$ whose determinant is nonzero. Then matrix $\mathbf{A}$ is said to have *rank* r, and we write $\text{rank}(\mathbf{A}) = r$.

If the determinant of $\mathbf{A}$ is nonzero, then $\mathbf{A}$ is said to be *nonsingular*. The rank of a nonsingular $n \times n$ matrix is equal to n.

Obviously the rank of $\mathbf{A}^T$ is the same as that of $\mathbf{A}$ since it is only necessary to transpose "row" and "column" in the definition.

The notion of rank can be extended to rectangular matrices as outlined in section §C.2.4 below. That extension, however, is not important for the material covered here.

EXAMPLE C.5

The 3×3 matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & 2 \\ 1 & 2 & -1 \\ 2 & -1 & 3 \end{bmatrix}, \tag{C.24}$$

has rank $r = 3$ because $|\mathbf{A}| = -3 \neq 0$.

EXAMPLE C.6

The matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 2 & -1 \\ 2 & -1 & 3 \end{bmatrix}, \tag{C.25}$$

already used as an example in §C.1.1 is singular because its first row and column may be expressed as linear combinations of the others through the relations (C.9) and (C.10). Removing the first row and column we are left with a 2×2 matrix whose determinant is $2 \times 3 - (-1) \times (-1) = 5 \neq 0$. Consequently (C.25) has rank $r = 2$.

§C.2.1 Rank Deficiency

If the square matrix **A** is supposed to be of rank r but in fact has a smaller rank $\bar{r} < r$, the matrix is said to be *rank deficient*. The number $r - \bar{r} > 0$ is called the *rank deficiency*.

EXAMPLE C.7

Suppose that the *unconstrained* master stiffness matrix **K** of a finite element has order n, and that the element possesses b independent rigid body modes. Then the expected rank of **K** is $r = n - b$. If the actual rank is less than r, the finite element model is said to be rank-deficient. This is an undesirable property.

EXAMPLE C.8

An an illustration of the foregoing rule, consider the two-node, 4-DOF plane beam element stiffness derived in Chapter 13:

$$\mathbf{K} = \frac{EI}{L^3} \begin{bmatrix} 12 & 6L & -12 & 6L \\ & 4L^2 & -6L & 2L^2 \\ & & 12 & -6L \\ symm & & & 4L^2 \end{bmatrix} \tag{C.26}$$

where EI and L are nonzero scalars. It can be verified that this 4×4 matrix has rank 2. The number of rigid body modes is 2, and the expected rank is $r = 4 - 2 = 2$. Consequently this model is rank sufficient.

§C.2.2 Rank of Matrix Sums and Products

In finite element analysis matrices are often built through sum and product combinations of simpler matrices. Two important rules apply to “rank propagation” through those combinations.

The rank of the product of two square matrices **A** and **B** cannot exceed the smallest rank of the multiplicand matrices. That is, if the rank of **A** is r_a and the rank of **B** is r_b,

$$Rank(\mathbf{AB}) \leq \min(r_a, r_b). \tag{C.27}$$

Regarding sums: the rank of a matrix sum cannot exceed the sum of ranks of the summand matrices. That is, if the rank of **A** is r_a and the rank of **B** is r_b,

$$Rank(\mathbf{A} + \mathbf{B}) \leq r_a + r_b. \tag{C.28}$$

§C.2.3 Singular Systems: Particular and Homegeneous Solutions

Having introduced the notion of rank we can now discuss what happens to the linear system (C.16) when the determinant of $\mathbf{A}$ vanishes, meaning that its rank is less than n. If so, the linear system (C.16) has either no solution or an infinite number of solution. Cramer's rule is of limited or no help in this situation.

To discuss this case further we note that if $|\mathbf{A}| = 0$ and the rank of $\mathbf{A}$ is $r = n - d$, where $d \geq 1$ is the *rank deficiency*, then there exist d nonzero independent vectors $\mathbf{z}_i$, $i = 1, \ldots d$ such that

$$\mathbf{A}\mathbf{z}_i = \mathbf{0}. \tag{C.29}$$

These d vectors, suitably orthonormalized, are called *null eigenvectors* of $\mathbf{A}$, and form a basis for its *null space*.

