Basic Principles of Numerical Analysis, Matrix Theory and Functional Analysis

Compiled by

R. Albanese
Ass. EURATOM/ENEA/CREATE, DIMET, Univ.Mediterranea di Reggio Calabria,
Via Graziella, Loc. Feo di Vito, I-89133, Reggio Calabria, Italy

Sources:

- Wikipedia, the free encyclopedia (http://en.wikipedia.org)
- Matrices and Linear Algebra (http://www.mathworks.com/access/helpdesk/help/techdoc/math/mat_linalg.html)
- Lectures of course "Modelli Numerici per Campi e Circuiti" (Computational Methods in Electrical Engineering) given by R. Albanese in 1995-96 at the Università degli Studi di Reggio Calabria, Italy
- PlanetMath (http://planetmath.org/encyclopedia)

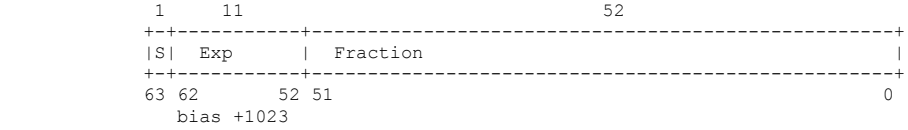
Floating-point numbers

MATLAB uses conventional decimal notation, with an optional decimal point and leading plus or minus sign, for numbers. Scientific notation uses the letter e to specify a power-of-ten scale factor. Imaginary numbers use either i or j as a suffix. Some examples of legal numbers are

3 -99 0.0001
9.6397238 1.60210e-20 6.02252e23
1i -3.14159j 3e5i

All numbers are stored internally using the long format specified by the IEEE floating-point standard. Floating-point numbers have a finite precision of roughly 16 significant decimal digits and a finite range of roughly 10^{-308} to 10^{+308}.

IEEE floating-point standard double-precision 64 bit



For normalised numbers, the most common, S is the sign, Exp is the biased exponent and Fraction is the fractional part of the significand. The number has value $v = s \times 2^e \times m$ where
$s = +1$ (positive numbers) when S is 0
$s = -1$ (negative numbers) when S is 1
e = Exp − 1023 (in other words the exponent is stored with 1023 added to it, also called "biased with 1023")
m = 1.Fraction in binary (that is, the significand is the binary number 1 followed by the radix point followed by the binary bits of Fraction). Therefore, $1 <= m < 2$.

eps

Floating-point relative accuracy, i.e., the distance from 1.0 to the next largest floating-point number.
In MATLAB eps = 2^(-52), which is roughly 2.22e-16.

realmax

Largest floating-point number your computer can represent.
In MATLAB realmax is one bit less than 21024 or about 1.7977e+308

realmin

Smallest floating-point number your computer can represent.
On machines with IEEE floating-point format, realmin is 2^(-1022) or about 2.2251e-308.

NOTICE:

1) CRAMER'S RULE

Cramer's rule it is of theoretical importance in that it gives an explicit expression for the solution of the system $Ax = c$ where the *n-by-n* square matrix A is invertible and the vector x is the column vector of the variables (x_i):

$$x_i = \frac{\det(A_i)}{\det(A)}$$

where A_i is the matrix formed by replacing the ith column of A by the column vector c.
However, Cramer's rule is generally inefficient and thus not used in practical applications which may involve many equations. The computation of a determinant starting from its definition requires a number of floating point operation proportional to $n!$. Since:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

and the most powerful computer nowadays are rated around hundred of MFLOPS, the simple calculation of a determinant would take billion years for a *30-by-30* matrix. Of course there are faster ways to compute determinants and inverse matrices.

2) ASSOCIATIVITY PROPERTY

Associativity, i.e. *(a+b)+c=a+(b+c)*, does not generally hold on a computer. For instance, if *a=10, b=-10* and *c=2*eps*, we have *(a+b)+c=2*eps*, whereas *a+(b+c)*=0.

3) SCALING

State space is the preferred model for LTI systems, especially with higher order models. Even with state-space models, however, accurate results are not guaranteed, because of the finite-word-length arithmetic of the computer. A well-conditioned problem is usually a prerequisite for obtaining accurate results.

You should generally normalize or scale the (A, B, C, D) matrices of a system to improve their conditioning. An example of a poorly scaled problem might be a dynamic system where two states in the state vector have units of light years and millimeters. You would expect the A matrix to contain both very large and very small numbers. Matrices containing numbers widely spread in value are often poorly conditioned both with respect to inversion and with respect to their eigenproblems, and inaccurate results can ensue.

Normalization also allows meaningful statements to be made about the degree of controllability and observability of the various inputs and outputs.

A set of (A, B, C, D) matrices can be normalized using diagonal scaling matrices N_u, N_x, and N_y to scale *u*, *x*, and *y*.

$$u = N_u u_n \qquad x = N_x x_n \qquad y = N_y y_n$$

so the normalized system is

$$\dot{x}_n = A_n x_n + B_n u_n$$
$$y_n = C_n x_n + D_n u_n$$

where

$$A_n = N_x^{-1} A N_x \qquad B_n = N_x^{-1} B N_u$$
$$C_n = N_y^{-1} C N_x \qquad D_n = N_y^{-1} D N_u$$

Choose the diagonal scaling matrices according to some appropriate normalization procedure. One criterion is to choose the maximum range of each of the input, state, and output variables. This method originated in the days of analog simulation computers when u_n, x_n, and y_n were forced to be between ± 10 Volts. A second method is to form scaling matrices where the diagonal entries are the smallest deviations that are significant to each variable. An excellent discussion of scaling is given in the introduction to the *LINPACK Users' Guide, [1].*

Choose scaling based upon physical insight to the problem at hand. If you choose not to scale, and for many small problems scaling is not necessary, be aware that this choice affects the accuracy of your answers.

Vector Space

A set V is a **vector space** over a field F (for example, the field of real or of complex numbers) if, given

- an operation *vector addition* defined in V, denoted $\mathbf{v} + \mathbf{w}$ (where $\mathbf{v}, \mathbf{w} \in V$), and
- an operation *scalar multiplication* in V, denoted $a * \mathbf{v}$ (where $\mathbf{v} \in$ V and $a \in$ F),

the following ten properties hold for all $a, b \in$ F and $\mathbf{u}, \mathbf{v}$, and $\mathbf{w} \in$ V:

1. $\mathbf{v} + \mathbf{w}$ belongs to V.
   (Closure of V under vector addition.)
2. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$.
   (Associativity of vector addition in V.)
3. There exists a neutral element $\mathbf{0}$ in V, such that for all elements $\mathbf{v}$ in V, $\mathbf{v} + \mathbf{0} = \mathbf{v}$.
   (Existence of an additive identity element in V.)
4. For all $\mathbf{v}$ in V, there exists an element $\mathbf{w}$ in V, such that $\mathbf{v} + \mathbf{w} = \mathbf{0}$.
   (Existence of additive inverses in V.)
5. $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$.
   (Commutativity of vector addition in V.)
6. $a * \mathbf{v}$ belongs to V.
   (Closure of V under scalar multiplication.)
7. $a * (b * \mathbf{v}) = (ab) * \mathbf{v}$.
   (Associativity of scalar multiplication in V.)
8. If 1 denotes the multiplicative identity of the field F, then $1 * \mathbf{v} = \mathbf{v}$.
   (Neutrality of one.)
9. $a * (\mathbf{v} + \mathbf{w}) = a * \mathbf{v} + a * \mathbf{w}$.
   (Distributivity with respect to vector addition.)
10. $(a + b) * \mathbf{v} = a * \mathbf{v} + b * \mathbf{v}$.
    (Distributivity with respect to field addition.)

