Online Appendix A

Introduction to Matrix Computations

A.1 Vectors and Matrices

A.1.1 Linear Vector Spaces

In this appendix we recall basic elements of finite-dimensional linear vector spaces and related matrix algebra, and introduce some notations to be used in the book. The exposition is brief and meant as a convenient reference.

We will be concerned with the vector spaces $\mathbf{R}^n$ and $\mathbf{C}^n$, that is, the set of real or complex n-tuples with $1 \leq n < \infty$. Let $v_1, v_2, \ldots, v_k$ be vectors and $\alpha_1, \alpha_2, \ldots, \alpha_k$ be scalars. The vectors are said to be **linearly independent** if none of them is a linear combination of the others, that is,

$$\sum_{i=1}^{k} \alpha_i v_i = 0 \;\Rightarrow\; \alpha_i = 0, \quad i = 1 : k.$$

Otherwise, if a nontrivial linear combination of $v_1, \ldots, v_k$ is zero, the vectors are said to be linearly dependent. Then at least one vector v_i will be a linear combination of the rest.

A **basis** in a vector space $\mathcal{V}$ is a set of linearly independent vectors $v_1, v_2, \ldots, v_n \in \mathcal{V}$ such that all vectors $v \in \mathcal{V}$ can be expressed as a linear combination:

$$v = \sum_{i=1}^{n} \xi_i v_i.$$

The scalars ξ_i are called the components or coordinates of v with respect to the basis $\{v_i\}$. If the vector space $\mathcal{V}$ has a basis of n vectors, then every system of linearly independent vectors of $\mathcal{V}$ has at most k elements and any other basis of $\mathcal{V}$ has the same number k of elements. The number k is called the **dimension** of $\mathcal{V}$ and denoted by $\dim(\mathcal{V})$.

The linear space of column vectors $x = (x_1, x_2, \ldots, x_n)^T$, where $x_i \in \mathbf{R}$ is denoted $\mathbf{R}^n$; if $x_i \in \mathbf{C}$, then it is denoted $\mathbf{C}^n$. The dimension of this space is n, and the unit vectors $e_1, e_2, \ldots, e_n$, where

$$e_1 = (1, 0, \ldots, 0)^T, \qquad e_2 = (0, 1, \ldots, 0)^T, \ldots, e_n = (0, 0, \ldots, 1)^T,$$

constitute the **standard basis**. Note that the components $x_1, x_2, \ldots, x_n$ are the coordinates when the vector x is expressed as a linear combination of the standard basis. We shall use the same name for a vector as for its coordinate representation by a column vector with respect to the standard basis.

An arbitrary basis can be characterized by the *nonsingular* matrix $V = (v_1, v_2, \ldots, v_n)$ composed of the basis vectors. The coordinate transformation reads $x = V\xi$. The standard basis itself is characterized by the **unit matrix**

$$I = (e_1, e_2, \ldots, e_n).$$

If $\mathcal{W} \subset \mathcal{V}$ is a vector space, then $\mathcal{W}$ is called a **vector subspace** of $\mathcal{V}$. The set of all linear combinations of $v_1, \ldots, v_k \in \mathcal{V}$ form a vector subspace denoted by

$$\text{span}\{v_1, \ldots, v_k\} = \sum_{i=1}^{k} \alpha_i v_i, \quad i = 1 : k,$$

where α_i are real or complex scalars. If $\mathcal{S}_1, \ldots, \mathcal{S}_k$ are vector subspaces of $\mathcal{V}$, then their sum defined by

$$S = \{v_1 + \cdots + v_k \mid v_i \in \mathcal{S}_i, \ i = 1 : k\}$$

is also a vector subspace. The intersection T of a set of vector subspaces is also a subspace,

$$T = \mathcal{S}_1 \cap \mathcal{S}_2 \cdots \cap \mathcal{S}_k.$$

(The union of vector spaces is generally not a vector space.) If the intersections of the subspaces are empty, $\mathcal{S}_i \cap \mathcal{S}_j = 0$, $i \neq j$, then the sum of the subspaces is called their **direct sum** and denoted by

$$S = \mathcal{S}_1 \oplus \mathcal{S}_2 \cdots \oplus \mathcal{S}_k.$$

A function F from one linear space to another (or the same) linear space is said to be **linear** if

$$F(\alpha u + \beta v) = \alpha F(u) + \beta F(v)$$

for all vectors $u, v \in V$ and all scalars α, β. Note that this terminology excludes nonhomogeneous functions like $\alpha u + \beta$, which are called **affine** functions. Linear functions are often expressed in the form Au, where A is called a **linear operator**.

A vector space for which an inner product is defined is called an **inner product space**. For the vector space $\mathbf{R}^n$ the **Euclidean inner product** is

$$(x, y) = \sum_{i=1}^{n} x_i y_i. \tag{A.1.1}$$

Similarly $\mathbf{C}^n$ is an inner product space with the inner product

$$(x, y) = \sum_{k=1}^{n} \bar{x}_k y_k, \tag{A.1.2}$$

where $\bar{x}_k$ denotes the complex conjugate of x_k.

Two vectors v and w in $\mathbf{R}^n$ are said to be **orthogonal** if $(v, w) = 0$. A set of vectors $v_1, \ldots, v_k$ in $\mathbf{R}^n$ is called orthogonal with respect to the Euclidean inner product if

$$(v_i, v_j) = 0, \quad i \neq j,$$

and **orthonormal** if also $(v_i, v_i) = 1$, $i = 1 : k$. An orthogonal set of vectors is linearly independent.

The **orthogonal complement** $S^\perp$ of a subspace $S \in \mathbf{R}^n$ is the subspace defined by

$$S^\perp = \{y \in \mathbf{R}^n \mid (y, x) = 0, \ x \in S\}.$$

More generally, the subspaces $S_1, \ldots, S_k$ of $\mathbf{R}^n$ are mutually orthogonal if, for all $1 \le i, j \le k$, $i \neq j$,

$$x \in S_i, \quad y \in S_j, \quad \Rightarrow \quad (x, y) = 0.$$

The vectors $q_1, \ldots, q_k$ form an orthonormal basis for a subspace $S \subset \mathbf{R}^n$ if they are orthonormal and span $\{q_1, \ldots, q_k\} = S$.

A.1.2 Matrix and Vector Algebra

A **matrix** A is a collection of $m \times n$ real or complex numbers ordered in m rows and n columns:

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}.$$

We write $A \in \mathbf{R}^{m \times n}$, where $\mathbf{R}^{m \times n}$ denotes the set of all real $m \times n$ matrices. For some problems it is more relevant and convenient to work with complex vectors and matrices; $\mathbf{C}^{m \times n}$ denotes the set of $m \times n$ matrices whose components are complex numbers. If $m = n$, then the matrix A is said to be square and of order n. If $m \neq n$, then A is said to be rectangular.

A matrix $A \in \mathbf{R}^{m \times n}$ or $\mathbf{C}^{m \times n}$ can be interpreted as representing a linear transformation on finite-dimensional vector spaces over $\mathbf{R}^n$ or $\mathbf{C}^n$. Consider a linear function $u = F(v)$, $v \in \mathbf{C}^n$, $u \in \mathbf{C}^m$. Let x and y be the column vectors representing the vectors v and $F(v)$, respectively, using the standard basis of the two spaces. Then there is a unique matrix $A \in \mathbf{C}^{m \times n}$ representing this map such that

$$y = Ax.$$

This gives a link between linear maps and matrices.