Let $\mathbf{Z}$ denote the $n \times d$ matrix obtained by collecting the $\mathbf{z}_i$ as columns. If $\mathbf{y}$ in (C.13) is in the *range* of $\mathbf{A}$, that is, there exists an nonzero $\mathbf{x}_p$ such that $\mathbf{y} = \mathbf{A}\mathbf{x}_p$, its general solution is

$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_h = \mathbf{x}_p + \mathbf{Z}\mathbf{w}, \tag{C.30}$$

where $\mathbf{w}$ is an arbitrary $d \times 1$ weighting vector. This statement can be easily verified by substituting this solution into $\mathbf{A}\mathbf{x} = \mathbf{y}$ and noting that $\mathbf{A}\mathbf{Z}$ vanishes.

The components $\mathbf{x}_p$ and $\mathbf{x}_h$ are called the *particular* and *homogeneous* part, respectively, of the solution $\mathbf{x}$. If $\mathbf{y} = \mathbf{0}$ only the homogeneous part remains.

If $\mathbf{y}$ is not in the range of $\mathbf{A}$, system (C.13) does not generally have a solution in the conventional sense, although least-square solutions can usually be constructed. The reader is referred to the many textbooks in linear algebra for further details.

§C.2.4 Rank of Rectangular Matrices

The notion of rank can be extended to rectangular matrices, real or complex, as follows. Let $\mathbf{A}$ be $m \times n$. Its *column range space* $\mathcal{R}(\mathbf{A})$ is the subspace spanned by $\mathbf{A}\mathbf{x}$ where $\mathbf{x}$ is the set of all complex n-vectors. Mathematically: $\mathcal{R}(\mathbf{A}) = \{\mathbf{A}\mathbf{x} : \mathbf{x} \in C^n\}$. The rank r of $\mathbf{A}$ is the dimension of $\mathcal{R}(\mathbf{A})$.

The *null space* $\mathcal{N}(\mathbf{A})$ of $\mathbf{A}$ is the set of n-vectors $\mathbf{z}$ such that $\mathbf{A}\mathbf{z} = \mathbf{0}$. The dimension of $\mathcal{N}(\mathbf{A})$ is $n - r$.

Using these definitions, the product and sum rules (C.27) and (C.28) generalize to the case of rectangular (but conforming) $\mathbf{A}$ and $\mathbf{B}$. So does the treatment of linear equation systems $\mathbf{A}\mathbf{x} = \mathbf{y}$ in which $\mathbf{A}$ is rectangular; such systems often arise in the fitting of observation and measurement data.

In finite element methods, rectangular matrices appear in change of basis through congruential transformations, and in the treatment of multifreedom constraints.

§C.3 MATRIX INVERSION

The *inverse* of a square nonsingular matrix $\mathbf{A}$ is represented by the symbol $\mathbf{A}^{-1}$ and is defined by the relation

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}. \tag{C.31}$$

The most important application of the concept of inverse is the solution of linear systems. Suppose that, in the usual notation, we have

$$\mathbf{A}\mathbf{x} = \mathbf{y} \tag{C.32}$$

Premultiplying both sides by $\mathbf{A}^{-1}$ we get the inverse relationship

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y} \tag{C.33}$$

More generally, consider the matrix equation for multiple (m) right-hand sides:

$$\underset{n\times n}{\mathbf{A}}\;\underset{n\times m}{\mathbf{X}} = \underset{n\times m}{\mathbf{Y}}, \tag{C.34}$$

which reduces to (C.32) for $m = 1$. The inverse relation that gives $\mathbf{X}$ as function of $\mathbf{Y}$ is

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{Y}. \tag{C.35}$$

In particular, the solution of

$$\mathbf{A}\mathbf{X} = \mathbf{I}, \tag{C.36}$$

is $\mathbf{X} = \mathbf{A}^{-1}$. Practical methods for computing inverses are based on directly solving this equation; see Remark C.4.

§C.3.1 Explicit Computation of Inverses

The explicit calculation of matrix inverses is seldom needed in large matrix computations. But occasionally the need arises for the explicit inverse of small matrices that appear in element computations. For example, the inversion of Jacobian matrices at Gauss points, or of constitutive matrices.