Properties 1 through 5 indicate that V is an abelian group under vector addition. The rest, properties 6 through 10, apply to scalar multiplication of a vector $\mathbf{v} \in$ V by a scalar $a \in$ F. Note that property 5 actually follows from the other 9.

Basis (linear algebra)

A subset B of a vector space V is said to be a **basis** of V if it satisfies one of the four equivalent conditions:

1. B is both a set of linearly independent vectors and a generating set of V.
2. B is a minimal generating set of V, i.e. it is a generating set but no proper subset of B is.
3. B is a maximal set of linearly independent vectors, i.e. it is a linearly independent set but no proper superset is.
4. every vector in V can be expressed as a linear combination of vectors in B in a unique way.

Recall that a set B is a generating set of V if every vector in V is a linear combination of vectors in B. This definition includes a finiteness condition: a linear combination is always a ***finite*** sum of the form $a_1v_1 + ... + a_nv_n$.

All bases of a vector space have the same cardinality (number of elements), called the dimension of the vector space.

In these definitions the fact that all linear combinations are *finite* is crucial. A set B is a basis of a vector space V if every member of V is a linear combination of just *finitely* many members of B. However, in Hilbert spaces and other Banach spaces, one often considers linear combinations of infinitely many vectors. In an infinite-dimensional Hilbert space, a set of vectors orthogonal to each other can never span the whole space via finite linear combinations, but what is called an orthonormal basis is a set of mutually orthogonal unit vectors that "span" the space via sometimes-infinite linear combinations. More generally, in topological vector spaces, one may define *infinite sums* (or series) and express elements of the space as *infinite linear combinations* of other elements.

Normed vector space

Definition

If V is a vector space over a field K (which must be either the real or complex numbers or another field of characteristic zero), a norm on V is a function from V to $\mathbf{R}$, the real numbers — that is, it associates to each vector $\mathbf{v}$ in V a real number, which is usually denoted $\|\mathbf{v}\|$. The norm must satisfy the following conditions:

For all a in K and all $\mathbf{u}$ and $\mathbf{v}$ in V,
1. $\|\mathbf{v}\| \geq 0$ with equality if and only if $\mathbf{v} = \mathbf{0}$.
2. $\|a\mathbf{v}\| = |a| \, \|\mathbf{v}\|$.
3. $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Most of property 1 follows from the other axioms, and in fact it can be replaced by the following condition:

1'. if $\|\mathbf{v}\| = 0$, then $\mathbf{v} = \mathbf{0}$

A useful consequence of the norm axioms is the inequality

$$\|\mathbf{u} \pm \mathbf{v}\| \geq | \, \|\mathbf{u}\| - \|\mathbf{v}\| \, |$$

for all vectors $\mathbf{u}$ and $\mathbf{v}$.

Distances in normed vector spaces

For any normed vector space we can define the *distance* between two vectors $\mathbf{u}$ and $\mathbf{v}$ as $\|\mathbf{u}\text{-}\mathbf{v}\|$. (Note that the Euclidean norm gives rise to the Euclidean distance in this fashion.) This turns the normed space into a metric space and allows the definition of notions such as continuity and convergence. The norm is then a continuous map. If this metric space is complete then the normed space is called a Banach space. Every normed vector space V sits as a dense subspace inside a Banach space; this Banach space is essentially uniquely defined by V and is called the *completion* of V.

Two norms $\|.\|_1$ and $\|.\|_2$ on a vector space V are called *equivalent* if there exist positive real numbers C and D such that

$$C\|x\|_1 \leq \|x\|_2 \leq D\|x\|_1$$

for all x in V. In this case, the two norms define the same notions of continuity and convergence and do not need to be distinguished for most purposes.

Finite-dimensional normed vector spaces

All norms on a finite-dimensional vector space V are equivalent. Since Euclidean space is complete, we can thus conclude that all finite-dimensional normed vector spaces are Banach spaces.

A normed vector space V is finite-dimensional if and only if the unit ball $B = \{x : \|x\| = 1\}$ is compact, which is the case if and only if V is locally compact.

Examples of norms

Euclidean norm

On $\mathbf{R}^n$, the intuitive notion of length of the vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is captured by the formula

$$\|x\| = \sqrt{|x_1|^2 + \cdots + |x_n|^2}.$$

This gives the ordinary distance from the origin to the point $\mathbf{x}$, a consequence of the Pythagorean theorem. The Euclidean norm is by far the most commonly used norm on $\mathbf{R}^n$, but there are other norms on this vector space as will be shown below.

Taxicab norm or Manhattan norm

Main article Taxicab geometry

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|.$$

The name comes from the fact that the norm gives the distance a taxi has to drive in a rectangular street grid to get from the origin to the point x.

p-norm

Let p=1 be a real number.

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$$

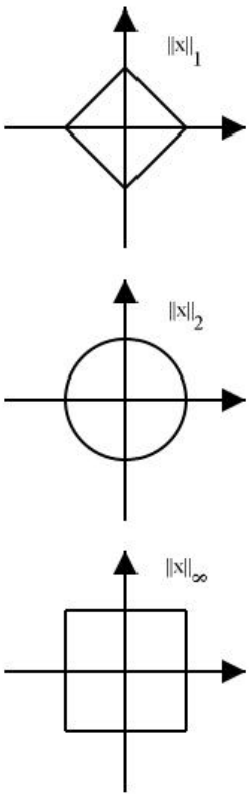
Note that for p=1 we get the taxicab norm and for p=2 we get the Euclidean norm. See also L^p space.

Infinity norm or maximum norm

Main article maximum norm

$$\|x\|_\infty = \max\left(|x_1|, \ldots, |x_n|\right).$$

The concept of unit circle (the set of all vectors of norm 1) is different in different norms: for the 1-norm the unit circle in $\mathbf{R}^2$ is a rhomboid, for the 2-norm (Euclidean norm) it is the well-known unit circle, while for the infinity norm it is a square. See the accompanying illustration.



Illustrations of unit circles in different norms.

Other norms

Other norms on $\mathbf{R}^n$ can be constructed by combining the above; for example

$$\|x\| = 2|x_1| + \sqrt{3|x_2|^2 + \max(|x_3|, 2|x_4|)^2}$$

is a norm on $\mathbf{R}^4$.

All the above formulas also yield norms on $\mathbf{C}^n$ without modification.

Examples of infinite dimensional normed vector spaces can be found in the Banach space article. In addition, inner product space becomes a normed vector space if we define the norm as

$$\|x\| = \sqrt{< x, x >}.$$

Another common norm in R^n is:

$$\|x\|_A = \sqrt{x^T A x}$$

where A is a symmetric positive definite matrix.