We will follow a convention introduced by Householder[191] and use uppercase letters (e.g., A, B) to denote matrices. The corresponding lowercase letters with subscripts ij then refer to the (i, j) component of the matrix (e.g., a_{ij}, b_{ij}). Greek letters $\alpha, \beta, \ldots$ are usually used to denote scalars. Column vectors are usually denoted by lower case letters (e.g., x, y).

---

[191]A. S. Householder (1904–1993), at Oak Ridge National Laboratory and University of Tennessee, was a pioneer in the use of matrix factorization and orthogonal transformations in numerical linear algebra.

Two matrices in $\mathbf{R}^{m \times n}$ are said to be *equal*, $A = B$, if

$$a_{ij} = b_{ij}, \quad i = 1 : m, \quad j = 1 : n.$$

The basic operations with matrices are defined as follows. The product of a matrix A with a scalar α is

$$B = \alpha A, \qquad b_{ij} = \alpha a_{ij}.$$

The **sum** of two matrices A and B in $\mathbf{R}^{m \times n}$ is

$$C = A + B, \qquad c_{ij} = a_{ij} + b_{ij}. \tag{A.1.3}$$

The **product** of two matrices A and B is defined if and only if the number of columns in A equals the number of rows in B. If $A \in \mathbf{R}^{m \times n}$ and $B \in \mathbf{R}^{n \times p}$, then

$$C = AB \in \mathbf{R}^{m \times p}, \qquad c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, \tag{A.1.4}$$

and can be computed with $2mnp$ flops. The product BA is defined only if $m = p$.

Matrix multiplication satisfies the distributive rules

$$A(BC) = (AB)C, \qquad A(B + C) = AB + AC. \tag{A.1.5}$$

Note, however, that *the number of arithmetic operations required to compute, respectively, the left- and right-hand sides of these equations can be very different.* Matrix multiplication is, however, *not commutative*, that is, even when both products are defined $AB \neq BA$, in general. In the special case that $AB = BA$ the matrices are said to **commute**.

The **transpose** A^T of a matrix $A = (a_{ij})$ is the matrix whose rows are the columns of A, i.e., if $C = A^T$, then $c_{ij} = a_{ji}$. For the transpose of a product we have

$$(AB)^T = B^T A^T, \tag{A.1.6}$$

i.e., the product of the transposed matrices in *reverse order*. For a complex matrix, A^H denotes the complex conjugate transpose of A

$$A = (a_{ij}), \qquad A^H = (\bar{a}_{ji}),$$

and it holds that $(AB)^H = B^H A^H$.

A **column vector** is a matrix consisting of just one column and we write $x \in \mathbf{R}^n$ instead of $x \in \mathbf{R}^{n \times 1}$. Note that the Euclidean inner product (A.1.1) can be written as

$$(x, y) = x^T y.$$

If $A \in \mathbf{R}^{m \times n}$, $x \in \mathbf{R}^n$, then

$$y = Ax \in \mathbf{R}^m, \qquad y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1 : m.$$

A **row vector** is a matrix consisting of just one row and is obtained by transposing a column vector (e.g., x^T).

It is useful to define **array operations**, which are carried out element by element on vectors and matrices. Let $A = (a_{ij})$ and $B = (b_{ij})$ be two matrices of the same dimensions. Then the **Hadamard product**[192] is defined by

$$C = A . * B \quad \Leftrightarrow \quad c_{ij} = a_{ij} \cdot b_{ij}. \tag{A.1.7}$$

Similarly $A ./ B$ is a matrix with elements a_{ij}/b_{ij}. For the operations $+$ and $-$ the array operations coincide with matrix operations so no distinction is necessary.

A.1.3   Rank and Linear Systems

For a matrix $A \in \mathbf{R}^{m \times n}$ the maximum number of independent row vectors is always equal to the maximum number of independent column vectors. This number r is called the **rank** of A and thus we have $r \leq \min(m, n)$. If rank $(A) = n$, A is said to have full **column rank**; if rank $(A) = m$, A is said to have full **row rank**.

The **outer product** of two vectors $x \in \mathbf{R}^m$ and $y \in \mathbf{R}^n$ is the matrix

$$xy^T = \begin{pmatrix} x_1 y_1 & \dots & x_1 y_n \\ \vdots & & \vdots \\ x_m y_1 & \dots & x_m y_n \end{pmatrix} \in \mathbf{R}^{m \times n}. \tag{A.1.8}$$

Clearly this matrix has rank equal to one.

A square matrix is **nonsingular** and invertible if there exists an **inverse matrix** denoted by A^{-1} with the property that

$$A^{-1}A = AA^{-1} = I.$$

This is the case if and only if A has full row (column) rank. The inverse of a product of two matrices is

$$(AB)^{-1} = B^{-1}A^{-1};$$

i.e., it equals the product of the inverse matrices taken in *reverse order*.

The operations of taking transpose and inverse commutes, i.e., $(A^{-1})^T = (A^T)^{-1}$. Therefore, we can denote the resulting matrix by A^{-T}.

The range and the nullspace of a matrix $A \in \mathbf{R}^{m \times n}$ are

$$\mathcal{R}(A) = \{z \in \mathbf{R}^m \,|\, z = Ax, \ x \in \mathbf{R}^n\}, \tag{A.1.9}$$

$$\mathcal{N}(A) = \{y \in \mathbf{R}^n \,|\, Ay = 0\}. \tag{A.1.10}$$

These are related to the range and nullspace of the transpose matrix A^T by

$$\mathcal{R}(A)^\perp = \mathcal{N}(A^T), \qquad \mathcal{N}(A)^\perp = \mathcal{R}(A^T); \tag{A.1.11}$$

i.e., $\mathcal{N}(A^T)$ is the orthogonal complement to $\mathcal{R}(A)$ and $\mathcal{N}(A)$ the orthogonal complement to $\mathcal{R}(A^T)$. This result is sometimes called the Fundamental Theorem of Linear Algebra.

---

[192]Jacques Salomon Hadamard (1865–1963) was a French mathematician active at the Sorbonne, Collège de France and Ècole Polytechnique in Paris. He made important contributions to geodesics of surfaces and functional analysis. He gave a proof of the result that the number of primes $\leq n$ tends to infinity as $n/\ln n$.

A square matrix $A \in \mathbf{R}^{n \times n}$ is nonsingular if and only if $\mathcal{N}(A) = \{0\}$. A linear system $Ax = b$, $A \in \mathbf{R}^{m \times n}$, is said to be **consistent** if $b \in \mathcal{R}(A)$ or, equivalently if rank $(A, b) = $ rank (A). A consistent linear system always has *at least one solution x*. If $b \notin \mathcal{R}(A)$ or equivalently, rank $(A, b) > $ rank (A), the system is **inconsistent** and has no solution. If $m > n$, there are always right-hand sides b such that $Ax = b$ is inconsistent.

A.1.4   Special Matrices

Any matrix D for which $d_{ij} = 0$ if $i \neq j$ is called a **diagonal matrix**. If $x \in \mathbf{R}^n$ is a vector, then $D = \mathrm{diag}\,(x) \in \mathbf{R}^{n \times n}$ is the diagonal matrix formed by the elements of x. For a matrix $A \in \mathbf{R}^{n \times n}$ the elements a_{ii}, $i = 1 : n$, form the **main diagonal** of A, and we write

$$\mathrm{diag}\,(A) = \mathrm{diag}\,(a_{11}, a_{22}, \ldots, a_{nn}).$$

For $k = 1 : n - 1$ the elements $a_{i,i+k}$ $(a_{i+k,i})$, $i = 1 : n - k$, form the kth **superdiagonal** (**subdiagonal**) of A. The elements $a_{i,n-i+1}$, $i = 1 : n$, form the (main) **antidiagonal** of A.

