Positive-definite matrix

From Wikipedia, the free encyclopedia

In linear algebra, a symmetric *n* × *n* real matrix M is said to be **positive definite** if $z^T M z$ is positive for every non-zero column vector z of n real numbers. Here z^T denotes the transpose of z.

More generally, an *n* × *n* Hermitian matrix M is said to be **positive definite** if $z^* M z$ is real and positive for all non-zero column vectors z of n complex numbers. Here z^* denotes the conjugate transpose of z.

The **negative definite**, **positive semi-definite**, and **negative semi-definite** matrices are defined in the same way, except that the expression $z^T M z$ or $z^* M z$ is required to be always negative, non-negative, and non-positive, respectively.

Positive definite matrices are closely related to positive-definite symmetric bilinear forms (or sesquilinear forms in the complex case), and to inner products of vector spaces.[1]

Some authors use more general definitions of "positive definite" that include some non-symmetric real matrices, or non-Hermitian complex ones.

Contents

Examples

- The identity matrix $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is positive definite. Seen as a real matrix, it is symmetric, and, for any non-zero column vector *z* with real entries *a* and *b*, one has $z^{\mathrm{T}} I z = \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = a^2 + b^2$. Seen as a complex matrix, for any non-zero column vector *z* with complex entries *a* and *b* one has $z^* I z = \begin{bmatrix} a^* & b^* \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = a^* a + b^* b = |a|^2 + |b|^2$. Either way, the result is positive since *z* is not the zero vector (that is, at least one of *a* and *b* is not zero).

- The real symmetric matrix

  $$M = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

  is positive definite since for any non-zero column vector *z* with entries *a*, *b* and *c*, we have

  $$\begin{aligned} z^{\mathrm{T}} M z = (z^{\mathrm{T}} M) z &= \begin{bmatrix} (2a-b) & (-a+2b-c) & (-b+2c) \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ &= 2a^2 - 2ab + 2b^2 - 2bc + 2c^2 \\ &= a^2 + (a-b)^2 + (b-c)^2 + c^2 \end{aligned}$$

  This result is a sum of squares, and therefore non-negative; and is zero only if *a* = *b* = *c* = 0, that is, when *z* is zero.

- The real symmetric matrix

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

  is not positive definite. If *z* is the vector $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, one has

  $$z^{\mathrm{T}} N z = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -2 \not> 0.$$

- For any real non-singular matrix A, the product $A^T A$ is a positive definite matrix. A simple proof is that for any non-zero vector z, the condition $z^T A^T A z = \|Az\|_2^2 > 0$, since the non-singularity of matrix A means that $Az \neq 0$.

The examples *M* and *N* above show that a matrix in which some elements are negative may still be positive-definite, and conversely a matrix whose entries are all positive may not be positive definite.

Connections

The general purely quadratic real function $f(z)$ on n real variables $z_1, ..., z_n$ can always be written as $z^T Mz$ where z is the column vector with those variables, and M is a symmetric real matrix. Therefore, the matrix being positive definite means that f has a unique minimum (zero) when z is zero, and is strictly positive for any other z.

More generally, a twice-differentiable real function f on n real variables has an isolated local minimum at arguments $z_1, ..., z_n$ if its gradient is zero and its Hessian (the matrix of all second derivatives) is positive definite at that point. Similar statements can be made for negative definite and semi-definite matrices.

In statistics, the covariance matrix of a multivariate probability distribution is always positive semi-definite; and it is positive definite unless one variable is an exact linear combination of the others. Conversely, every positive semi-definite matrix is the covariance matrix of some multivariate distribution.