Examples of functional spaces

When a variational formulation is introduced, the field problem reduces to give the minimum (or the stationary point) of a functional in a proper functional space. The fields **E**, **H**, **B**, **D**, and **J** should belong to $\mathbf{L}^2(V)$, the space of the square integrable vector fields over V, because the associated energy and power are finite. Moreover, according to the previous discussion, the continuity of the tangential component of **E**, **H**, **T**, and **A**, as well as the normal component of **B**, **D**, and **J**, is required at any material interface. The normal components of **E**, **H**, **T**, and **A** are free to jump as well as the tangential component of **B**, **D**, and **J**. On the other hand, the scalar potentials Ω and ϕ should be continuous. These properties can be formally stated as follows:

$$\mathbf{E}, \mathbf{H}, \mathbf{A}, \mathbf{T} \in \mathbf{L}^2_{\text{rot}}(V) = \{\mathbf{W} \in \mathbf{L}^2(V), \nabla \times \mathbf{W} \in \mathbf{L}^2(V)\} \quad , \qquad (50)$$

$$\mathbf{B}, \mathbf{D}, \mathbf{J} \in \mathbf{L}^2_{\text{div}}(V) = \{\mathbf{W} \in \mathbf{L}^2(V), \nabla \cdot \mathbf{W} \in L^2(V)\} \qquad (51)$$

$$\Omega, \phi \in L^2_{\text{grad}}(V) = \{w \in L^2(V), \nabla w \in \mathbf{L}^2(V)\} \qquad (52)$$

Matrix norm

- A **sub-multiplicative vector norm** is any vector norm on square matrices compatible with matrix multiplication in the sense that

$$\|AB\| \leq \|A\|\|B\|$$

Operator norm or induced norm

If norms on K^m and K^n are given (K is real or complex), then one defines the corresponding *induced norm* or *operator norm* on the space of m-by-n matrices as the following suprema:

$$\begin{aligned}\|A\| &= \sup\{\|Ax\| : x \in K^n \text{ with } \|x\| \leq 1\} \\ &= \sup\{\|Ax\| : x \in K^n \text{ with } \|x\| = 1\} \\ &= \sup\left\{\frac{\|Ax\|}{\|x\|} : x \in K^n \text{ with } x \neq 0\right\}\end{aligned}$$

If $m = n$ and one uses the same norm on domain and range, then these operator norms are all (submultiplicative) matrix norms.

An induced norm is consistent with the vector norm $\varphi(x)$ in the sense that $\varphi(Ax) \leq ||A||\ \varphi(x)$.

A submultiplicative norm $||A||$ is consistent with the vector norm $\varphi(x) = ||[x\ 0\ \ldots\ 0]||$.

Spectral norm or spectral radius

If $m=n$ and the norm on K^n is the Euclidean norm, then the induced matrix norm is the *spectral norm*.

Spectral norm is the only minimal matrix norm which is an induced norm. The spectral norm of A equals to the square root of the spectral radius of AA_* or the largest singular value of A.

An important property for matrix norm is

$$\lim_{r\to\infty} \|A^r\|^{1/r} = \rho(A)$$

where $\rho(A)$ is the spectral radius of A.

The spectral radius has also the following property (Householder's theorem): $\rho(A) = inf\ ||A||$ where $||A||$ is the set of submultiplicative norms of A.

Matrix (mathematics)

From Wikipedia, the free encyclopedia. http://en.wikipedia.org

In mathematics, a **matrix** (plural **matrices**) is a rectangular table of numbers or, more generally, of elements of a fixed ring.

The horizontal lines in a matrix are called **rows** and the vertical lines are called **columns**. A matrix with m rows and n columns is called an m-by-n matrix (or $m \times n$ matrix) and m and n are called its **dimensions**.

The entry of a matrix A that lies in the ith row and the j-th column is called the i,j-entry or (i,j)th entry of A. This is written as $A[i,j]$ or $A_{i,j}$, or in notation of the C programming language, `A[i][j]`.

If a matrix A and a number c are given, we may define the **scalar multiplication** cA by $(cA)[i, j] = cA[i, j]$. For example

$$2\begin{bmatrix} 1 & 8 & -3 \\ 4 & -2 & 5 \end{bmatrix} = \begin{bmatrix} 2 \times 1 & 2 \times 8 & 2 \times -3 \\ 2 \times 4 & 2 \times -2 & 2 \times 5 \end{bmatrix} = \begin{bmatrix} 2 & 16 & -6 \\ 8 & -4 & 10 \end{bmatrix}$$

If a matrix A and a number c are given, we may define the **scalar multiplication** cA by $(cA)[i, j] = cA[i, j]$. For example

$$2\begin{bmatrix} 1 & 8 & -3 \\ 4 & -2 & 5 \end{bmatrix} = \begin{bmatrix} 2 \times 1 & 2 \times 8 & 2 \times -3 \\ 2 \times 4 & 2 \times -2 & 2 \times 5 \end{bmatrix} = \begin{bmatrix} 2 & 16 & -6 \\ 8 & -4 & 10 \end{bmatrix}$$

Multiplication of two matrices is well-defined only if the number of columns of the first matrix is the same as the number of rows of the second matrix. If A is an m-by-n matrix (m rows, n columns) and B is an n-by-p matrix (n rows, p columns), then their **product** AB is the m-by-p matrix (m rows, p columns) given by

$(AB)[i, j] = A[i, 1] * B[1, j] + A[i, 2] * B[2, j] + ... + A[i, n] * B[n, j]$ for each pair i and j.

For instance

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

This multiplication has the following properties:

- $(AB)C = A(BC)$ for all k-by-m matrices A, m-by-n matrices B and n-by-p matrices C ("associativity").
- $(A + B)C = AC + BC$ for all m-by-n matrices A and B and n-by-k matrices C ("distributivity").
- $C(A + B) = CA + CB$ for all m-by-n matrices A and B and k-by-m matrices C ("distributivity").

It is important to note that commutativity does ***not*** generally hold; that is, given matrices A and B and their product defined, then generally $AB \neq BA$.

Linear transformations, ranks and transpose

Matrices can conveniently represent linear transformations because matrix multiplication neatly corresponds to the composition of maps, as will be described next.

Here and in the sequel we identify $\mathbf{R}^n$ with the set of "rows" or n-by-1 matrices. For every linear map $f : \mathbf{R}^n \to \mathbf{R}^m$ there exists a unique m-by-n matrix A such that $f(x) = Ax$ for all x in $\mathbf{R}^n$. We say that the matrix A "represents" the linear map f. Now if the k-by-m matrix B represents another linear map $g : \mathbf{R}^m \to \mathbf{R}^k$, then the linear map $g \circ f$ is represented by BA. This follows from the above-mentioned associativity of matrix multiplication.

The rank of a matrix A is the dimension of the image of the linear map represented by A; this is the same as the dimension of the space generated by the rows of A, and also the same as the dimension of the space generated by the columns of A.