The **unit matrix** $I = I_n \in \mathbf{R}^{n \times n}$ is defined by

$$I_n = \mathrm{diag}\,(1, 1, \ldots, 1) = (e_1, e_2, \ldots, e_n),$$

and the kth column of I_n is denoted by e_k. We have that $I_n = (\delta_{ij})$, where δ_{ij} is the **Kronecker symbol** $\delta_{ij} = 0$, $i \neq j$, and $\delta_{ij} = 1$, $i = j$. For all square matrices of order n, it holds that $AI = IA = A$. If desirable, we set the size of the unit matrix as a subscript of I, e.g., I_n.

A matrix A for which all nonzero elements are located in consecutive diagonals is called a **band matrix**. A is said to have **upper bandwidth** r if r is the smallest integer such that

$$a_{ij} = 0, \quad j > i + r,$$

and similarly to have **lower bandwidth** s if s is the smallest integer such that

$$a_{ij} = 0, \quad i > j + s.$$

The number of nonzero elements in each row of A is then at most equal to $w = r + s + 1$, which is the **bandwidth** of A. For a matrix $A \in \mathbf{R}^{m \times n}$ which is not square, we define the bandwidth as

$$w = \max_{1 \leq i \leq m} \{j - k + 1 \mid a_{ij} a_{ik} \neq 0\}.$$

Several classes of band matrices that occur frequently have special names. Thus, a matrix for which $r = s = 1$ is called **tridiagonal**; if $r = 0$, $s = 1$ $(r = 1, s = 0)$, it is called lower (upper) **bidiagonal**, etc. A matrix with $s = 1$ $(r = 1)$ is called an upper (lower) **Hessenberg** matrix.

An **upper triangular** matrix is a matrix R for which $r_{ij} = 0$ whenever $i > j$. A square upper triangular matrix has the form

$$R = \begin{pmatrix} r_{11} & r_{12} & \ldots & r_{1n} \\ 0 & r_{22} & \ldots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & r_{nn} \end{pmatrix}.$$

If also $r_{ij} = 0$ when $i = j$, then R is **strictly** upper triangular. Similarly a matrix L is **lower triangular** if $l_{ij} = 0$, $i < j$, and strictly lower triangular if $l_{ij} = 0$, $i \leq j$. Sums, products, and inverses of square upper (lower) triangular matrices are again triangular matrices of the same type.

A square matrix A is called **symmetric** if its elements are symmetric about its main diagonal, i.e., $a_{ij} = a_{ji}$, or equivalently, $A^T = A$. The product of two symmetric matrices is symmetric if and only if A and B commute, that is, $AB = BA$. If $A^T = -A$, then A is called **skew-symmetric**.

For any square nonsingular matrix A, there is a unique **adjoint** matrix A^* such that

$$(x, A^*y) = (Ax, y).$$

The matrix $A \in \mathbf{C}^{n \times n}$ is called **self-adjoint** if $A^* = A$. In particular, for $A \in \mathbf{R}^{n \times n}$ with the standard inner product, we have

$$(Ax, y) = (Ax)^T y = x^T A^T y.$$

Hence $A^* = A^T$, the transpose of A, and A is self-adjoint if it is symmetric. A symmetric matrix A is called **positive definite** if

$$x^T A x > 0 \quad \forall x \in \mathbf{R}^n, \quad x \neq 0, \tag{A.1.12}$$

and **positive semidefinite** if $x^T A x \geq 0$ for all $x \in \mathbf{R}^n$. Otherwise it is called **indefinite**.

Similarly, $A \in \mathbf{C}^{n \times n}$ is self-adjoint or **Hermitian** if $A = A^H$, the conjugate transpose of A. A Hermitian matrix has analogous properties to a real symmetric matrix. If A is Hermitian, then $(x^H A x)^H = x^H A x$ is real, and A is **positive definite** if

$$x^H A x > 0 \quad \forall x \in \mathbf{C}^n, \quad x \neq 0. \tag{A.1.13}$$

Any matrix $A \in \mathbf{C}^{n \times n}$ can be written as the sum of its Hermitian and a skew-Hermitian part, $A = H(A) + S(A)$, where

$$H(A) = \frac{1}{2}(A + A^H), \qquad S(A) = \frac{1}{2}(A - A^H).$$

A is Hermitian if and only if $S(A) = 0$. It is easily seen that A is positive definite if and only if its symmetric part $H(A)$ is positive definite. For the vector space $\mathbf{R}^n$ ($\mathbf{C}^n$), any inner product can be written as

$$(x, y) = x^T G y \qquad ((x, y) = x^H G y),$$

where the matrix G is positive definite.

Let $q_1, \ldots, q_n \in \mathbf{R}^m$ be orthonormal and form the matrix

$$Q = (q_1, \ldots, q_n) \in \mathbf{R}^{m \times n}, \quad m \geq n.$$

Then Q is called an **orthogonal matrix** and $Q^T Q = I_n$. If Q is square ($m = n$), then it also holds that $Q^{-1} = Q^T$, $Q Q^T = I_n$.

Two vectors x and y in $\mathbf{C}^n$ are called orthogonal if $x^H y = 0$. A square matrix U for which $U^H U = I$ is called **unitary**, and from (A.1.2) we find that

$$(Ux)^H U y = x^H U^H U y = x^H y.$$

A.2   Submatrices and Block Matrices

A matrix formed by the elements at the intersection of a set of rows and columns of a matrix A is called a **submatrix**. For example, the matrices

$$\begin{pmatrix} a_{22} & a_{24} \\ a_{42} & a_{44} \end{pmatrix}, \qquad \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}$$

are submatrices of A. The second submatrix is called a contiguous submatrix since it is formed by contiguous elements of A.

Definition A.2.1.

   A **submatrix** *of* $A = (a_{ij}) \in \mathbf{R}^{m \times n}$ *is a matrix* $B \in \mathbf{R}^{p \times q}$ *formed by selecting p rows and q columns of A,*

$$B = \begin{pmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \cdots & a_{i_1 j_q} \\ a_{i_2 j_1} & a_{i_2 j_2} & \cdots & a_{i_2 j_q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_p j_1} & a_{i_p j_2} & \cdots & a_{i_p j_q} \end{pmatrix},$$

where

$$1 \le i_1 \le i_2 \le \cdots \le i_p \le m, \qquad 1 \le j_1 \le j_2 \le \cdots \le j_q \le n.$$

If $p = q$ and $i_k = j_k$, $k = 1 : p$, then B is a **principal submatrix** *of A. If in addition, $i_k = j_k = k$, $k = 1 : p$, then B is a* **leading** *principal submatrix of A.*

   It is often convenient to think of a matrix (vector) as being built up of contiguous submatrices (subvectors) of lower dimensions. This can be achieved by **partitioning** the matrix or vector into blocks. We write, e.g.,

$$A = \begin{matrix} \\ p_1 \{ \\ p_2 \{ \\ \vdots \\ p_M \{ \end{matrix} \begin{pmatrix} q_1 & q_2 & \cdots & q_N \\ A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} \end{pmatrix}, \qquad x = \begin{matrix} p_1 \{ \\ p_2 \{ \\ \vdots \\ p_M \{ \end{matrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{pmatrix}, \qquad (A.2.1)$$

where A_{IJ} is a matrix of dimension $p_I \times q_J$. We call such a matrix a **block matrix**. The partitioning can be carried out in many ways and is often suggested by the structure of the underlying problem. For square matrices the most important case is when $M = N$, and $p_I = q_I$, $I = 1 : N$. Then the diagonal blocks A_{II}, $I = 1 : N$, are square matrices.

   The great convenience of block matrices lies in the fact that the operations of addition and multiplication can be performed by treating the blocks A_{IJ} as *non-commuting scalars*.

