

Diagonalizable matrix

In linear algebra, a square matrix A is called **diagonalizable** or **non-defective** if it is similar to a diagonal matrix. That is, if there exists an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$. This is equivalent to $A = PDP^{-1}$. (Such P , D are not unique.) This property exists for any linear map: for a finite-dimensional vector space V , a linear map $T : V \rightarrow V$ is called **diagonalizable** if there exists an ordered basis of V consisting of eigenvectors of T . These definitions are equivalent: if T has a matrix representation $T = PDP^{-1}$ as above, then the column vectors of P form a basis consisting of eigenvectors of T , and the diagonal entries of D are the corresponding eigenvalues of T ; with respect to this eigenvector basis, T is represented by D .

Diagonalization is the process of finding the above P and D and makes many subsequent computations easier. One can raise a diagonal matrix D to a power by simply raising the diagonal entries to that power. The determinant of a diagonal matrix is simply the product of all diagonal entries. Such computations generalize easily to $A = PDP^{-1}$.

The geometric transformation represented by a diagonalizable matrix is an inhomogeneous dilation (or anisotropic scaling). That is, it can scale the space by a different amount in different directions. The direction of each eigenvector is scaled by a factor given by the corresponding eigenvalue.

A square matrix that is not diagonalizable is called defective. It can happen that a matrix A with real entries is defective over the real numbers, meaning that $A = PDP^{-1}$ is impossible for any invertible P and diagonal D with real entries, but it is possible with complex entries, so that A is diagonalizable over the complex numbers. For example, this is the case for a generic rotation matrix.

Many results for diagonalizable matrices hold only over an algebraically closed field (such as the complex numbers). In this case, diagonalizable matrices are dense in the space of all matrices, which means any defective matrix can be deformed into a diagonalizable matrix by a small perturbation; and the Jordan normal form theorem states that any matrix is uniquely the sum of a diagonalizable matrix and a nilpotent matrix. Over an algebraically closed field, diagonalizable matrices are equivalent to semi-simple matrices.

Definition

A square $n \times n$ matrix, A , with entries in a field F is called **diagonalizable** or **nondefective** if there exists an $n \times n$ invertible matrix (i.e. an element of the general linear group $\text{GL}_n(F)$), P , such that $P^{-1}AP$ is a diagonal matrix. Formally,

$$A \in F^{n \times n} \text{ diagonalizable} \iff \exists P \in \text{GL}_n(F) : P^{-1}AP \text{ diagonal}$$

Characterization

The fundamental fact about diagonalizable maps and matrices is expressed by the following:

- An $n \times n$ matrix A over a 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 whose diagonal entries are the eigenvalues of A . The matrix P is known as a modal matrix for A .
- A linear map $T : V \rightarrow V$ is diagonalizable if and only if the sum of the dimensions of its eigenspaces is equal to $\dim(V)$, which is the case if and only if there exists a basis of V consisting of eigenvectors of T . With respect to such a basis, T will be represented by a diagonal matrix. The diagonal entries of this matrix are the eigenvalues of T .

The following sufficient (but not necessary) condition is often useful.

- An $n \times n$ matrix A is diagonalizable over the field F if it has n distinct eigenvalues in F , i.e. if its characteristic polynomial has n distinct roots in F ; however, the converse may be false. Consider

$$\begin{bmatrix} -1 & 3 & -1 \\ -3 & 5 & -1 \\ -3 & 3 & 1 \end{bmatrix},$$

which has eigenvalues 1, 2, 2 (not all distinct) and is diagonalizable with diagonal form (similar to A)

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

and change of basis matrix P :

$$\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}.$$

The converse fails when A has an eigenspace of dimension higher than 1. In this example, the eigenspace of A associated with the eigenvalue 2 has dimension 2.

- A linear map $T : V \rightarrow V$ with $n = \dim(V)$ is diagonalizable if it has n distinct eigenvalues, i.e. if its characteristic polynomial has n distinct roots in F .