The transpose of an m-by-n matrix A is the n-by-m matrix A^{tr} (also sometimes written as A^T or ${}^tA$) gotten by turning rows into columns and columns into rows, i.e. $A^{tr}[i, j] = A[j, i]$ for all indices i and j. If A describes a linear map with respect to two bases, then the matrix A^{tr} describes the transpose of the linear map with respect to the dual bases, see dual space.

We have $(A + B)^{tr} = A^{tr} + B^{tr}$ and $(AB)^{tr} = B^{tr} * A^{tr}$.

Square matrices and related definitions

A **square matrix** is a matrix which has the same number of rows as columns. The set of all square n-by-n matrices, together with matrix addition and matrix multiplication is a ring. Unless $n = 1$, this ring is not commutative.

M(n, **R**) , the ring of real square matrices, is a real unitary associative algebra. M(n, **C**), the ring of complex square matrices, is a complex associative algebra.

The **unit matrix** or **identity matrix** I_n, with elements on the main diagonal set to 1 and all other elements set to 0, satisfies $MI_n=M$ and $I_nN=N$ for any m-by-n matrix M and n-by-k matrix N. For example, if $n = 3$:

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The identity matrix is the identity element in the ring of square matrices.

Invertible elements in this ring are called **invertible matrices** or **non-singular matrices**. An n by n matrix A is invertible if and only if there exists a matrix B such that

$$AB = \mathrm{I}_n \ (= BA).$$

In this case, B is the **inverse matrix** of A, denoted by A^{-1}. The set of all invertible n-by-n matrices forms a group (specifically a Lie group) under matrix multiplication, the general linear group.

If λ is a number and $\mathbf{v}$ is a non-zero vector such that $A\mathbf{v} = \lambda\mathbf{v}$, then we call $\mathbf{v}$ an eigenvector of A and λ the associated eigenvalue. (Eigen means "own" in German.) The number λ is an eigenvalue of A if and only if $A-\lambda I_n$ is not invertible, which happens if and only if $p_A(\lambda) = 0$. Here $p_A(x)$ is the characteristic polynomial of A. This is a polynomial of degree n and has therefore n complex roots (counting multiple roots according to their multiplicity). In this sense, every square matrix has n complex eigenvalues.

The determinant of a square matrix A is the product of its n eigenvalues, but it can also be defined by the *Leibniz formula*. Invertible matrices are precisely those matrices with nonzero determinant.

The Gauss-Jordan elimination algorithm is of central importance: it can be used to compute determinants, ranks and inverses of matrices and to solve systems of linear equations.

The trace of a square matrix is the sum of its diagonal entries, which equals the sum of its n eigenvalues.

Every orthogonal matrix is a square matrix.

Submatrix

From Wikipedia, the free encyclopedia.

A **submatrix** is a matrix formed by taking certain rows and columns from a bigger matrix.

For example

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

Then

$$A[1,2;1,3,4] = \begin{bmatrix} a_{11} & a_{13} & a_{14} \\ a_{21} & a_{23} & a_{24} \end{bmatrix}$$

is a submatrix of A formed by rows 1,2 and columns 1,3,4. This submatrix can also be denoted by $A(3;2)$ which means that it is formed by *deleting* row 3 and column 2.

Triangular matrix

A matrix

$$\mathbf{L} = \begin{bmatrix} l_{1,1} & & & & 0 \\ l_{2,1} & l_{2,2} & & & \\ l_{3,1} & l_{3,2} & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \\ l_{n,1} & l_{n,2} & \dots & l_{n,n-1} & l_{n,n} \end{bmatrix}$$

is called **lower triangular matrix** or **left triangular matrix**, and analogously a matrix of the form

$$\mathbf{U} = \begin{bmatrix} u_{1,1} & u_{1,2} & u_{1,3} & \dots & u_{1,n} \\ & u_{2,2} & u_{2,3} & \dots & u_{2,n} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & u_{n-1,n} \\ 0 & & & & u_{n,n} \end{bmatrix}$$

is called **upper triangular matrix** or **right triangular matrix**.

The identity matrix is a normed upper and lower triangular matrix.

The product of two upper triangular matrices is upper triangular, so the set of upper triangular matrices forms an algebra. Algebras of upper triangular matrices have a natural generalisation in functional analysis which yields nest algebras.

The transpose of a upper triangular matrix is a lower triangular matrix and vice versa.

A matrix equation in the form

$$\mathbf{Lx} = \mathbf{b}$$

or

$$\mathbf{Ux} = \mathbf{b}$$

is very easy to solve. The matrix equation *Lx*= *b* can be written as a system of linear equations

$$\begin{array}{ccccccc} x_1 & & & & & = & b_1 \\ l_{2,1}x_1 & + & x_2 & & & = & b_2 \\ \vdots & & \vdots & \ddots & & & \vdots \\ l_{m,1}x_1 & + & l_{m,2}x_2 & + \ldots + & x_m & = & b_m \end{array}$$

which can be solved by the following recursive relation

$$\begin{array}{rcl} x_1 & = & b_1 \\ x_2 & = & b_2 - l_{2,1}b_1 \\ & \vdots & \\ x_m & = & b_m - \sum_{i=1}^{m-1} l_{m,i}x_i \end{array} \quad ;$$

Rank

In linear algebra, the column rank (row rank respectively) of a matrix A with entries in some field is defined to be the maximal number of columns (rows respectively) of A which are linearly independent.

Given an m×n matrix) and rank r, then there exists at least one non-zero r×r minor, while all larger minors are zero (a minor of a matrix is the determinant of a submatrix of its).

Diagonal matrix

In linear algebra, a **diagonal matrix** is a square matrix in which only the entries in the main diagonal are non-zero. The diagonal entries themselves may or may not be zero. Thus, the matrix D = ($d_{i,j}$) with n columns and n rows is diagonal if:

$$d_{i,j} = 0 \text{ if } i \neq j \qquad \forall i,j \in \{1,2,\ldots,n\}$$

For example, the following matrix is diagonal:

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

Any diagonal matrix is also a symmetric matrix, a triangular matrix, and (if the entries come from the field **R** or **C**) also a normal matrix. The identity matrix I_n is diagonal.

A diagonal matrix with all its main diagonal entries equal is a **scalar matrix**, that is, a scalar multiple λI of the identity matrix I. Its effect on a vector is scalar multiplication by λ. The scalar matrices are the center of the algebra of matrices: that is, they are precisely the matrices that commute with all other square matrices of the same size.

Sparse matrix

From Wikipedia, the free encyclopedia.

In the mathematical subfield of numerical analysis a **sparse matrix** is a matrix populated primarily with zeros. A **sparse graph** is a graph with a sparse adjacency matrix.

Sparsity is a concept, useful in combinatorial mathematics and application areas such as network theory, of a low density of significant data or connections. This concept is amenable to quantitative reasoning. It is also noticeable in everyday life.

Huge sparse matrices often appear in science or engineering when solving problems for linear models.

When storing and manipulating sparse matrices on the computer it is often neccessary to modify the standard algorithms and take advantage of the sparse structure of the matrix. Sparse data is by its nature more easily compressed which can yield enormous savings in memory usage. And more importantly manipulating huge sparse matrices with the standard algorithms may be impossible due to their sheer size. The definition of huge depends on the hardware and the computer programs available to manipulate the matrix.