   Let $A = (A_{IK})$ and $B = (B_{KJ})$ be two block matrices of block dimensions $M \times N$ and $N \times P$, respectively, where the partitioning corresponding to the index K is the same for each matrix. Then we have $C = AB = (C_{IJ})$, where

$$C_{IJ} = \sum_{K=1}^{N} A_{IK} B_{KJ}, \quad 1 \le I \le M, \quad 1 \le J \le P. \qquad (A.2.2)$$

Therefore many algorithms defined for matrices with scalar elements have another simple generalization to partitioned matrices. Of course the dimensions of the blocks must correspond in such a way that the operations can be performed. When this is the case, the matrices are said to be partitioned **conformally**.

The **colon notation** used in MATLAB is very convenient for handling partitioned matrices and will be used throughout this volume:

$j : k$ is the same as the vector $[j, j + 1, \ldots, k]$,

$j : k$ is empty if $j > k$,

$j : i : k$ is the same as the vector $[j, j + i, , j + 2i, \ldots, k]$,

$j : i : k$ is empty if $i > 0$ and $j > k$ or if $i < 0$ and $j < k$.

The colon notation is used to pick out selected rows, columns, and elements of vectors and matrices, for example,

$x(j : k)$ is the vector$[x(j), x(j + 1), \ldots, x(k)]$,

$A(:, j)$ is the jth column of A,

$A(i, :)$ is the ith row of A,

$A(:, :)$ is the same as A,

$A(:, j : k)$ is the matrix$[A(:, j), A(:, j + 1), \ldots, A(:, k)]$,

$A(:)$ is all the elements of the matrix A regarded as a single column.

The various special forms of matrices have analogue block forms. For example, R is block upper triangular if it has the form

$$
R = \begin{pmatrix}
R_{11} & R_{12} & R_{13} & \cdots & R_{1N} \\
0 & R_{22} & R_{23} & \cdots & R_{2N} \\
0 & 0 & R_{33} & \cdots & R_{3N} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & R_{NN}
\end{pmatrix}.
$$

Example A.2.1.

Partitioning a matrix into a block 2×2 matrix with square diagonal blocks is particularly useful. For this case we have

$$
\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}. \tag{A.2.3}
$$

Be careful to note that since matrix multiplication is not commutative the *order* of the factors in the products cannot be changed! In the special case of block upper triangular matrices this reduces to

$$
\begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix} \begin{pmatrix} S_{11} & S_{12} \\ 0 & S_{22} \end{pmatrix} = \begin{pmatrix} R_{11}S_{11} & R_{11}S_{12} + R_{12}S_{22} \\ 0 & R_{22}S_{22} \end{pmatrix}.
$$

Note that the product is again block upper triangular and its block diagonal simply equals the products of the diagonal blocks of the factors.

A.2.1   Block Gaussian Elimination

Let

$$L = \begin{pmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{pmatrix}, \qquad U = \begin{pmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{pmatrix} \tag{A.2.4}$$

be 2×2 block lower and block upper triangular matrices, respectively. We assume that the diagonal blocks are square but not necessarily triangular. Generalizing (A.3.5) it then holds that

$$\det(L) = \det(L_{11}) \det(L_{22}), \qquad \det(U) = \det(U_{11}) \det(U_{22}). \tag{A.2.5}$$

Hence L and U are nonsingular if and only if the diagonal blocks are nonsingular. If they are nonsingular, their inverses are given by

$$L^{-1} = \begin{pmatrix} L_{11}^{-1} & 0 \\ -L_{22}^{-1} L_{21} L_{11}^{-1} & L_{22}^{-1} \end{pmatrix}, \qquad U^{-1} = \begin{pmatrix} U_{11}^{-1} & -U_{11}^{-1} U_{12} U_{22}^{-1} \\ 0 & U_{22}^{-1} \end{pmatrix}. \tag{A.2.6}$$

This can be verified by forming the products $L^{-1}L$ and $U^{-1}U$ using the rule for multiplying partitioned matrices.

We now give some formulas for the inverse of a block 2×2 matrix,

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{A.2.7}$$

where A and D are square matrices. If A is nonsingular, we can factor M in a product of a block lower and a block upper triangular matrix,

$$M = \begin{pmatrix} I & 0 \\ CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & B \\ 0 & S \end{pmatrix}, \quad S = D - CA^{-1}B. \tag{A.2.8}$$

This identity, which is equivalent to block Gaussian elimination, can be verified directly. The matrix S is the **Schur complement** of A in M.[193]

From $M^{-1} = (LU)^{-1} = U^{-1}L^{-1}$, using the formulas (A.2.6) for the inverses of 2×2 block triangular matrices we get the **Banachiewicz** inversion formula[194]

$$M^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}BS^{-1} \\ 0 & S^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -CA^{-1} & I \end{pmatrix}$$

$$= \begin{pmatrix} A^{-1} + A^{-1}BS^{-1}CA^{-1} & -A^{-1}BS^{-1} \\ -S^{-1}CA^{-1} & S^{-1} \end{pmatrix}. \tag{A.2.9}$$

Similarly, assuming that D is nonsingular, we can factor M into a product of a block upper and a block lower triangular matrix

$$M = \begin{pmatrix} I & BD^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} T & 0 \\ C & D \end{pmatrix}, \quad T = A - BD^{-1}C, \tag{A.2.10}$$

---

[193]Issai Schur (1875–1941) was born in Russia but studied at the University of Berlin, where he became full professor in 1919. Schur is mainly known for his fundamental work on the theory of groups, but he also worked in the field of matrices.

[194]Tadeusz Banachiewicz (1882–1954) was a Polish astronomer and mathematician. In 1919 he became the director of Cracow Observatory. In 1925 he developed a special kind of matrix algebra for "cracovians" which brought him international recognition.

where T is the Schur complement of D in M. (This is equivalent to block Gaussian elimination in reverse order.) From this factorization an alternative expression of M^{-1} can be derived,

$$M^{-1} = \begin{pmatrix} T^{-1} & -T^{-1}BD^{-1} \\ -D^{-1}CT^{-1} & D^{-1} + D^{-1}CT^{-1}BD^{-1} \end{pmatrix}. \qquad \text{(A.2.11)}$$

If A and D are nonsingular, then both triangular factorizations (A.2.8) and (A.2.10) exist.

An important special case of the first Banachiewicz inversion formula (A.2.9) is when the block D is a scalar,

$$M = \begin{pmatrix} A & b \\ c^T & \delta \end{pmatrix}. \qquad \text{(A.2.12)}$$

Then if the Schur complement $\sigma = \delta - c^T A^{-1} b \neq 0$, we obtain for the inverse the formula

$$M^{-1} = \begin{pmatrix} A^{-1} + \sigma^{-1} A^{-1} b c^T A^{-1} & -\sigma^{-1} A^{-1} b \\ -\sigma^{-1} c^T A^{-1} & \sigma^{-1} \end{pmatrix}. \qquad \text{(A.2.13)}$$

This formula is convenient to use in case it is necessary to solve a system for which the truncated system, obtained by crossing out one equation and one unknown, has been solved earlier. Such a situation is often encountered in applications.

The formula can also be used to invert a matrix by successive bordering, where one constructs in succession the inverse of matrices

$$(a_{11}), \quad \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \ldots.$$

Each step is then carried by using the formula (A.2.13).