Let A be a matrix over F . If A is diagonalizable, then so is any power of it. Conversely, if A is invertible, F is algebraically closed, and A^n is diagonalizable for some n that is not an integer multiple of the characteristic of F , then A is diagonalizable. Proof: If A^n is diagonalizable, then A is annihilated by some polynomial $(x^n - \lambda_1) \cdots (x^n - \lambda_k)$, which has no multiple root (since $\lambda_j \neq 0$) and is divided by the minimal polynomial of A .

Over the complex numbers \mathbb{C} , almost every matrix is diagonalizable. More precisely: the set of complex $n \times n$ matrices that are *not* diagonalizable over \mathbb{C} , considered as a subset of $\mathbb{C}^{n \times n}$, has Lebesgue measure zero. One can also say that the diagonalizable matrices form a dense subset with respect to the Zariski topology: the non-diagonalizable matrices lie inside the vanishing set of the discriminant of the characteristic polynomial, which is a hypersurface. From that follows also density in the usual (strong) topology given by a norm. The same is not true over \mathbb{R} .

The Jordan–Chevalley decomposition expresses an operator as the sum of its semisimple (i.e., diagonalizable) part and its nilpotent part. Hence, a matrix is diagonalizable if and only if its nilpotent part is zero. Put in another way, a matrix is diagonalizable if each block in its Jordan form has no nilpotent part;

i.e., each "block" is a one-by-one matrix.

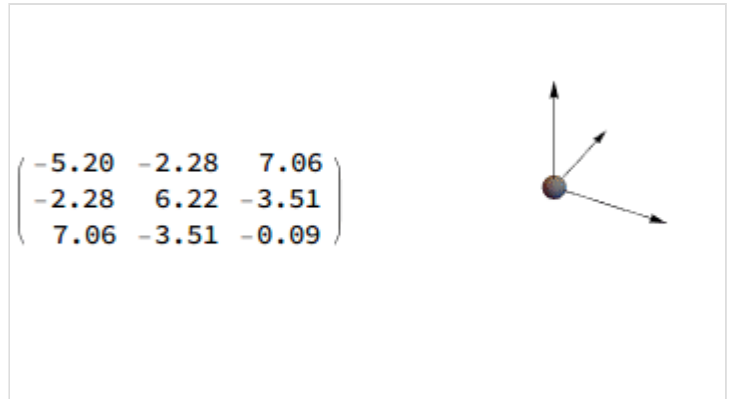
Diagonalization

If a matrix A can be diagonalized, that is,

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

then:

$$AP = P \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$



The diagonalization of a symmetric matrix can be interpreted as a rotation of the axes to align them with the eigenvectors.

Writing P as a block matrix of its column vectors α_i

$$P = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n],$$

the above equation can be rewritten as

$$A\alpha_i = \lambda_i\alpha_i \quad (i = 1, 2, \dots, n).$$

So the column vectors of P are right eigenvectors of A , and the corresponding diagonal entry is the corresponding eigenvalue. The invertibility of P also suggests that the eigenvectors are linearly independent and form a basis of \mathbb{F}^n . This is the necessary and sufficient condition for diagonalizability and the canonical approach of diagonalization. The row vectors of P^{-1} are the left eigenvectors of A .

When a complex matrix $A \in \mathbb{C}^{n \times n}$ is a Hermitian matrix (or more generally a normal matrix), eigenvectors of A can be chosen to form an orthonormal basis of \mathbb{C}^n , and P can be chosen to be a unitary matrix. If in addition, $A \in \mathbb{R}^{n \times n}$ is a real symmetric matrix, then its eigenvectors can be chosen to be an orthonormal basis of \mathbb{R}^n and P can be chosen to be an orthogonal matrix.

For most practical work matrices are diagonalized numerically using computer software. Many algorithms exist to accomplish this.