Given a sparse $N \times M$ matrix A the **row bandwidth** for the n-th row is defined as

$$b_n(\mathbf{A}) := \min_{1 \leq m \leq M} \{m \mid a_{n,m} \neq 0\}$$

The **bandwidth** for the matrix is defined as

$$B(\mathbf{A}) := \max_{1 \leq n \leq N} b_n(\mathbf{A})$$

The Cuthill-McKee_algorithm can be used to reduce the bandwith of a sparse symmetric matrix.

The inverse of a sparse matrix is generally full. The LU factorization of a sparse matrix **A** provides two triangular matrices **L** and **U** having the same bandwidth as **A**.

In MATLAB:

S = SPARSE(i,j,s,m,n,nzmax)
uses the rows of [i,j,s] to generate an m-by-n sparse matrix with space allocated for nzmax nonzeros. The two integer index vectors, i and j, and the real or complex entries vector, s, all have the same length, nnz, which is the number of nonzeros in the resulting sparse matrix S . Any elements of s which have duplicate values of i and j are added together.

To dissect and then reassemble a sparse matrix:

```
[i,j,s] = find(S);
[m,n] = size(S);
S = sparse(i,j,s,m,n);
```

S = SPARSE(X) converts a sparse or full matrix to sparse form by squeezing out any zero elements.

A = FULL(X) converts a sparse matrix S to full storage organization. If X is a full matrix, it is left unchanged.

All of MATLAB's built-in arithmetic, logical and indexing operations can be applied to sparse matrices, or to mixtures of sparse and full matrices.

Determinant

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} A_{i,\sigma(i)}$$

The sum is computed over all permutations s of the numbers $\{1, 2, \ldots, n\}$ and sgn(s) denotes the signature of the permutation s: +1 if s is an even permutation and -1 if it is odd. See symmetric group for an explanation of even/odd permutations.

This formula contains $n!$ summands and is therefore impractical to use to calculate determinants for large n.

In general, determinants can be computed with the Gauss algorithm using the following rules:

- If A is a triangular matrix, i.e. $A_{i,j} = 0$ whenver $i > j$, then $\det(A) = A_{1,1} A_{2,2} \ldots A_{n,n}$
- If B results from A by exchanging two rows or columns, then $\det(B) = -\det(A)$
- If B results from A by multiplying one row or column with the number c, then $\det(B) = c \det(A)$
- If B results from A by adding a multiple of one row or column to another row or column, then $\det(B) = \det(A)$.

Explicitly, starting out with some matrix, use the last three rules to convert it into a triangular matrix, then use the first rule to compute its determinant.

It is also possible to expand a determinant along a row or column using *Laplace's formula*, which is efficient for relatively small matrices. To do this along row i, say, we write

$$\det(A) = \sum_{j=1}^{n} A_{i,j} C_{i,j}$$

where the $C_{i,j}$ represent the matrix cofactors, i.e. $C_{i,j}$ is $(-1)^{i+j}$ times the determinant of the matrix that results from A by removing the i-th row and the j-th column.

The determinant is a *multiplicative map* in the sense that

$\det(AB) = \det(A)\det(B)$ for all $n \times n$ matrices A and B.

This is generalized by the Cauchy-Binet formula to products of non-square matrices.

It is easy to see that $\det(rI^n) = r^n$ and thus

$\det(rA) = r^n \det(A)$ for all n-by-n matrices A and all scalars r.

If A is invertible, then

$\det(A^{-1}) = \det(A)^{-1}$.

A matrix and its transpose have the same determinant:

$\det(A) = \det(A^T)$.

If A and B are similar, i.e. if there exists an invertible matrix X such that $A = X^{-1}BX$, then by the multiplicative property,

$\det(A) = \det(B)$.

This means that the determinant is a similarity invariant. Because of this, the determinant of some linear transformation $T : V \to V$ for some finite dimensional vector space V, is independent of the basis for V. The relationship is one-way, however, there exist matrices which have the same determinant but are not similar.

If A is a square n-by-n matrix with real or complex entries and if $\lambda_1, \ldots, \lambda_n$ are the (complex) eigenvalues of A listed according to their algebraic multiplicities, then

$$\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$$

This follows from the fact that A is always similar to its Jordan normal form, an upper triangular matrix with the eigenvalues on the main diagonal.

Matrix Inversion

Matrix inversion is the following problem in linear algebra: given a square n-by-n matrix A, find a square n-by-n matrix B (if one exists) such that $AB = BA = I_n$, the n-by-n identity matrix.

The Gauss-Jordan elimination is an algorithm that can be used to determine whether a given matrix is invertible and to find the inverse. An alternative is the Cholesky decomposition which generates two upper triangular matrices which are easier to invert. For special purposes, it may be convenient to invert matrices by treating mn-by-mn matrices as m-by-m matrices of n-by-n matrices, and applying one or another formula recursively (other sized matrices can be padded out with dummy rows and columns). For other purposes, a variant of Newton's method may be convenient (particularly when dealing with families of related matrices, so inverses of earlier matrices can be used to seed generating inverses of later matrices).

Writing another special matrix of cofactors, known as an adjoint matrix, can also be an efficient way to calculate the inverse of *small* matrices (since this method is essentially recursive, it becomes inefficient for large matrices). To determine the inverse, we calculate a matrix of cofactors:

$$A^{-1} = \frac{1}{|A|}(C_{ij})^T = \frac{1}{|A|}\begin{pmatrix} C_{11} & C_{21} & \cdots & C_{j1} \\ C_{12} & \ddots & \vdots & C_{j2} \\ \vdots & \cdots & \ddots & \vdots \\ C_{1i} & \cdots & \cdots & C_{ji} \end{pmatrix}$$

where $|A|$ is the determinant of A, C_{ij} is the matrix cofactor, and A^T represents the matrix transpose.

In most practical applications, it is in fact not necessary to invert a matrix, but only to solve a system of linear equations. Various fast algorithms for special classes of such systems have been developed.

The cofactor C_{ij} of A is defined as $(-1)^{i+j}$ times the minor M_{ij} of A.

Let A be a square n by n matrix over a field K (for example the field R of real numbers). The following statements are equivalent and must all be true for A to be invertible:

- det A ≠ 0
- rank A = n
- The equation Ax = 0 has only the trivial solution x = 0 (i.e. Nul A = {0}).
- The equation Ax = b has at most one solution for each b in K^n
- The equation Ax = b has at least one solution for each b in K^n
- The equation Ax = b has exactly one solution for each b in K^n
- The columns of A are linearly independent.
- The columns of A span K^n (i.e. Col A = K^n)
- The columns of A form a basis of K^n
- The linear transformation x |-> Ax from K^n to K^n is one-to-one
- The linear transformation x |-> Ax from K^n to K^n is onto
- The linear transformation x |-> Ax from K^n to K^n is bijective
- There is an n by n matrix B such that $BA = I_n$
- There is an n by n matrix B such that $AB = I_n$
- The transpose A^T is an invertible matrix.
- The number 0 is not an eigenvalue of A

The inverse of an invertible matrix A is itself invertible, with

$$(A^{-1})^{-1} = A.$$

The product of two invertible matrices A and B of the same size is invertible, with the inverse given by

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

Eigenvalues

In linear algebra, a scalar λ is called an **eigenvalue** (in some older texts, a **characteristic value**) of a linear mapping A if there exists a nonzero vector x such that $Ax=\lambda x$. The vector x is called an eigenvector.