The formulas for the inverse of a block 2×2 matrix can be used to derive expressions for the inverse of a matrix $A \in \mathbf{R}^{n \times n}$ modified by a matrix of rank p. Any matrix of rank $p \leq n$ can be written as $BD^{-1}C$, where $B \in \mathbf{R}^{p \times n}$, $C \in \mathbf{R}^{p \times n}$, and $D \in \mathbf{R}^{n \times n}$ is nonsingular. (The factor D is not necessary but included for convenience.) Assuming that $A - BD^{-1}C$ is nonsingular and equating the $(1, 1)$ blocks in the inverse M^{-1} in (A.2.9) and (A.2.11), we obtain the **Woodbury formula**,

$$(A - BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1}. \qquad \text{(A.2.14)}$$

This gives an expression for the inverse of a matrix A after it has been modified by a matrix of rank p, a very useful result in situations where $p \ll n$.

If we specialize the Woodbury formula to the case where D is a scalar and

$$M = \begin{pmatrix} A & u \\ v^T & 1/\sigma \end{pmatrix},$$

we get the well-known **Sherman–Morrison formula**,

$$(A - \sigma u v^T)^{-1} = A^{-1} + \alpha(A^{-1}u)(v^T A^{-1}), \quad \alpha = \frac{\sigma}{1 - \sigma\, v^T A^{-1} u}. \qquad \text{(A.2.15)}$$

It follows that $A - \sigma u v^T$ is nonsingular if and only if $\sigma \neq 1/v^T A^{-1} u$. The Sherman–Morrison formula can be used to compute the new inverse when a matrix A is modified by a matrix of rank one.

Frequently it is required to solve a linear problem where the matrix has been modified by a correction of low rank. Consider first a linear system $Ax = b$, where A is modified by a correction of rank one,

$$(A - \sigma u v^T)\hat{x} = b. \tag{A.2.16}$$

Using the Sherman–Morrison formula, we can write the solution as

$$(A - \sigma u v^T)^{-1}b = A^{-1}b + \alpha(A^{-1}u)(v^T A^{-1}b), \quad \alpha = 1/(\sigma^{-1} - v^T A^{-1}u).$$

Here $x = A^{-1}b$ is the solution to the original system and $v^T A^{-1}b = v^T x$ is a scalar. Hence,

$$\hat{x} = x + \beta w, \qquad \beta = v^T x/(\sigma^{-1} - v^T w), \quad w = A^{-1}u, \tag{A.2.17}$$

which shows that the solution $\hat{x}$ can be obtained from x by solving the system $Aw = u$. Note that computing A^{-1} can be avoided.

We caution that the updating formulas given here cannot be expected to be numerically stable in all cases. This is related to the fact that pivoting is necessary in Gaussian elimination.

A.3 Permutations and Determinants

The classical definition of the **determinant**[195] requires some elementary facts about permutations which we now state.

Let $\alpha = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a permutation of the integers $\{1, 2, \ldots, n\}$. The pair α_r, α_s, $r < s$, is said to form an inversion in the permutation if $\alpha_r > \alpha_s$. For example, in the permutation $\{2, \ldots, n, 1\}$ there are $(n-1)$ inversions $(2, 1), (3, 1), \ldots, (n, 1)$. A permutation α is said to be even and sign $(\alpha) = 1$ if it contains an even number of inversions; otherwise the permutation is odd and sign $(\alpha) = -1$.

The product of two permutations σ and τ is the composition $\sigma\tau$ defined by

$$\sigma\tau(i) = \sigma[\tau(i)], \quad i = 1:n.$$

A **transposition** τ is a permutation which interchanges only two elements. Any permutation can be decomposed into a sequence of transpositions, but this decomposition is not unique.

A **permutation matrix** $P \in \mathbf{R}^{n\times n}$ is a matrix whose columns are a permutation of the columns of the unit matrix, that is,

$$P = (e_{p_1}, \ldots, e_{p_n}),$$

where $p_1, \ldots, p_n$ is a permutation of $1, \ldots, n$. Notice that in a permutation matrix every row and every column contains just one unity element. Since P is uniquely represented by the integer vector $p = (p_1, \ldots, p_n)$ it need never be explicitly stored. For example, the

---

[195]Determinants were first introduced by Leibniz in 1693 and then by Cayley in 1841. Determinants arise in many parts of mathematics such as combinatorial enumeration, graph theory, representation theory, statistics, and theoretical computer science. The theory of determinants is covered in the monumental five-volume work *The Theory of Determinants in the Historical Order of Development* by Thomas Muir (1844–1934).

vector $p = (2, 4, 1, 3)$ represents the permutation matrix

$$P = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

If P is a permutation matrix, then PA is the matrix A with its rows permuted and AP is A with its columns permuted. Using the colon notation, we can write these permuted matrices as PA = A(p,:) and PA = A(:,p), respectively.

The transpose P^T of a permutation matrix is again a permutation matrix. Any permutation may be expressed as a sequence of transposition matrices. Therefore any permutation matrix can be expressed as a product of transposition matrices $P = I_{i_1,j_1} I_{i_2,j_2} \cdots I_{i_k,j_k}$. Since $I_{i_p,j_p}^{-1} = I_{i_p,j_p}$, we have

$$P^{-1} = I_{i_k,j_k} \cdots I_{i_2,j_2} I_{i_1,j_1} = P^T;$$

that is, permutation matrices are orthogonal and P^T performs the reverse permutation, and thus,

$$P^T P = P P^T = I. \tag{A.3.1}$$

Lemma A.3.1.

A transposition τ of a permutation will change the number of inversions in the permutation by an odd number, and thus $\text{sign}\,(\tau) = -1$.

Proof. If τ interchanges two adjacent elements α_r and α_{r+1} in the permutation $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, this will not affect inversions in other elements. Hence the number of inversions increases by 1 if $\alpha_r < \alpha_{r+1}$ and decreases by 1 otherwise. Suppose now that τ interchanges α_r and α_{r+q}. This can be achieved by first successively interchanging α_r with α_{r+1}, then with α_{r+2}, and finally with α_{r+q}. This takes q steps. Next the element α_{r+q} is moved in $q - 1$ steps to the position which α_r previously had. In all it takes an *odd number* $2q - 1$ of transpositions of adjacent elements, in each of which the sign of the permutation changes. $\square$

Definition A.3.2.

*The **determinant** of a square matrix $A \in \mathbf{R}^{n \times n}$ is the scalar*

$$\det(A) = \sum_{\alpha \in S_n} \text{sign}\,(\alpha)\, a_{1,\alpha_1} a_{2,\alpha_2} \cdots a_{n,\alpha_n}, \tag{A.3.2}$$

where the sum is over all $n!$ permutations of the set $\{1, \ldots, n\}$ and $\text{sign}\,(\alpha) = \pm 1$ according to whether α is an even or odd permutation.

Note that there are $n!$ terms in (A.3.2) and each term contains exactly one factor from each row and each column in A. For example, if $n = 2$, there are two terms, and

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

From the definition, it follows easily that

$$\det(\alpha A) = \alpha^n \det(A), \qquad \det(A^T) = \det(A).$$

If we collect all terms in (A.3.2) that contain the element a_{rs}, the sum of these terms can be written as $a_{rs} A_{rs}$, where A_{rs} is called the complement of a_{rs}. Since the determinant contains only one element from row r and column s, the complement A_{rs} does not depend on any elements in row r and column s. Since each product in (A.3.2) contains precisely one element of the elements $a_{r1}, a_{r2}, \ldots, a_{rn}$ in row r, it follows that

$$\det(A) = a_{r1} A_{r1} + a_{r2} A_{r2} + \cdots + a_{rn} A_{rn}. \tag{A.3.3}$$

This is called to expand the determinant after the row r. It is not difficult to verify that

$$A_{rs} = (-1)^{r+s} D_{rs}, \tag{A.3.4}$$

where D_{rs} is the determinant of the matrix of order $n-1$ obtained by striking out row r and column s in A. Since $\det(A) = \det(A^T)$, it is clear that we can similarly expand $\det(A)$ after a column.