Characterizations

Let M be an $n \times n$ Hermitian matrix. The following properties are equivalent to M being positive definite:

1. **All its eigenvalues are positive.** Let $P^{-1}DP$ be an eigendecomposition of M, where P is a unitary complex matrix whose rows comprise an orthonormal basis of eigenvectors of M, and D is a *real* diagonal matrix whose main diagonal contains the corresponding eigenvalues. The matrix M may be regarded as a diagonal matrix D that has been re-expressed in coordinates of the basis P. In particular, the one-to-one change of variable $y = Pz$ shows that z^*Mz is real and positive for any complex vector z if and only if y^*Dy is real and positive for any y; in other words, if D is positive definite. For a diagonal matrix, this is true only if each element of the main diagonal—that is, every eigenvalue of M—is positive. Since the spectral theorem guarantees all eigenvalues of a Hermitian matrix to be real, the positivity of eigenvalues can be checked using Descartes' rule of alternating signs when the characteristic polynomial of a real, symmetric matrix M is available.
2. **The associated sesquilinear form is an inner product.** The sesquilinear form defined by M is the function $\langle \cdot, \cdot \rangle$ from $\mathbf{C}^n \times \mathbf{C}^n$ to $\mathbf{C}$ such that $\langle x, y \rangle := y^* M x$ for all x and y in $\mathbf{C}^n$, where y^* is the complex conjugate of y. For any complex matrix M, this form is linear in each argument separately. Therefore the form is an inner product on $\mathbf{C}^n$ if and only if $\langle z, z \rangle$ is real and positive for all nonzero z; that is if and only if M is positive definite. (In fact, every inner product on $\mathbf{C}^n$ arises in this fashion from a Hermitian positive definite matrix.)
3. **It is the Gram matrix of linearly independent vectors.** Let $x_1, \ldots, x_n$ be a list of n linearly independent vectors of some complex vector space with an inner product $\langle \cdot, \cdot \rangle$. It can be verified that the Gram matrix M of those vectors, defined by $M_{ij} = \langle x_i, x_j \rangle$, is always positive definite. Conversely, if M is positive definite, it has an eigendecomposition $P^{-1}DP$ where P is unitary, D diagonal, and all diagonal elements $D_{ii} = \lambda_i$ of D are real and positive. Let E be the real diagonal matrix with entries $E_{ii} = \sqrt{\lambda_i}$ so $E^2 = D$; then $P^{-1}DP = P^*DP = P^*EEP = (EP)^*EP$. Now we let $x_1, \ldots, x_n$ be the columns of EP. These vectors are linearly independent, and by the above M is their Gram matrix, under the standard inner product of $\mathbf{C}^n$, namely $\langle x_i, x_j \rangle = x_i^T x_j$
4. **Its leading principal minors are all positive.** The kth leading principal minor of a matrix M is the

determinant of its upper-left k by k sub-matrix. It turns out that a matrix is positive definite if and only if all these determinants are positive. This condition is known as Sylvester's criterion, and provides an efficient test of positive-definiteness of a symmetric real matrix. Namely, the matrix is reduced to an upper triangular matrix by using elementary row operations, as in the first part of the Gaussian elimination method, taking care to preserve the sign of its determinant during pivoting process. Since the kth leading principal minor of a triangular matrix is the product of its diagonal elements up to row k, Sylvester's criterion is equivalent to checking whether its diagonal elements are all positive. This condition can be checked each time a new row k of the triangular matrix is obtained.

5. **It has a unique Cholesky decomposition.** The matrix M is positive definite if and only if there exists a unique lower triangular matrix L, with real and strictly positive diagonal elements, such that $M = LL^*$. This factorization is called the Cholesky decomposition of M.

Quadratic forms

The (purely) quadratic form associated with a real matrix M is the function $Q : \mathbf{R}^n \to \mathbf{R}$ such that $Q(x) = x^{\mathrm{T}}Mx$ for all x. It turns out that the matrix M is positive definite if and only if it is symmetric and its quadratic form is a strictly convex function.

More generally, any quadratic function from $\mathbf{R}^n$ to $\mathbf{R}$ can be written as $x^{\mathrm{T}}Mx + x^{\mathrm{T}}b + c$ where M is a symmetric $n \times n$ matrix, b is a real n-vector, and c a real constant. This quadratic function is strictly convex when M is positive definite, and hence has a unique finite global minimum, if and only if M is positive definite. For this reason, positive definite matrices play an important role in optimization problems.

Simultaneous diagonalization

A symmetric matrix and another symmetric and positive-definite matrix can be simultaneously diagonalized, although not necessarily via a similarity transformation. This result does not extend to the case of three or more matrices. In this section we write for the real case. Extension to the complex case is immediate.