A general formula for elements of the inverse can be obtained by specializing Cramer's rule to (C.36). Let $\mathbf{B} = [b_{ij}] = \mathbf{A}^{-1}$. Then

$$b_{ij} = \frac{A_{ji}}{|\mathbf{A}|}, \tag{C.37}$$

in which A_{ji} denotes the so-called *adjoint* of entry a_{ij} of $\mathbf{A}$. The adjoint A_{ji} is defined as the determinant of the submatrix of order $(n-1)\times(n-1)$ obtained by deleting the j^{th} row and i^{th} column of $\mathbf{A}$, multiplied by $(-1)^{i+j}$.

This direct inversion procedure is useful only for small matrix orders: 2 or 3. In the examples below the inversion formulas for second and third order matrices are listed.

EXAMPLE C.9

For order $n = 2$:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \qquad \mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{22} \end{bmatrix}, \tag{C.38}$$

in which $|\mathbf{A}|$ is given by (C.2).

EXAMPLE C.10

For order $n = 3$:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \qquad \mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}, \tag{C.39}$$

where

$$\begin{aligned} b_{11} &= \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, & b_{21} &= -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}, & b_{31} &= \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}, \\ b_{12} &= -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}, & b_{22} &= \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}, & b_{32} &= -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}, \\ b_{13} &= \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}, & b_{23} &= -\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}, & b_{33} &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \end{aligned} \tag{C.40}$$

in which $|\mathbf{A}|$ is given by (C.3).

EXAMPLE C.11

$$\mathbf{A} = \begin{bmatrix} 2 & 4 & 2 \\ 3 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \qquad \mathbf{A}^{-1} = -\frac{1}{8} \begin{bmatrix} 1 & -4 & 2 \\ -2 & 0 & 4 \\ -1 & 4 & -10 \end{bmatrix}. \tag{C.41}$$

If the order exceeds 3, the general inversion formula based on Cramer's rule becomes rapidly useless as it displays combinatorial complexity. For numerical work it is preferable to solve the system (C.38) after $\mathbf{A}$ is factored. Those techniques are described in detail in linear algebra books; see also Remark C.4.

§C.3.2 Some Properties of the Inverse

I. The inverse of the transpose is equal to the transpose of the inverse:

$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T, \tag{C.42}$$

because

$$(\mathbf{A}\mathbf{A}^{-1}) = (\mathbf{A}\mathbf{A}^{-1})^T = (\mathbf{A}^{-1})^T\,\mathbf{A}^T = \mathbf{I}. \tag{C.43}$$

II. The inverse of a symmetric matrix is also symmetric. Because of the previous rule, $(\mathbf{A}^T)^{-1} = \mathbf{A}^{-1} = (\mathbf{A}^{-1})^T$, hence $\mathbf{A}^{-1}$ is also symmetric.

III. The inverse of a matrix product is the reverse product of the inverses of the factors:

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}. \tag{C.44}$$

This is easily verified by substituting both sides of (C.39) into (C.31). This property generalizes to an arbitrary number of factors.

IV. For a diagonal matrix $\mathbf{D}$ in which all diagonal entries are nonzero, $\mathbf{D}^{-1}$ is again a diagonal matrix with entries $1/d_{ii}$. The verification is straightforward.

V. If $\mathbf{S}$ is a block diagonal matrix:

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{22} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{S}_{33} & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{S}_{nn} \end{bmatrix} = \mathrm{diag}\,[\,\mathbf{S}_{ii}\,], \tag{C.45}$$

then the inverse matrix is also block diagonal and is given by

$$\mathbf{S}^{-1} = \begin{bmatrix} \mathbf{S}_{11}^{-1} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{22}^{-1} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{S}_{33}^{-1} & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{S}_{nn}^{-1} \end{bmatrix} = \mathrm{diag}\,[\,\mathbf{S}_{ii}^{-1}\,]. \tag{C.46}$$

VI. The inverse of an upper triangular matrix is also an upper triangular matrix. The inverse of a lower triangular matrix is also a lower triangular matrix. Both inverses can be computed in $O(n^2)$ floating-point operations.