The direct use of the definition (A.3.2) to evaluate $\det(A)$ would require about $nn!$ operations, which rapidly becomes infeasible as n increases. A much more efficient way to compute $\det(A)$ is by repeatedly using the following properties.

Theorem A.3.3.
  (i) *The value of the* $\det(A)$ *is unchanged if a row (column) in A multiplied by a scalar is added to another row (column).*

 (ii) *The determinant of a triangular matrix equals the product of the elements in the main diagonal; i.e., if U is upper triangular,*

$$\det(U) = u_{11} u_{22} \cdots u_{nn}. \tag{A.3.5}$$

 (iii) *If two rows (columns) in A are interchanged, the value of* $\det(A)$ *is multiplied by* (-1).

 (iv) *The product rule* $\det(AB) = \det(A) \det(B)$.

If Q is an orthogonal matrix, then $Q^T Q = I_n$. Then using (iv) it follows that

$$1 = \det(I) = \det(Q^T Q) = \det(Q^T) \det(Q) = (\det(Q))^2,$$

and hence $\det(Q) = \pm 1$. If $\det(Q) = 1$, then Q is a rotation.

Theorem A.3.4.
  The matrix A is nonsingular if and only if $\det(A) \neq 0$. *If the matrix A is nonsingular, then the solution of the linear system* $Ax = b$ *can be expressed as*

$$x_j = \det(B_j)/\det(A), \quad j = 1 : n. \tag{A.3.6}$$

Here B_j is the matrix A, where the jth column has been replaced by the right-hand side vector b.

Proof. We have

$$a_{1j}A_{1r} + a_{2j}A_{2r} + \cdots + a_{nj}A_{nr} = \begin{cases} 0 & \text{if } j \neq r, \\ \det(A) & \text{if } j = r, \end{cases} \qquad \text{(A.3.7)}$$

where the linear combination is formed with elements from column j and the complements of column r. If $j = r$, this is an expansion after column r of $\det(A)$. If $j \neq r$, the expression is the expansion of the determinant of a matrix equal to A except that column r is equal to column j. Such a matrix has a determinant equal to 0.

Now take the ith equation in $Ax = b$,

$$a_{i1}x_1 + a_{i2}x_2 + \cdots a_{in}x_n = b_i,$$

multiply by A_{ir}, and sum over $i = 1 : n$. Then by (A.3.7) the coefficients of x_j, $j \neq r$, vanish and we get

$$\det(A)x_r = b_1 A_{1r} + b_2 A_{2r} + \cdots b_n A_{nr}.$$

The right-hand side equals $\det(B_r)$ expanded by its rth column, which proves (A.3.6). □

The expression (A.3.6) is known as **Cramer's rule**.[196] Although elegant, it is both computationally expensive and numerically unstable, even for $n = 2$.

Let U be an upper block triangular matrix with square diagonal blocks U_{II}, $I = 1 : N$. Then

$$\det(U) = \det(U_{11}) \det(U_{22}) \cdots \det(U_{NN}), \qquad \text{(A.3.8)}$$

and thus U is nonsingular if and only if all its diagonal blocks are nonsingular. Since $\det(L) = \det(L^T)$, a similar result holds for a lower block triangular matrix.

Example A.3.1.

For the 2×2 block matrix M in (A.2.8) and (A.2.10), it follows by using (A.3.8) that

$$\det(M) = \det(A - BD^{-1}C)\det(D) = \det(A)\det(D - CA^{-1}B).$$

In the special case that $D^{-1} = \lambda$, $B = x$, and $B = y$, this gives

$$\det(A - \lambda xy^T) = \det(A)(1 - \lambda y^T A^{-1}x). \qquad \text{(A.3.9)}$$

This shows that $\det(A - \lambda xy^T) = 0$ if $\lambda = 1/y^T A^{-1}x$, a fact which is useful for the solution of eigenvalue problems.

---

[196]Named after the Swiss mathematician Gabriel Cramer (1704–1752).

A.4   Eigenvalues and Norms of Matrices

A.4.1   The Characteristic Equation

Of central importance in the study of matrices are the special vectors whose directions are not changed when multiplied by A. A complex scalar λ such that

$$Ax = \lambda x, \quad x \neq 0, \tag{A.4.1}$$

is called an **eigenvalue** of A and x is an **eigenvector** of A. Eigenvalues and eigenvectors give information about the behavior of evolving systems governed by a matrix or operator and are fundamental tools in the mathematical sciences and in scientific computing.

From (A.4.1) it follows that λ is an eigenvalue if and only if the linear homogeneous system $(A - \lambda I)x = 0$ has a nontrivial solution $x \neq 0$, or equivalently, if and only if $A - \lambda I$ is singular. It follows that the eigenvalues satisfy the **characteristic equation**

$$p(\lambda) = \det(A - \lambda I) = 0. \tag{A.4.2}$$

Obviously, if x is an eigenvector, so is αx for any scalar $\alpha \neq 0$.

The polynomial $p(\lambda) = \det(A - \lambda I)$ is the **characteristic polynomial** of the matrix A. Expanding the determinant in (A.4.2), it follows that $p(\lambda)$ has the form

$$p(\lambda) = (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) + q(\lambda), \tag{A.4.3}$$

where $q(\lambda)$ has degree at most $n - 2$. Hence $p(\lambda)$ is a polynomial of degree n in λ with leading term $(-1)^n \lambda^n$. By the fundamental theorem of algebra the matrix A has exactly n (possibly complex) eigenvalues λ_i, $i = 1, 2, \ldots, n$, counting multiple roots according to their multiplicities. The set of eigenvalues of A is called the **spectrum** of A. The largest modulus of an eigenvalue is called the **spectral radius** and denoted by

$$\rho(A) = \max_i |\lambda_i(A)|. \tag{A.4.4}$$

Putting $\lambda = 0$ in $p(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda)$ and (A.4.2), it follows that

$$p(0) = \lambda_1 \lambda_2 \cdots \lambda_n = \det(A). \tag{A.4.5}$$

Consider the linear transformation $y = Ax$, where $A \in \mathbf{R}^{n \times n}$. Let V be nonsingular and suppose we change the basis by setting $x = V\xi$, $y = V\eta$. The column vectors ξ and η then represent the vectors x and y with respect to the basis $V = (v_1, \ldots, v_n)$. Now $V\eta = AV\xi$, and hence $\eta = V^{-1}AV\xi$. This shows that the matrix

$$B = V^{-1}AV$$

represents the operator A in the new basis. The mapping $A \rightarrow B = V^{-1}AV$ is called a **similarity transformation**. If $Ax = \lambda x$, then

$$V^{-1}AVy = By = \lambda y, \quad y = V^{-1}x,$$

which shows the important facts that B has the same eigenvalues as A and that the eigenvectors of B can be easily computed from those of A. In other words, eigenvalues are

properties of the operator itself and are independent of the basis used for its representation by a matrix.

The **trace** of a square matrix of order n is the sum of its diagonal elements

$$\text{trace}\,(A) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i. \tag{A.4.6}$$

The last equality follows by using the relation between the coefficients and roots of the characteristic equation. Hence the trace of the matrix is invariant under similarity transformations.