Let M be a symmetric and N a symmetric and positive-definite matrix. Write the generalized eigenvalue equation as $(M-\lambda N)x = 0$ where we impose that x be normalized, i.e. $x^{\mathrm{T}}Nx = 1$. Now we use Cholesky decomposition to write the inverse of N as $Q^{\mathrm{T}}Q$. Multiplying by Q and Q^{T}, we get $Q(M-\lambda N)Q^{\mathrm{T}}x = 0$, which can be rewritten as $(QMQ^{\mathrm{T}})y = \lambda y$ where $y^{\mathrm{T}}y = 1$. Manipulation now yields $MX = NX\Lambda$ where X is a matrix having as columns the generalized eigenvectors and Λ is a diagonal matrix with the generalized eigenvalues. Now premultiplication with X^{T} gives the final result: $X^{\mathrm{T}}MX = \Lambda$ and $X^{\mathrm{T}}NX = I$, but note that this is no longer an orthogonal diagonalization.

Note that this result does not contradict what is said on simultaneous diagonalization in the article Diagonalizable matrix, which refers to simultaneous diagonalization by a similarity transformation. Our result here is more akin to a simultaneous diagonalization of two quadratic forms, and is useful for optimization of one form under conditions on the other. For this result see Horn&Johnson, 1985, page 218 and following.

Negative-definite, semidefinite and indefinite matrices

A Hermitian matrix is negative-definite, negative-semidefinite, or positive-semidefinite if and only if all of its eigenvalues are negative, non-positive, or non-negative, respectively.

Negative-definite

The $n \times n$ Hermitian matrix M is said to be *negative-definite* if

$$x^* M x < 0$$

for all non-zero x in $\mathbf{C}^n$ (or, all non-zero x in $\mathbf{R}^n$ for the real matrix), where x^* is the conjugate transpose of x.

A matrix is negative definite if its k-th order leading principal minor is negative when k is odd, and positive when k is even.

Positive-semidefinite

M is called *positive-semidefinite* (or sometimes *nonnegative-definite*) if

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

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

A matrix M is positive-semidefinite if and only if it arises as the Gram matrix of some set of vectors. In contrast to the positive-definite case, these vectors need not be linearly independent.

For any matrix A, the matrix A^*A is positive semidefinite, and rank(A) = rank(A^*A). Conversely, any Hermitian positive semi-definite matrix M can be written as $M = LL^*$, where L is lower triangular; this is the Cholesky decomposition. If M is not positive definite, then some of the diagonal elements of L may be zero.

A Hermitian matrix is positive semidefinite if and only if all of its principal minors are nonnegative. It is however not enough to consider the leading principal minors only, as is checked on the diagonal matrix with entries 0 and -1.

Negative-semidefinite

It is called *negative-semidefinite* if

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

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

Indefinite

A Hermitian matrix which is neither positive definite, negative definite, positive-semidefinite, nor negative-semidefinite is called *indefinite*. Indefinite matrices are also characterized by having both positive and negative eigenvalues.

Further properties

If M is a Hermitian positive-semidefinite matrix, one sometimes writes $M \geq 0$ and if M is positive-definite one writes $M > 0$.[2] The notion comes from functional analysis where positive-semidefinite matrices define positive operators.

For arbitrary square matrices M, N we write $M \geq N$ if $M - N \geq 0$; i.e., $M - N$ is positive semi-definite. This defines a partial ordering on the set of all square matrices. One can similarly define a strict partial ordering $M > N$.

1. Every positive definite matrix is invertible and its inverse is also positive definite.[3] If $M \geq N > 0$ then $N^{-1} \geq M^{-1} > 0$, and $\sqrt{M} > \sqrt{N} > 0$.[4] Moreover, by the min-max theorem, the kth largest eigenvalue of M is greater than the kth largest eigenvalue of N
2. If M is positive definite and $r > 0$ is a real number, then rM is positive definite.[5] If M and N are positive definite, then the sum $M + N$[5] and the products MNM and NMN are also positive definite. If $MN = NM$, then MN is also positive definite.
3. Every principal submatrix of a positive definite matrix is positive definite.
4. $Q^T M Q$ is positive-semidefinite. If Q is invertible, then $Q^T M Q$ is positive definite. Note that $Q^{-1} M Q$ need not be positive definite.
5. The determinant of M is bounded by the product of its diagonal elements.
6. The diagonal entries m_{ii} are real and non-negative. As a consequence the trace, $\mathrm{tr}(M) \geq 0$. Furthermore,[6] since every principal sub matrix (in particular, 2-by-2) is positive definite,