REMARK C.4

The practical numerical calculation of inverses is based on triangular factorization. Given a nonsingular $n \times n$ matrix $\mathbf{A}$, calculate its LU factorization $\mathbf{A} = \mathbf{LU}$, which can be obtained in $O(n^3)$ operations. Then solve the linear triangular systems:

$$\mathbf{UY} = \mathbf{I}, \qquad \mathbf{LX} = \mathbf{Y}, \tag{C.47}$$

and the computed inverse $\mathbf{A}^{-1}$ appears in $\mathbf{X}$. One can overwrite $\mathbf{I}$ with $\mathbf{Y}$ and $\mathbf{Y}$ with $\mathbf{X}$. The whole process can be completed in $O(n^3)$ floating-point operations. For symmetric matrices the alternative decomposition $\mathbf{A} = \mathbf{LDL}^T$, where $\mathbf{L}$ is unit lower triangular and $\mathbf{D}$ is diagonal, is generally preferred to save computing time and storage.

§C.4 EIGENVALUES AND EIGENVECTORS

Consider the special form of the linear system (C.13) in which the right-hand side vector $\mathbf{y}$ is a multiple of the solution vector $\mathbf{x}$:

$$\mathbf{Ax} = \lambda\mathbf{x}, \tag{C.48}$$

or, written in full,

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & \lambda x_1 \\ a_{21}x_2 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & \lambda x_2 \\ \cdots & & \cdots & & \cdots & & \cdots & & \cdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n & = & \lambda x_n \end{array} \tag{C.49}$$

This is called the standard (or classical) *algebraic eigenproblem*. System (C.48) can be rearranged into the homogeneous form

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}. \tag{C.50}$$

A nontrivial solution of this equation is possible if and only if the coefficient matrix $\mathbf{A} - \lambda\mathbf{I}$ is singular. Such a condition can be expressed as the vanishing of the determinant

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0. \tag{C.51}$$

When this determinant is expanded, we obtain an algebraic polynomial equation in λ of degree n:

$$P(\lambda) = \lambda^n + \alpha_1\lambda^{n-1} + \cdots + \alpha_n = 0. \tag{C.52}$$

This is known as the *characteristic equation* of the matrix $\mathbf{A}$. The left-hand side is called the *characteristic polynomial*. We known that a polynomial of degree n has n (generally complex) roots λ_1, λ_2, ..., λ_n. These n numbers are called the *eigenvalues*, *eigenroots* or *characteristic values* of matrix $\mathbf{A}$.

With each eigenvalue λ_i there is an associated vector $\mathbf{x}_i$ that satisfies

$$\mathbf{A}\mathbf{x}_i = \lambda_i \mathbf{x}_i. \tag{C.53}$$

This $\mathbf{x}_i$ is called an *eigenvector* or *characteristic vector*. An eigenvector is unique only up to a scale factor since if $\mathbf{x}_i$ is an eigenvector, so is $\beta\mathbf{x}_i$ where β is an arbitrary nonzero number. Eigenvectors are often *normalized* so that their Euclidean length is 1, or their largest component is unity.

§C.4.1 Real Symmetric Matrices

Real symmetric matrices are of special importance in the finite element method. In linear algebra books dealing with the algebraic eigenproblem it is shown that:

(a) The n eigenvalues of a real symmetric matrix of order n are real.

(b) The eigenvectors corresponding to distinct eigenvalues are orthogonal. The eigenvectors corresponding to multiple roots may be orthogonalized with respect to each other.

(c) The n eigenvectors form a complete orthonormal basis for the Euclidean space E_n.

§C.4.2 Positivity

Let $\mathbf{A}$ be an $n \times n$ square *symmetric* matrix. $\mathbf{A}$ is said to be *positive definite* (p.d.) if

$$\mathbf{x}^T\mathbf{A}\mathbf{x} > 0, \qquad \mathbf{x} \neq 0 \tag{C.54}$$

A positive definite matrix has rank n. This property can be checked by computing the n eigenvalues λ_i of $\mathbf{A}\mathbf{z} = \lambda\mathbf{z}$. If all $\lambda_i > 0$, $\mathbf{A}$ is p.d.