A.4.2 The Schur and Jordan Normal Forms

Given $A \in \mathbf{C}^{n \times n}$ there exists a unitary matrix $U \in \mathbf{C}^{n \times n}$ such that

$$U^H A U = T = \begin{pmatrix} \lambda_1 & t_{12} & \ldots & t_{1n} \\ & \lambda_2 & \ldots & t_{2n} \\ & & \ddots & \vdots \\ & & & \lambda_n \end{pmatrix},$$

where T is upper triangular. This is the **Schur normal form** of A. (A proof will be given in Chapter 9 of Volume II.) Since

$$\det(T - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda),$$

the diagonal elements $\lambda_1, \ldots, \lambda_n$ of T are the eigenvalues of A.

Each *distinct* eigenvalue λ_i has at least one eigenvector v_i. Let $V = (v_1, \ldots, v_k)$ be eigenvectors corresponding to the eigenvalues $\Lambda = \text{diag}\,(\lambda_1, \ldots, \lambda_k)$ of a matrix A. Then, we can write

$$AV = V\Lambda.$$

If there are n linearly independent eigenvectors, then $V = (v_1, \ldots, v_n)$ is nonsingular and

$$A = V\Lambda V^{-1}, \qquad \Lambda = V^{-1}AV.$$

Then A is said to be **diagonalizable**.

A matrix $A \in \mathbf{C}^{n \times n}$ is said to be **normal** if $A^H A = AA^H$. For a normal matrix the upper triangular matrix T in the Schur normal form is also normal, i.e.,

$$T^H T = TT^H.$$

It can be shown that this relation implies that all nondiagonal elements in T vanish, i.e., $T = \Lambda$. Then we have $AU = UT = U\Lambda$, where $\Lambda = \text{diag}\,(\lambda_i)$, or with $U = (u_1, \ldots, u_n)$,

$$Au_i = \lambda_i u_i, \quad i = 1 : n.$$

This shows the important result that *a normal matrix always has a set of mutually unitary (orthogonal) eigenvectors*.

Important classes of normal matrices are Hermitian ($A = A^H$), skew-Hermitian ($A^H = -A$), and unitary ($A^{-1} = A^H$). Hermitian matrices have real eigenvalues, skew-Hermitian matrices have imaginary eigenvalues, and unitary matrices have eigenvalues on the unit circle (see Chapter 9 of Volume II).

An example of a nondiagonalizable matrix is

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in \mathbf{C}^{m \times m}.$$

The matrix $J_m(\lambda)$ is called a **Jordan block**. It has one eigenvalue λ of multiplicity m to which corresponds only one eigenvector,

$$J_m(\lambda)e_1 = \lambda e_1, \quad e_1 = (1, 0, \ldots, 0)^T.$$

A.4.3   Norms of Vectors and Matrices

In many applications it is useful to have a measure of the size of a vector or a matrix. An example is the quantitative discussion of errors in matrix computation. Such measures are provided by vector and matrix norms, which can be regarded as generalizations of the absolute value function on $\mathbf{R}$.

A **norm** on the vector space $\mathbf{C}^n$ is a function $\mathbf{C}^n \to \mathbf{R}$ denoted by $\| \cdot \|$ that satisfies the following three conditions:

1. $\|x\| > 0 \quad \forall x \in \mathbf{C}^n, \quad x \neq 0$ (definiteness),

2. $\|\alpha x\| = |\alpha| \|x\| \quad \forall \alpha \in \mathbf{C}, \quad x \in \mathbf{C}^n$ (homogeneity),

3. $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in \mathbf{C}^n$ (triangle inequality).

The triangle inequality is often used in the form (see Problem A.11)

$$\|x \pm y\| \geq \big| \|x\| - \|y\| \big|.$$

The most common vector norms are special cases of the family of **Hölder** norms, or p-norms,

$$\|x\|_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}, \qquad 1 \leq p < \infty. \tag{A.4.7}$$

The p-norms have the property that $\|x\|_p = \| |x| \|_p$. Vector norms with this property are said to be **absolute**. The three most important particular cases are $p = 1$ (the 1-norm), $p = 2$ (the Euclidean norm), and the limit when $p \to \infty$ (the maximum norm):

$$\begin{aligned} \|x\|_1 &= |x_1| + \cdots + |x_n|, \\ \|x\|_2 &= (|x_1|^2 + \cdots + |x_n|^2)^{1/2} = (x^H x)^{1/2}, \\ \|x\|_\infty &= \max_{1 \leq i \leq n} |x_i|. \end{aligned} \tag{A.4.8}$$

2008/3/27
page A-19

If Q is unitary, then

$$\|Qx\|_2^2 = x^H Q^H Q x = x^H x = \|x\|_2^2,$$

that is, the Euclidean norm is invariant under unitary (orthogonal) transformations.

The proof that the triangle inequality is satisfied for the p-norms depends on the following inequality. Let $p > 1$ and q satisfy $1/p + 1/q = 1$. Then it holds that

$$\alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q}.$$

Indeed, let x and y be any real number and λ satisfy $0 < \lambda < 1$. Then by the convexity of the exponential function, it holds that

$$e^{\lambda x + (1-\lambda)y} \leq \lambda e^x + (1-\lambda)e^y.$$

We obtain the desired result by setting $\lambda = 1/p$, $x = p \log \alpha$, and $y = q \log \beta$.

Another important property of the p-norms is the **Hölder inequality**

$$|x^H y| \leq \|x\|_p \|y\|_q, \quad \frac{1}{p} + \frac{1}{q} = 1, \quad p \geq 1. \tag{A.4.9}$$

For $p = q = 2$ this becomes the **Cauchy–Schwarz inequality**

$$|x^H y| \leq \|x\|_2 \|y\|_2.$$

Norms can be obtained from inner products by taking

$$\|x\|^2 = (x, x) = x^H G x,$$

where G is Hermitian and positive definite. It can be shown that the unit ball $\{x : \|x\| \leq 1\}$ corresponding to this norm is an ellipsoid, and hence such norms are also called elliptic norms. A special useful case involves the **scaled** p-norms defined by

$$\|x\|_{p,D} = \|Dx\|_p, \quad D = \text{diag}\,(d_1, \ldots, d_n), \quad d_i \neq 0, \quad i = 1 : n. \tag{A.4.10}$$

All norms on $\mathbf{C}^n$ are equivalent in the following sense: For each pair of norms $\|\cdot\|$ and $\|\cdot\|'$ there are positive constants c and c' such that

$$\frac{1}{c}\|x\|' \leq \|x\| \leq c'\|x\|' \quad \forall x \in \mathbf{C}^n. \tag{A.4.11}$$

In particular, it can be shown that for the p-norms we have

$$\|x\|_q \leq \|x\|_p \leq n^{\left(\frac{1}{p} - \frac{1}{q}\right)} \|x\|_q, \quad 1 \leq p \leq q \leq \infty. \tag{A.4.12}$$

We now consider **matrix norms**. We can construct a matrix norm from a vector norm by defining

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\|. \tag{A.4.13}$$

This norm is called the **operator norm**, or the matrix norm **subordinate** to the vector norm. From this definition it follows directly that

$$\|Ax\| \le \|A\| \, \|x\|, \quad x \in \mathbf{C}^n.$$

Whenever this inequality holds, we say that the matrix norm is **consistent** with the vector norm. For any operator norm it holds that $\|I_n\|_p = 1$.

It is an easy exercise to show that operator norms are **submultiplicative**; i.e., whenever the product AB is defined it satisfies the condition

    4. $\|AB\| \le \|A\| \, \|B\|$.

The matrix norms

$$\|A\|_p = \sup_{\|x\|=1} \|Ax\|_p, \quad p = 1, 2, \infty,$$

subordinate to the vector p-norms are especially important. The 1-norm and ∞-norm are easily computable from

$$\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|, \qquad \|A\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|, \tag{A.4.14}$$

respectively. Note that the 1-norm equals the maximal column sum and the ∞-norm equals the maximal row sum of the magnitude of the elements. Consequently $\|A\|_1 = \|A^H\|_\infty$.