   $$|m_{ij}| \leq \sqrt{m_{ii}m_{jj}} \leq \frac{m_{ii}+m_{jj}}{2}$$

   and thus

   $$\max |m_{ij}| \leq \max |m_{ii}|$$

7. A matrix M is positive semi-definite if and only if there is a positive semi-definite matrix B with $B^2 = M$. This matrix B is unique,[7] is called the square root of M, and is denoted with $B = M^{1/2}$ (the square root B is not to be confused with the matrix L in the Cholesky factorization $M = LL^*$, which is also sometimes called the square root of M). If $M > N > 0$ then $M^{1/2} > N^{1/2} > 0$.
8. If M is a symmetric matrix of the form $m_{ij} = m(i-j)$, and the *strict* inequality holds

   $$\sum_{j\neq 0} |m(j)| < m(0)$$

   then M is *strictly* positive definite.

9. Let $M > 0$ and N Hermitian. If $MN + NM \geq 0$ (resp., $MN + NM > 0$) then $N \geq 0$ (resp., $N > 0$).
10. If $M > 0$ is real, then there is a $\delta > 0$ such that $M > \delta I$, where I is the identity matrix.
11. If M_k denotes the leading k by k minor, $\det(M_k)/\det(M_{k-1})$ is the kth pivot during LU decomposition.
12. The set of positive semidefinite symmetric matrices is convex. That is, if M and N are positive semidefinite, then for any α between 0 and 1, $\alpha M + (1-\alpha)N$ is also positive semidefinite. For any vector x:

$$x^\top(\alpha M+(1-\alpha)N)x=\alpha x^\top Mx+(1-\alpha)x^\top Nx\geq 0.$$

This property guarantees that semidefinite programming problems converge to a globally optimal solution.

13. If $M,N \geq 0$, although MN is not necessary positive-semidefinite, the Kronecker product $M \otimes N \geq 0$, the Hadamard product $M \circ N \geq 0$ (this result is often called the Schur product theorem).,[8] and the Frobenius product $M : N \geq 0$ (Lancaster-Tismenetsky, The Theory of Matrices, p. 218).
14. Regarding the Hadamard product of two positive-semidefinite matrices $M = (m_{ij}) \geq 0$, $N \geq 0$, there are two notable inequalities:
    - Oppenheim's inequality: $\det(M\circ N)\geq \det(N)\prod_i m_{ii}$.[9]
    - $\det(M \circ N) \geq \det(M)\det(N)$.[10]

Block matrices

A positive $2n \times 2n$ matrix may also be defined by blocks:

$$M=\begin{bmatrix}A & B\\ C & D\end{bmatrix}$$

where each block is $n \times n$. By applying the positivity condition, it immediately follows that A and D are hermitian, and $C = B^*$.

We have that $z^*Mz \geq 0$ for all complex z, and in particular for $z = (v, 0)^\mathrm{T}$. Then

$$\begin{bmatrix}v^* & 0\end{bmatrix}\begin{bmatrix}A & B\\ B^* & D\end{bmatrix}\begin{bmatrix}v\\ 0\end{bmatrix}=v^*Av\geq 0.$$

A similar argument can be applied to D, and thus we conclude that both A and D must be positive definite matrices, as well.

Converse results can be proved with stronger conditions on the blocks, for instance using the Schur complement.

On the definition

Consistency between real and complex definitions

Since every real matrix is also a complex matrix, the definitions of "positive definite" for the two classes must agree.

For complex matrices, the most common definition says that "M is positive definite if and only if z^*Mz is real and positive for all non-zero *complex* column vectors z". This condition implies that M is Hermitian, that is, its transpose is equal to its conjugate. To see this, consider the matrices $A = (M+M^*)/2$ and $B = (M$

$-M^*)/(2i)$, so that $M = A+iB$ and $z^*Mz = z^*Az + iz^*Bz$. The matrices A and B are Hermitian, therefore z^*Az and z^*Bz are individually real. If z^*Mz is real, then z^*Bz must be zero for all z. Then B is the zero matrix and $M = A$, proving that M is Hermitian.