In MATLAB [V,D] = EIG(X) produces a diagonal matrix D of eigenvalues and a full matrix V whose columns are the corresponding eigenvectors so that X*V = V*D.

The characteristic polynomial of A, denoted by $p_A(t)$, is the polynomial defined by $p_A(t) = \det(tI - A)$ where I denotes the n-by-n identity matrix. Due to Rouché-Capelli theorem, an eigenvalue of A is a root of p_A.

The **algebraic multiplicity** (or simply **multiplicity**) of an eigenvalue λ of A is the number of factors t-λ of the characteristic polynomial of A. The **geometric multiplicity** of λ is the number of factors t-λ of the minimal polynomial of A or equivalently the nullity of (λI-A). the nullity of a matrix M is the number of columns of M minus the rank of M.

Notice that A and A^T have same the eigenvalues but in general different eigenvectors.

In case both multiplications that A and A^T have same the eigenvalues but in general different eigenvectors.

Given matrices A and B and their product defined, then generally $AB \neq BA$. However, AB and BA have the same nonzero eigenvalues. The eigenvectors of BA are obtained multiplying by B those of AB.

Cayley-Hamilton theorem states that every square matrix over the real or complex field, satisfies its own characteristic equation. This means the following: if A is the given square matrix and $p_A(t) = det(A - tI)$ is its characteristic polynomial (a polynomial in the variable t), then replacing t by the matrix A results in the zero matrix: $p_A(A) = 0$. An important corollary of the Cayley-Hamilton theorem is that the minimal polynomial of a given matrix is a divisor of its characteristic polynomial. This is very useful in finding the Jordan form of a matrix.

If $p(t)$ is any polynomial in the variable t and λ is an eigenvalue of matrix A, then $p(\lambda)$ is an eigenvalue of matrix $p(A)$. In particular, λ^n is an eigenvalue of A^n, and (σ+μλ) is an eigenvalue of (σI+μA).

Real symmetric matrices have real eigenvalues, as $\lambda=(x*Ax)/(x*x)$ is the ratio of two real numbers. On the other hand, skew-symmetric matrices have imaginary eigenvalues.

Gerschgorin's Theorem

Let A be a square complex matrix. Around every element a_{ij} on the diagonal of the matrix, we draw a circle with radius the sum of the norms of the other elements on the same row $\sum_{j\neq i}|a_{ij}|$. Such circles are called Gershgorin discs. Every eigenvalue of A lies in one of these Gershgorin discs.

Proof: Let λ be an eigenvalue of A and x its corresponding eigenvector. Choose i such that $|x_i| = \max_j |x_j|$. Since x can't be 0, $|x_i|$>0. Now $Ax =\lambda x$, or looking at the i-th component $(\lambda - a_{ii})x_i = \sum_{j\neq i} a_{ij}x_j$, taking the norm on both sides gives:

$$|\lambda - a_{ii}| = \left|\sum_{j\neq i} \frac{a_{ij}x_j}{x_i}\right| \leq \sum_{j\neq i}|a_{ij}|$$

As the eigenvalues of A and A^T are the same, a similar result holds for the columns of A.

Gaussian Elimination

Gaussian (or Gauss-Jordan) elimination is an algorithm in linear algebra for determining the solutions of a system of linear equations, for determining the rank of a matrix, and for calculating the inverse of an invertible square matrix.

The first step for the solution of the *n*-by-*n* system $Ax=b$ is aimed at eliminating x_1 from all but the first equation. This is obtained by pre-multiplying both left and right hand side by the sparse lower matrix L_1:

$$L_1 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -A_{21}/A_{11} & 1 & 0 & \dots & 0 \\ -A_{31}/A_{11} & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -A_{n1}/A_{11} & 0 & 0 & \dots & 1 \end{bmatrix}, \text{ with } L_1^{-1} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ A_{21}/A_{11} & 1 & 0 & \dots & 0 \\ A_{31}/A_{11} & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ A_{n1}/A_{11} & 0 & 0 & \dots & 1 \end{bmatrix}$$

The second step eliminates x_2 from all but the first two equations using a suitable lower matrix L_2, having all diagonal entries equal to *1*, and non-zero off-diagonal terms only in the second column. After the *n-th* step, there is an upper triangular matrix *U* on the left hand side and therefore the system is easily solved via a recursive relation, which is equivalent to n steps requiring pre-multiplications by suitable sparse upper matrices. Notice that $A=LU$, where $L=L_1^{-1}L_2^{-1}...L_n^{-1}$.

The computational complexity of Gaussian elimination is $O(n^3)$, that is, the number of operations required is proportional to n^3 if the matrix size is *n*-by-*n*. This is because the matrices L_1, L_2, etc. used in the process are sparse.

Sometimes it is necessary to switch two equations: for instance if *y* hadn't occurred in the second equation after our first step above, we would have switched the second and third equation and then eliminated *y* from the first equation. It is possible that the algorithm gets "stuck": for instance if *y* hadn't occurred in the second and the third equation after our first step above. In this case, the system doesn't have a unique solution. In MATLAB [L,U] = LU(X) stores an upper triangular matrix in U and a "psychologically lower triangular matrix" (i.e. a product of lower triangular and permutation matrices) in L, so that X = L*U. X can be rectangular.

Orthogonal Matrix

An orthogonal matrix is a square matrix *G* whose transpose is its inverse, i.e.,

$$GG^T = G^TG = I_n.$$

This definition can be given for matrices with entries from any field, but the most common case is the one of matrices with real entries, and only that case will be considered here. A real square matrix is orthogonal if and only if its columns form an orthonormal basis of $\mathbf{R}^n$ with the ordinary Euclidean dot product, which is the case if and only if its rows form an orthonormal basis of $\mathbf{R}^n$.

Geometrically, orthogonal matrices describe linear transformations of $\mathbf{R}^n$ which preserve angles and lengths, such as rotations and reflections. They are compatible with the Euclidean inner product in the following sense: if *G* is orthogonal and *x* and *y* are vectors in $\mathbf{R}^n$, then

$$\langle Gx, Gy \rangle = \langle x, y \rangle.$$

The inverse of every orthogonal matrix is again orthogonal, as is the matrix product of two orthogonal matrices.

The determinant of any orthogonal matrix is 1 or −1:

$$1 = \det(I) = \det(GG^T) = \det(G)\det(G^T) = (\det(G))^2.$$

In three dimensions, the orthogonal matrices with determinant 1 correspond to proper rotations and those with determinant −1 to improper rotations.

All eigenvalues of an orthogonal matrix, even the complex ones, have absolute value 1. Eigenvectors for different eigenvalues are orthogonal.

Unitary matrix

A **unitary matrix** is a *n* by *n* complex matrix *U* satisfying the condition

$$U^*U = UU^* = I_n$$

where I_n is the identity matrix and U^* is the conjugate transpose (also called the Hermitian adjoint) of *U*. Note this condition says that a matrix *U* is unitary if it has an inverse which is equal to its conjugate transpose U^*.