$\mathbf{A}$ is said to be *nonnegative*[1] if zero equality is allowed in (C.54):

$$\mathbf{x}^T\mathbf{A}\mathbf{x} \leq 0, \qquad \mathbf{x} \neq 0 \tag{C.55}$$

A p.d. matrix is also nonnegative but the converse is not necessarily true. This property can be checked by computing the n eigenvalues λ_i of $\mathbf{A}\mathbf{z} = \lambda\mathbf{z}$. If r eigenvalues $\lambda_i > 0$ and $n - r$ eigenvalues are zero, $\mathbf{A}$ is nonnegative with rank r.

A symmetric square matrix $\mathbf{A}$ that has at least one negative eigenvalue is called *indefinite*.

[1] A property called *positive semi-definite* by some authors.

§C.4.3 Normal and Orthogonal Matrices

Let $\mathbf{A}$ be an $n \times n$ real square matrix. This matrix is called *normal* if

$$\mathbf{A}^T\mathbf{A} = \mathbf{A}\mathbf{A}^T \tag{C.56}$$

A normal matrix is called *orthogonal* if

$$\mathbf{A}^T\mathbf{A} = \mathbf{A}\mathbf{A}^T = \mathbf{I} \qquad \text{or} \quad \mathbf{A}^T = \mathbf{A}^{-1} \tag{C.57}$$

All eigenvalues of an orthogonal matrix have modulus one, and the matrix has rank n.

The generalization of the orthogonality property to complex matrices, for which transposition is replaced by conjugation, leads to *unitary* matrices. These are not required, however, for the material covered in the text.

§C.5 THE SHERMAN-MORRISON AND RELATED FORMULAS

The Sherman-Morrison formula gives the inverse of a matrix modified by a rank-one matrix. The Woodbury formula extends the Sherman-Morrison formula to a modification of arbitrary rank. In structural analysis these formulas are of interest for problems of *structural modifications*, in which a finite-element (or, in general, a discrete model) is changed by an amount expressable as a low-rank correction to the original model.

§C.5.1 The Sherman-Morrison Formula

Let $\mathbf{A}$ be a square $n \times n$ invertible matrix, whereas $\mathbf{u}$ and $\mathbf{v}$ are two n-vectors and β an arbitrary scalar. Assume that $\sigma = 1 + \beta\,\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u} \neq 0$. Then

$$\left(\mathbf{A} + \beta\mathbf{u}\mathbf{v}^T\right)^{-1} = \mathbf{A}^{-1} - \frac{\beta}{\sigma}\,\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}. \tag{C.58}$$

This is called the Sherman-Morrison formula[2] when $\beta = 1$. Since any rank-one correction to $\mathbf{A}$ can be written as $\beta\mathbf{u}\mathbf{v}^T$, (C.58) gives the rank-one change to its inverse. The proof is by direct multiplication as in Exercise C.5.

For practical computation of the change one solves the linear systems $\mathbf{A}\mathbf{a} = \mathbf{u}$ and $\mathbf{A}\mathbf{b} = \mathbf{v}$ for $\mathbf{a}$ and $\mathbf{b}$, using the known $\mathbf{A}^{-1}$. Compute $\sigma = 1 + \beta\mathbf{v}^T\mathbf{a}$. If $\sigma \neq 0$, the change to $\mathbf{A}^{-1}$ is the dyadic $-(\beta/\sigma)\mathbf{a}\mathbf{b}^T$.

§C.5.2 The Woodbury Formula

Let again $\mathbf{A}$ be a square $n \times n$ invertible matrix, whereas $\mathbf{U}$ and $\mathbf{V}$ are two $n \times k$ matrices with $k \leq n$ and β an arbitrary scalar. Assume that the $k \times k$ matrix $\boldsymbol{\Sigma} = \mathbf{I}_k + \beta\mathbf{V}^T\mathbf{A}^{-1}\mathbf{U}$, in which $\mathbf{I}_k$ denotes the $k \times k$ identity matrix, is invertible. Then

$$\left(\mathbf{A} + \beta\mathbf{U}\mathbf{V}^T\right)^{-1} = \mathbf{A}^{-1} - \beta\,\mathbf{A}^{-1}\mathbf{U}\boldsymbol{\Sigma}^{-1}\mathbf{V}^T\mathbf{A}^{-1}. \tag{C.59}$$

This is called the Woodbury formula.[3] It reduces to (C.58) if $k = 1$, in which case $\boldsymbol{\Sigma} \equiv \sigma$ is a scalar. The proof is by direct multiplication.