By this definition, a positive definite *real* matrix M is Hermitian, hence symmetric; and z^TMz is positive for all non-zero *real* column vectors z. However the last condition alone is not sufficient for M to be positive definite. For example, if

$$M = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix},$$

then for any real vector z with entries a and b we have $z^TMz = (a-b)a + (a+b)b = a^2 + b^2$, which is always positive if z is not zero. However, if z is the complex vector with entries 1 and i, one gets

$$z^*Mz = [1, -i]M[1, i]^T = [1+i, 1-i][1, i]^T = 2 + 2i,$$

which is not real. Therefore, M is not positive definite.

On the other hand, for a *symmetric* real matrix M, the condition "$z^TMz > 0$ for all nonzero real vectors z" *does* imply that M is positive definite in the complex sense.

Extension for non symmetric matrices

Some authors choose to say that a complex matrix M is positive definite if $\mathrm{Re}(z^*Mz) > 0$ for all non-zero complex vectors z, where $\mathrm{Re}(c)$ denotes the real part of a complex number c.[11] This weaker definition encompasses some non-Hermitian complex matrices, including some non-symmetric real ones, such as $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$.

Indeed, with this definition, a real matrix is positive definite if and only if $z^TMz > 0$ for all nonzero real vectors z, even if M is not symmetric.

In general, we have $\mathrm{Re}(z^*Mz) > 0$ for all complex nonzero vectors z if and only if the Hermitian part $(M + M^*)/2$ of M is positive definite in the narrower sense. Similarly, we have $x^TMx > 0$ for all real nonzero vectors x if and only if the symmetric part $(M + M^T)/2$ of M is positive definite in the narrower sense.

In summary, the distinguishing feature between the real and complex case is that, a bounded positive operator on a complex Hilbert space is necessarily Hermitian, or self adjoint. The general claim can be argued using the polarization identity. That is no longer true in the real case.

See also

- Cholesky decomposition
- Covariance matrix
- M-matrix
- Positive-definite function
- Positive-definite kernel

- Schur complement
- Square root of a matrix
- Sylvester's criterion

Notes

1. Stewart, J. (1976). Positive definite functions and generalizations, an historical survey. Rocky Mountain J. Math, 6(3). (http://projecteuclid.org/DPubS?verb=Display&version=1.0&service=UI&handle=euclid.rmjm/1250130219&page=record)
2. This may be confusing, as sometimes nonnegative matrices are also denoted in this way. A common alternative notation is $M \succeq 0$ and $M \succ 0$ for positive semidefinite and positive definite matrices, respectively.
3. Horn & Johnson (1985), p. 397
4. Horn & Johnson (1985), Corollary 7.7.4(a)
5. Horn & Johnson (1985), Observation 7.1.3
6. Horn & Johnson (1985), p. 398
7. Horn & Johnson (1985), Theorem 7.2.6 with $k = 2$
8. Horn & Johnson (1985), Theorem 7.5.3
9. Horn & Johnson (1985), Theorem 7.8.6
10. (Styan 1973)
11. Weisstein, Eric W. *Positive Definite Matrix. (http://mathworld.wolfram.com/PositiveDefiniteMatrix.html)* From *MathWorld--A Wolfram Web Resource*. Accessed on 2012-07-26

References

- Horn, Roger A.; Johnson, Charles R. (1990), *Matrix Analysis*, Cambridge University Press, ISBN 978-0-521-38632-6.
- Rajendra Bhatia. *Positive definite matrices*. Princeton Series in Applied Mathematics, 2007. ISBN 978-0-691-12918-1.

External links

- Hazewinkel, Michiel, ed. (2001), "Positive-definite form" (http://www.encyclopediaofmath.org/index.php?title=p/p073880), *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4
- Wolfram MathWorld: Positive Definite Matrix (http://mathworld.wolfram.com/PositiveDefiniteMatrix.html)