A unitary matrix in which all entries are real is the same thing as an orthogonal matrix .

Permutation Matrix

In linear algebra, a permutation matrix is a binary matrix that has exactly one entry 1 in each row and each column and 0s elsewhere. Permutation matrices are the matrix representation of permutations.

Any permutation matrix is orthogonal.

Similarity

Two n-by-n matrices A and B over the field K are called similar if there exists an invertible n-by-n matrix P over K such that

$P^{-1}AP = B$.

Similar matrices share many properties: they have the same rank, the same determinant, the same trace, the same eigenvalues (but not necessarily the same eigenvectors), the same characteristic polynomial and the same minimal polynomial. There are two reasons for these facts:

- two similar matrices can be thought of as describing the same linear map, but with respect to different bases
- the map $X \mapsto P^{-1}XP$ is an automorphism of the associative algebra of all n-by-n matrices

Because of this, for a given matrix A, one is interested in finding a simple "normal form" B which is similar to A -- the study of A then reduces to the study of the simpler matrix B. For example, A is called diagonalizable if it is similar to a diagonal matrix. Not all matrices are diagonalizable, but at least over the complex numbers (or any algebraically closed field), every matrix is similar to a matrix in Jordan form. Another normal form, the rational canonical form, works over any field. By looking at the Jordan forms or rational canonical forms of A and B, one can immediately decide whether A and B are similar.

Similarity of matrices does not depend on the base field: if L is a field containing K as a subfield, and A and B are two matrices over K, then A and B are similar as matrices over K if and only if they are similar as matrices over L. This is quite useful: one may safely enlarge the field K, for instance to get an algebraically closed field; Jordan forms can then be computed over the large field and can be used to determine whether the given matrices are similar over the small field. This approach can be used, for example, to show that every matrix is similar to its transpose.

If in the definition of similarity, the matrix P can be chosen to be a permutation matrix then A and B are *permutation-similar*; if P can be chosen to be a unitary matrix then A and B are *unitarily equivalent*. The spectral theorem says that every normal matrix is unitarily equivalent to some diagonal matrix.

Properties:

- $P^{-1}(A+B)P = P^{-1}AP + P^{-1}BP$
- $P^{-1}(AB)P = (P^{-1}AP)(P^{-1}BP)$
- $P^{-1}(\sigma A)P = \sigma(P^{-1}AP)$
- $P^{-1}(A^n)P = (P^{-1}AP)^n$
- $P^{-1}AP$ and $P^{-1}AP$ have the same rank, the same determinant, the same trace, the same eigenvalues (but not necessarily the same eigenvectors), the same characteristic polynomial and the same minimal polynomial

Normal Matrix

A complex square matrix A is a **normal matrix** iff

$$A^* A = AA^*$$

where A^* is the conjugate transpose of A (if A is a real matrix, this is the same as the transpose of A).

Examples of normal matrices are unitary matrices, hermitian matrices and positive definite matrices.

It is useful to think of normal matrices in analogy to complex numbers, invertible normal matrices in analogy to non-zero complex numbers, the conjugate transpose in analogy to the complex conjugate, unitary matrices in analogy to complex numbers of absolute value 1, hermitian matrices in analogy to real numbers and positive definite matrices in analogy to positive real numbers.

The concept of normality is mainly important because normal matrices are precisely the ones to which the spectral theorem applies; in other words, normal matrices are precisely those matrices that can be represented by a diagonal matrix with respect to a properly chosen orthonormal basis of $\mathbf{C}^n$. Phrased differently: a matrix is normal if and only if its eigenspaces span $\mathbf{C}^n$ and are pairwise orthogonal with respect to the standard inner product of $\mathbf{C}^n$.

In general, the sum or product of two normal matrices need not be normal. However, if A and B are normal with $AB = BA$, then both AB and $A + B$ are also normal and furthermore we can simultaneously diagonalize A and B in the following sense: there exists a unitary matrix U such UAU^* and UBU^* are both diagonal matrices. In this case, the columns of U^* are eigenvectors of both A and B and form an orthonormal basis of $\mathbf{C}^n$.

Diagonalizable matrix

A square matrix A is called **diagonalizable** if it is similar to a diagonal matrix, i.e. if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix. If V is a finite-dimensional vector space, then a linear map $T : V \to V$ is called **diagonalizable** if there exists a basis of V with respect to which T is represented by a diagonal matrix. **Diagonalization** is the process of finding a corresponding diagonal matrix for a diagonalizable matrix or linear map.

Diagonalizable matrices and maps are of interest because diagonal matrices are especially easy to handle: their eigenvalues and eigenvectors are known and one can raise a diagonal matrix to a power by simply raising the diagonal entries to that same power.

An n-by-n matrix A over the field F is diagonalizable if and only if the sum of the dimensions of its eigenspaces is equal to n, which is the case if and only if there exists a basis of F^n consisting of eigenvectors of A. If such a basis has been found, one can form the matrix P having these basis vectors as columns, and $P^{-1}AP$ will be a diagonal matrix. The diagonal entries of this matrix are the eigenvalues of A.

Another characterization: A matrix or linear map is diagonalizable over the field F if and only if its minimal polynomial is a product of distinct linear factors over F.

The following sufficient (but not necessary) conditions are often useful:

- an n-by-n matrix A is diagonalizable if its characteristic polynomial has n distinct roots
- any normal matrix A (such that $AA^*=A^*A$) is diagonalizable

Positive-definite Matrix

In linear algebra, the **positive-definite matrices** are (in several ways) analogous to the positive real numbers. An $n \times n$ Hermitian matrix M is said to be **positive definite** if it has one (and therefore all) of the following 6 equivalent properties:

(1) For all non-zero vectors z in $\mathbf{C}^n$ we have

$$z^* M z > 0.$$

Here we view z as a column vector with n complex entries and z^* as the complex conjugate of its transpose. (M being Hermitian, $z^* M z$ is always real.)

(2) For all non-zero vectors x in $\mathbf{R}^n$ we have

$$x^{\mathrm{T}} M x > 0$$

(where x^{T} denotes the transpose of the column vector x).

(3) For all non-zero vectors u in $\mathbf{Z}^n$ (all components being integers), we have

$$u^{\mathrm{T}} M u > 0.$$

(4) All eigenvalues of M are positive.

(5) The form

$$\langle x, y \rangle = x^* M y$$

defines an inner product on $\mathbf{C}^n$. (In fact, every inner product on $\mathbf{C}^n$ arises in this fashion from a Hermitian positive definite matrix.)

(6) All the following matrices have positive determinant: the upper left 1-by-1 corner of M, the upper left 2-by-2 corner of M, the upper left 3-by-3 corner of M, ..., and M itself.

Further properties

Every positive definite matrix is invertible and its inverse is also positive definite. If M is positive definite and $r > 0$ is a real number, then rM is positive definite. If M and N are positive definite, then $M + N$ is also positive definite, and if $MN = NM$, then MN is also positive definite. Every positive definite matrix M, has at least one square root matrix N such that $N^2 = M$. In fact, M may have infinitely many square roots, but exactly one positive definite square root.