---

[2] J. Sherman and W. J. Morrison, Adjustment of an inverse matrix corresponding to changes in the elements of a given column or a given row of the original matrix, *Ann. Math. Statist.*, **20**, (1949), 621. For a history of this remarkable formula and its extensions, which are quite important in many applications such as statistics, see the paper by Henderson and Searle in *SIAM Review*, **23**, 53–60.

[3] M.A.Woodbury, Inverting modified matrices, *Statist. Res. Group, Mem. Rep.* No. 42, Princeton Univ., Princeton, N.J., 1950

§C.5.3 Formulas for Modified Determinants

Let $\widetilde{\mathbf{A}}$ denote the adjoint of $\mathbf{A}$. Taking the determinants from both sides of $\mathbf{A} + \beta \mathbf{u}\mathbf{v}^T$ one obtains

$$|\mathbf{A} + \beta \mathbf{u}\mathbf{v}^T| = |\mathbf{A}| + \beta\, \mathbf{v}^T\, \widetilde{\mathbf{A}}\, \mathbf{u}. \tag{C.60}$$

If $\mathbf{A}$ is invertible, replacing $\widetilde{\mathbf{A}} = |\mathbf{A}|\, \mathbf{A}^{-1}$ this becomes

$$|\mathbf{A} + \beta \mathbf{u}\mathbf{v}^T| = |\mathbf{A}|\, (1 + \beta\, \mathbf{v}^T \mathbf{A}^{-1}\, \mathbf{u}). \tag{C.61}$$

Similarly, one can show that if $\mathbf{A}$ is invertible, and $\mathbf{U}$ and $\mathbf{V}$ are $n \times k$ matrices,

$$|\mathbf{A} + \beta \mathbf{U}\mathbf{V}^T| = |\mathbf{A}|\;\; |\mathbf{I}_k + \beta\, \mathbf{V}^T \mathbf{A}^{-1} \mathbf{U}|. \tag{C.62}$$

Exercises for Appendix C: Determinants, Inverses, Eigenvalues

EXERCISE C.1

If $\mathbf{A}$ is a square matrix of order n and c a scalar, show that $\det(c\mathbf{A}) = c^n \det \mathbf{A}$.

EXERCISE C.2

Let $\mathbf{u}$ and $\mathbf{v}$ denote real n-vectors normalized to unit length, so that $\mathbf{u}^T = \mathbf{u} = 1$ and $\mathbf{v}^T\mathbf{v} = 1$, and let $\mathbf{I}$ denote the $n \times n$ identity matrix. Show that

$$\det(\mathbf{I} - \mathbf{u}\mathbf{v}^T) = 1 - \mathbf{v}^T\mathbf{u} \tag{EC.1}$$

EXERCISE C.3

Let $\mathbf{u}$ denote a real n-vector normalized to unit length, so that $\mathbf{u}^T = \mathbf{u} = 1$ and $\mathbf{I}$ denote the $n \times n$ identity matrix. Show that

$$\mathbf{H} = \mathbf{I} - 2\mathbf{u}\mathbf{u}^T \tag{EC.2}$$

is orthogonal: $\mathbf{H}^T\mathbf{H} = \mathbf{I}$, and idempotent: $\mathbf{H}^2 = \mathbf{H}$. This matrix is called a *elementary Hermitian*, a *Householder matrix*, or a *reflector*. It is a fundamental ingredient of many linear algebra algorithms; for example the QR algorithm for finding eigenvalues.

EXERCISE C.4

The *trace* of a $n \times n$ square matrix $\mathbf{A}$, denoted $\textbf{trace}(\mathbf{A})$ is the sum $\sum_{i=1}^{n} a_{ii}$ of its diagonal coefficients. Show that if the entries of $\mathbf{A}$ are real,

$$\textbf{trace}(\mathbf{A}^T\mathbf{A}) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2 \tag{EC.3}$$

EXERCISE C.5

Prove the Sherman-Morrison formula (C.59) by direct matrix multiplication.