The 2-norm is also called the **spectral norm**,

$$\|A\|_2 = \sup_{\|x\|=1} (x^H A^H A x)^{1/2} = \sigma_1(A), \tag{A.4.15}$$

where $\sigma_1(A)$ is the largest singular value of A. Its major drawback is that it is expensive to compute. Since the nonzero eigenvalues of $A^H A$ and $A A^H$ are the same it follows that $\|A\|_2 = \|A^H\|_2$. A useful upper bound for the matrix 2-norm is

$$\|A\|_2 \le (\|A\|_1 \|A\|_\infty)^{1/2}. \tag{A.4.16}$$

The proof of this bound is left as an exercise.

Another way to proceed in defining norms for matrices is to regard $\mathbf{C}^{m \times n}$ as an mn-dimensional vector space and apply a vector norm over that space. With the exception of the **Frobenius norm**[197] derived from the vector 2-norm,

$$\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}; \tag{A.4.17}$$

such norms are not much used. Note that $\|A^H\|_F = \|A\|_F$. Useful alternative characterizations of the Frobenius norm are

$$\|A\|_F^2 = \text{trace}\,(A^H A) = \sum_{i=1}^{k} \sigma_i^2(A), \quad k = \min(m, n), \tag{A.4.18}$$

---

[197]Ferdinand George Frobenius (1849–1917) a German mathematician, was a professor at ETH Zürich from 1875 to 1892 before he succeeded Weierstrass at Berlin University.

where $\sigma_i(A)$ are the nonzero singular values of A. The Frobenius norm is submultiplicative. However, it is often larger than necessary, e.g., $\|I_n\|_F = n^{1/2}$. This tends to make bounds derived in terms of the Frobenius norm not as sharp as they might be. From (A.4.15) and (A.4.18) we also get lower and upper bounds for the matrix 2-norm,

$$\frac{1}{\sqrt{k}}\|A\|_F \leq \|A\|_2 \leq \|A\|_F, \quad k = \min(m, n).$$

An important property of the Frobenius norm and the 2-norm is that both are invariant with respect to unitary (real orthogonal) transformations.

Lemma A.4.1. *For all unitary (real orthogonal) matrices U and V ($U^H U = I$ and $V^H V = I$) of appropriate dimensions, it holds that*

$$\|UAV\| = \|A\| \tag{A.4.19}$$

for the Frobenius norm and the 2-norm.

We finally remark that the 1-, ∞- and Frobenius norms satisfy

$$\| |A| \| = \|A\|, \qquad |A| = (|a_{ij}|),$$

but for the 2-norm the best result is that $\| |A| \|_2 \leq n^{1/2}\|A\|_2$.

One use of norms is in the study of *limits of sequences of vectors and matrices* (see Sec. 9.2.4 in Volume II). Consider an infinite sequence $x_1, x_2, \ldots$ of elements of a vector space $\mathcal{V}$ and let $\| \cdot \|$ be a norm on $\mathcal{V}$. The sequence is said to converge (strongly if V is infinite-dimensional) to a limit $x \in \mathcal{V}$, and so we write $\lim_{k\to\infty} x_k = x$ if

$$\lim_{k\to\infty} \|x_k - x\| = 0.$$

For a finite-dimensional vector space the equivalence of norms (A.4.11) shows that convergence is independent of the choice of norm. The particular choice of $\| \cdot \|_\infty$ shows that convergence of vectors in $\mathbf{C}^n$ is equivalent to convergence of the n sequences of scalars formed by the components of the vectors. By considering matrices in $\mathbf{C}^{m\times n}$ as vectors in $\mathbf{C}^{mn}$, we see that the same conclusion holds for matrices.

Review Questions

A.1. Define the following concepts:

   (i) Real symmetric matrix.   (ii) Real orthogonal matrix.

  (iii) Real skew-symmetric matrix.   (iv) Triangular matrix.

   (v) Hessenberg matrix.

A.2. (a) What is the Schur normal form of a matrix $A \in \mathbf{C}^{n\times n}$?

(b) What is meant by a normal matrix? How does the Schur form simplify for a normal matrix?

A.3. Define the matrix norm subordinate to a given vector norm.

A.4. Define the p-norm of a vector x. Give explicit expressions for the matrix p-norms for $p = 1, 2, \infty$. Show that

$$\|x\|_1 \leq \sqrt{n}\|x\|_2 \leq n\|x\|_\infty,$$

which are special cases of (A.4.12).

Problems

A.1. Let $A \in \mathbf{R}^{m \times n}$ have rows a_i^T, i.e., $A^T = (a_1, \ldots, a_m)$. Show that

$$A^T A = \sum_{i=1}^{m} a_i a_i^T.$$

If A is instead partitioned into columns, what is the corresponding expression for $A^T A$?

A.2. (a) Show that if A and B are square upper triangular matrices, then AB is upper triangular, and that A^{-1} is upper triangular, if it exists. Is the same true for lower triangular matrices?

(b) Let $A, B \in \mathbf{R}^{n \times n}$ have lower bandwidth r and s, respectively. Show that the product AB has lower bandwidth $r + s$.

(c) An upper Hessenberg matrix H is a matrix with lower bandwidth $r = 1$. Using the result in (a) deduce that the product of H and an upper triangular matrix is again an upper Hessenberg matrix.

(d) Show that if $R \in \mathbf{R}^{n \times n}$ is strictly upper triangular, then $R^n = 0$.

A.3. Use row operations to verify that the Vandermonde determinant is

$$\det \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{pmatrix} = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2).$$

A.4. To solve a linear system $Ax = b$, $A \in \mathbf{R}^{n \times n}$, by Cramer's rule (A.3.6) requires the evaluation of $n + 1$ determinants of order n. Estimate the number of multiplications needed for $n = 50$ if the determinants are evaluated in the naive way. Estimate the time it will take on a computer performing 10^9 floating point operations per second!

A.5. Consider an upper block triangular matrix,

$$R = \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix},$$

and suppose that R_{11} and R_{22} are nonsingular. Show that R is nonsingular and give an expression for R^{-1} in terms of its blocks.

A.6. (a) Show that if $w \in \mathbf{R}^n$ and $w^T w = 1$, then the matrix $P(w) = I - 2ww^T$ is both symmetric and orthogonal.

(b) Let $x, y \in \mathbf{R}^n$, $x \neq y$, be two given vectors with $\|x\|_2 = \|y\|_2$. Show that $P(w)x = y$ if $w = (y - x)/\|y - x\|_2$.

A.7. Show that for any matrix norm there exists a consistent vector norm.

Hint: Take $\|x\| = \|xy^T\|$ for any vector $y \in \mathbf{R}^n$, $y \neq 0$.

A.8. Derive the formula for $\|A\|_\infty$ given in (A.4.14).

A.9. Show that $\|A\|_2 = \|PAQ\|_2$ if $A \in \mathbf{R}^{m \times n}$ and P and Q are orthogonal matrices of appropriate dimensions.

A.10. Use the result $\|A\|_2^2 = \rho(A^T A) \leq \|A^T A\|$, valid for any matrix operator norm $\| \cdot \|$, where $\rho(A^T A)$ denotes the spectral radius of $A^T A$, to deduce the upper bound in (A.4.16).

A.11. (a) Let T be a nonsingular matrix, and let $\| \cdot \|$ be a given vector norm. Show that the function $N(x) = \|Tx\|$ is a vector norm.

(b) What is the matrix norm subordinate to $N(x)$?

(c) If $N(x) = \max_i |k_i x_i|$, what is the subordinate matrix norm?