Negative-definite, semidefinite and indefinite matrices

The Hermitian matrix M is said to be **negative-definite** if

$$x^* M x < 0$$

for all non-zero x in $\mathbf{R}^n$ (or, equivalently, all non-zero x in $\mathbf{C}^n$). It is called **positive-semidefinite** if

$$x^* M x \geq 0$$

for all x in $\mathbf{R}^n$ (or $\mathbf{C}^n$) and **negative-semidefinite** if

$$x^* M x \leq 0$$

for all x in $\mathbf{R}^n$ (or $\mathbf{C}^n$).

A Hermitian matrix which is neither positive- nor negative-semidefinite is called **indefinite**.

Rayleigh quotient

For a given real symmetric matrix A and real nonzero vector x, the Rayleigh quotient $R(A,x)$ is defined as:

$$\frac{x^T A x}{x^T x}$$

Note that $R(A,c \cdot x) = R(A,x)$ for any real scalar c.

It can be shown that this quotient reaches its minimum value λ_{min} (the smallest eigenvalue of A) when x is v_{min} (the corresponding eigenvector). Similarly, $R(A,x) \leq \lambda_{max}$ and $R(A,v_{max}) = \lambda_{max}$:

$$\frac{x^T A x}{x^T x} = \frac{x^T P^T D P x}{x^T P^T P x} = \frac{z^T D z}{z^T z} = \frac{\sum_i D_{ii} z_i^2}{\sum_i z_i^2} \leq \frac{\lambda_{\max} \sum_i z_i^2}{\sum_i z_i^2} = \lambda_{\max}$$

The Rayleigh quotient is used in eigenvalue algorithms to obtain an eigenvalue approximation from an eigenvector approximation. Specifically, this is the basis for Rayleigh quotient iteration.

Singular value decomposition

Any real $m \times n$ matrix A can be decomposed into

$$A = USV^T$$

where U is an $m \times m$ orthogonal matrix, V is an $n \times n$ orthogonal matrix, and S is a unique $m \times n$ diagonal matrix with real, non-negative elements σ_i, $i = 1, \ldots, \min(m,n)$, in descending order:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(m,n)} \geq 0$$

The σ_i are the *singular values* of A and the first $\min(m,n)$ columns of U and V are the left and right (respectively) *singular vectors* of A. S has the form:

$$\begin{bmatrix} \Sigma \\ 0 \end{bmatrix} \text{ if } m \geq n \text{ and } \begin{bmatrix} \Sigma & 0 \end{bmatrix} \text{ if } m < n,$$

where Σ is a diagonal matrix with the diagonal elements $\sigma_1, \sigma_2, \ldots, \sigma_{\min(m,n)}$. We assume now $m \geq n$. If $r = \operatorname{rank}(A) < n$, then

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

If $\sigma_r \neq 0$ and $\sigma_{r+1} = \cdots = \sigma_n = 0$, then r is the rank of A. In this case, S becomes an $r \times r$ matrix, and U and V shrink accordingly. SVD can thus be used for rank determination.

The SVD provides a numerically robust solution to the least-squares problem. The matrix-algebraic phrasing of the least-squares solution x is

$$x = (A^T A)^{-1} A^T b$$

Then utilizing the SVD by making the replacement $A = USV^T$ we have

$$x = V \begin{bmatrix} \Sigma^{-1} & 0 \end{bmatrix} U^T b.$$

In MATLAB the svd command computes the matrix singular value decomposition.

s = svd(X) returns a vector of singular values.

[U,S,V] = svd(X) produces a diagonal matrix S of the same dimension as X, with nonnegative diagonal elements in decreasing order, and unitary matrices U and V so that X = U*S*V'.

[U,S,V] = svd(X,0) produces the "economy size" decomposition. If X is m-by-n with m > n, then svd computes only the first n columns of U and S is n-by-n.

Scalar product

In the following article, the field of scalars denoted **F** is either the field of real numbers **R** or the field of complex numbers **C**. Formally, an inner product space is a vector space *V* over the field **F** together with a bilinear form, called an *inner product*

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbf{F}$$

satisfying the following axioms:

- Nonnegativity:
  $$\forall x \in V,\ \langle x, x \rangle \geq 0.$$
- Nondegeneracy:
  $$\forall x \in V,\ \langle x, x \rangle = 0 \text{ iff } x = 0.$$
- Conjugate symmetry:
  $$\forall x, y \in V,\ \langle x, y \rangle = \overline{\langle y, x \rangle}$$
  (Conjugation is also often written with an asterisk, as in <y,x>, as is the conjugate transpose.)*
- Sesquilinearity:
  $$\forall b \in F,\ \forall x, y \in V,\ \langle x, by \rangle = b\langle x, y \rangle$$
  $$\forall x, y, z \in V,\ \langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle.$$
  By combining these with conjugate symmetry, we get:
  $$\forall a \in F,\ \forall x, y \in V,\ \langle ax, y \rangle = \overline{a}\langle x, y \rangle$$
  $$\forall x, y, z \in V,\ \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle.$$

Note that if **F**=**R**, then the conjugate symmetry property is simply *symmetry* of the inner product, i.e.

$$\langle x, y \rangle = \langle y, x \rangle$$

In this case, sesquilinearity becoms standard linearity.

Remark. Many mathematical authors require an inner product to be linear in the first argument and conjugate-linear in the second argument, contrary to the convention adopted above. This change is immaterial, but the definition above ensures a smoother connection to the bra-ket notation used by physicists in quantum mechanics and is now often used by mathematicians as well. Some authors adopt the convention that < , > is linear in the first component while < | > is linear in the second component, although this is by no means universal. For instance the G. Emch reference does not follow this convention.

In some cases we need to consider non-negative *semi-definite* sesquilinear forms. This means that $\langle x, x \rangle$ is only required to be non-negative. We show how to treat these below.

[edit]

Examples

A trivial example are the real numbers with the standard multiplication as the inner product

$$\langle x, y \rangle := xy$$

More generally any Euclidean space $\mathbf{R}^n$ with the dot product is an inner product space

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle := \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \ldots + x_n y_n$$

Even more generally any positive-definite matrix *M* can be used to define an inner product on $\mathbf{C}^n$ as

$$\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^* \mathbf{M} \mathbf{y}$$

with $\mathbf{x}^*$ the conjugate transpose of **x**.

The article on Hilbert space has several examples of inner product spaces where the metric induced by the inner product yields a complete metric spaces. An example of an inner product which induces an incomplete metric is is the space $C[a, b]$ of continuous complex valued functions on the interval $[a,b]$. The inner product is

$$\langle f, g \rangle := \int_a^b \overline{f(t)} g(t)\, dt$$

This space is not complete; consider for example, for the interval [0,1] the sequence of functions $\{ f_k \}_k$ where

- $f_k(t)$ is 1 for t in the subinterval [0, 1/2]
- $f_k(t)$ is 0 for t in the subinterval $[1/2 + 1/k, 1]$
- f_k is affine in $[1/2, 1/2 + 1/k]$

This sequence is a Cauchy sequence which does not converge to a *continuous* function.