EXERCISE C.6

Prove the Sherman-Morrison formula (C.59) for $\beta = 1$ by considering the following block bordered system

$$\begin{bmatrix} \mathbf{A} & \mathbf{U} \\ \mathbf{V}^T & \mathbf{I}_k \end{bmatrix} \begin{bmatrix} \mathbf{B} \\ \mathbf{C} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_n \\ \mathbf{0} \end{bmatrix} \tag{EC.4}$$

in which $\mathbf{I}_k$ and $\mathbf{I}_n$ denote the identy matrices of orders k and n, respectively. Solve (C.62) two ways: eliminating first $\mathbf{B}$ and then $\mathbf{C}$, and eliminating first $\mathbf{C}$ and then $\mathbf{B}$. Equate the results for $\mathbf{B}$.

EXERCISE C.7

Show that the eigenvalues of a real symmetric square matrix are real, and that the eigenvectors are real vectors.

EXERCISE C.8

Let the n real eigenvalues λ_i of a real $n \times n$ symmetric matrix $\mathbf{A}$ be classified into two subsets: r eigenvalues are nonzero whereas $n - r$ are zero. Show that $\mathbf{A}$ has rank r.

EXERCISE C.9

Show that if $\mathbf{A}$ is p.d., $\mathbf{A}\mathbf{x} = \mathbf{0}$ implies that $\mathbf{x} = \mathbf{0}$.

EXERCISE C.10

Show that for any real $m \times n$ matrix $\mathbf{A}$, $\mathbf{A}^T\mathbf{A}$ exists and is nonnegative.

EXERCISE C.11

Show that a triangular matrix is normal if and only if it is diagonal.

EXERCISE C.12

Let $\mathbf{A}$ be a real orthogonal matrix. Show that all of its eigenvalues λ_i, which are generally complex, have unit modulus.

EXERCISE C.13

Let $\mathbf{A}$ and $\mathbf{T}$ be real $n \times n$ matrices, with $\mathbf{T}$ nonsingular. Show that $\mathbf{T}^{-1}\mathbf{AT}$ and $\mathbf{A}$ have the same eigenvalues. (This is called a similarity transformation in linear algebra).

EXERCISE C.14

(Tough) Let $\mathbf{A}$ be $m \times n$ and $\mathbf{B}$ be $n \times m$. Show that the nonzero eigenvalues of $\mathbf{AB}$ are the same as those of $\mathbf{BA}$ (Kahan).

EXERCISE C.15

Let $\mathbf{A}$ be real skew-symmetric, that is, $\mathbf{A} = -\mathbf{A}^T$. Show that all eigenvalues of $\mathbf{A}$ are purely imaginary or zero.

EXERCISE C.16

Let $\mathbf{A}$ be real skew-symmetric, that is, $\mathbf{A} = -\mathbf{A}^T$. Show that $\mathbf{U} = (\mathbf{I}+\mathbf{A})^{-1}(\mathbf{I}-\mathbf{A})$, called a Cayley transformation, is orthogonal.

EXERCISE C.17

Let $\mathbf{P}$ be a real square matrix that satisfies

$$\mathbf{P}^2 = \mathbf{P}. \tag{EC.5}$$

Such matrices are called *idempotent*, and also *orthogonal projectors*. Show that all eigenvalues of $\mathbf{P}$ are either zero or one.

EXERCISE C.18

The necessary and sufficient condition for two square matrices to commute is that they have the same eigenvectors.

EXERCISE C.19

A matrix whose elements are equal on any line parallel to the main diagonal is called a Toeplitz matrix. (They arise in finite difference or finite element discretizations of regular one-dimensional grids.) Show that if $\mathbf{T}_1$ and $\mathbf{T}_2$ are any two Toeplitz matrices, they commute: $\mathbf{T}_1\mathbf{T}_2 = \mathbf{T}_2\mathbf{T}_1$. Hint: do a Fourier transform to show that the eigenvectors of any Toeplitz matrix are of the form $\{e^{i\omega nh}\}$; then apply the previous Exercise.