Simultaneous diagonalization

A set of matrices is said to be *simultaneously diagonalizable* if there exists a single invertible matrix P such that $P^{-1}AP$ is a diagonal matrix for every A in the set. The following theorem characterizes simultaneously diagonalizable matrices: A set of diagonalizable matrices commutes if and only if the set is simultaneously diagonalizable.^{[1]:p. 64}

The set of all $n \times n$ diagonalizable matrices (over \mathbb{C}) with $n > 1$ is not simultaneously diagonalizable. For instance, the matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

are diagonalizable but not simultaneously diagonalizable because they do not commute.

A set consists of commuting normal matrices if and only if it is simultaneously diagonalizable by a unitary matrix; that is, there exists a unitary matrix U such that U^*AU is diagonal for every A in the set.

In the language of Lie theory, a set of simultaneously diagonalizable matrices generates a toral Lie algebra.

Examples

Diagonalizable matrices

- Involutions are diagonalizable over the reals (and indeed any field of characteristic not 2), with ± 1 on the diagonal.
- Finite order endomorphisms are diagonalizable over \mathbb{C} (or any algebraically closed field where the characteristic of the field does not divide the order of the endomorphism) with roots of unity on the diagonal. This follows since the minimal polynomial is separable, because the roots of unity are distinct.
- Projections are diagonalizable, with 0s and 1s on the diagonal.
- Real symmetric matrices are diagonalizable by orthogonal matrices; i.e., given a real symmetric matrix A , $Q^T A Q$ is diagonal for some orthogonal matrix Q . More generally, matrices are diagonalizable by unitary matrices if and only if they are normal. In the case of the real symmetric matrix, we see that $A = A^T$, so clearly $AA^T = A^T A$ holds. Examples of normal matrices are real symmetric (or skew-symmetric) matrices (e.g. covariance matrices) and Hermitian matrices (or skew-Hermitian matrices). See spectral theorems for generalizations to infinite-dimensional vector spaces.

Matrices that are not diagonalizable

In general, a rotation matrix is not diagonalizable over the reals, but all rotation matrices are diagonalizable over the complex field. Even if a matrix is not diagonalizable, it is always possible to "do the best one can", and find a matrix with the same properties consisting of eigenvalues on the leading diagonal, and either ones or zeroes on the superdiagonal – known as Jordan normal form.

Some matrices are not diagonalizable over any field, most notably nonzero nilpotent matrices. This happens more generally if the algebraic and geometric multiplicities of an eigenvalue do not coincide. For instance, consider

$$C = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

This matrix is not diagonalizable: there is no matrix U such that $U^{-1}CU$ is a diagonal matrix. Indeed, C has one eigenvalue (namely zero) and this eigenvalue has algebraic multiplicity 2 and geometric multiplicity 1.

Some real matrices are not diagonalizable over the reals. Consider for instance the matrix

$$B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

The matrix B does not have any real eigenvalues, so there is no **real** matrix Q such that $Q^{-1}BQ$ is a diagonal matrix. However, we can diagonalize B if we allow complex numbers. Indeed, if we take

$$Q = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix},$$

then $Q^{-1}BQ$ is diagonal. It is easy to find that B is the rotation matrix which rotates counterclockwise by angle $\theta = -\frac{\pi}{2}$

Note that the above examples show that the sum of diagonalizable matrices need not be diagonalizable.

How to diagonalize a matrix

Diagonalizing a matrix is the same process as finding its eigenvalues and eigenvectors, in the case that the eigenvectors form a basis. For example, consider the matrix

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

The roots of the characteristic polynomial $p(\lambda) = \det(\lambda I - A)$ are the eigenvalues $\lambda_1 = 1, \lambda_2 = 1, \lambda_3 = 2$. Solving the linear system $(I - A)\mathbf{v} = \mathbf{0}$ gives the eigenvectors $\mathbf{v}_1 = (1, 1, 0)$ and $\mathbf{v}_2 = (0, 2, 1)$, while $(2I - A)\mathbf{v} = \mathbf{0}$ gives $\mathbf{v}_3 = (1, 0, -1)$; that is, $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$ for $i = 1, 2, 3$. These vectors form a basis of $V = \mathbb{R}^3$, so we can assemble them as the column vectors of a change-of-basis matrix P to get:

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 & -2 \\ 0 & 1 & 0 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = D.$$

We may see this equation in terms of transformations: P takes the standard basis to the eigenbasis, $P\mathbf{e}_i = \mathbf{v}_i$, so we have:

$$P^{-1}AP\mathbf{e}_i = P^{-1}A\mathbf{v}_i = P^{-1}(\lambda_i\mathbf{v}_i) = \lambda_i\mathbf{e}_i,$$

so that $P^{-1}AP$ has the standard basis as its eigenvectors, which is the defining property of D .

Note that there is no preferred order of the eigenvectors in P ; changing the order of the eigenvectors in P just changes the order of the eigenvalues in the diagonalized form of A .^[2]

Application to matrix functions

Diagonalization can be used to efficiently compute the powers of a matrix $A = PDP^{-1}$:

$$\begin{aligned} A^k &= (PDP^{-1})^k = (PDP^{-1})(PDP^{-1}) \dots (PDP^{-1}) \\ &= PD(P^{-1}P)D(P^{-1}P) \dots (P^{-1}P)DP^{-1} = PD^kP^{-1}, \end{aligned}$$

and the latter is easy to calculate since it only involves the powers of a diagonal matrix. For example, for the matrix A with eigenvalues $\lambda = 1, 1, 2$ in the example above we compute:

$$\begin{aligned} A^k &= PD^kP^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1^k & 0 & 0 \\ 0 & 1^k & 0 \\ 0 & 0 & 2^k \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 2 - 2^k & -1 + 2^k & 2 - 2^{k+1} \\ 0 & 1 & 0 \\ -1 + 2^k & 1 - 2^k & -1 + 2^{k+1} \end{bmatrix}. \end{aligned}$$

This approach can be generalized to matrix exponential and other matrix functions that can be defined as power series. For example, defining $\exp(A) = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \dots$, we have:

$$\begin{aligned} \exp(A) &= P \exp(D) P^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} e^1 & 0 & 0 \\ 0 & e^1 & 0 \\ 0 & 0 & e^2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 2e - e^2 & -e + e^2 & 2e - 2e^2 \\ 0 & e & 0 \\ -e + e^2 & e - e^2 & -e + 2e^2 \end{bmatrix}. \end{aligned}$$

This is particularly useful in finding closed form expressions for terms of linear recursive sequences, such as the Fibonacci numbers.

Particular application

For example, consider the following matrix:

$$M = \begin{bmatrix} a & b - a \\ 0 & b \end{bmatrix}.$$

Calculating the various powers of M reveals a surprising pattern:

$$M^2 = \begin{bmatrix} a^2 & b^2 - a^2 \\ 0 & b^2 \end{bmatrix}, \quad M^3 = \begin{bmatrix} a^3 & b^3 - a^3 \\ 0 & b^3 \end{bmatrix}, \quad M^4 = \begin{bmatrix} a^4 & b^4 - a^4 \\ 0 & b^4 \end{bmatrix}, \quad \dots$$

The above phenomenon can be explained by diagonalizing M . To accomplish this, we need a basis of \mathbb{R}^2 consisting of eigenvectors of M . One such eigenvector basis is given by

$$\mathbf{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \mathbf{e}_1, \quad \mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{e}_1 + \mathbf{e}_2,$$

where \mathbf{e}_i denotes the standard basis of \mathbf{R}^n . The reverse change of basis is given by

$$\mathbf{e}_1 = \mathbf{u}, \quad \mathbf{e}_2 = \mathbf{v} - \mathbf{u}.$$

Straightforward calculations show that

$$M\mathbf{u} = a\mathbf{u}, \quad M\mathbf{v} = b\mathbf{v}.$$

Thus, a and b are the eigenvalues corresponding to \mathbf{u} and \mathbf{v} , respectively. By linearity of matrix multiplication, we have that

$$M^n \mathbf{u} = a^n \mathbf{u}, \quad M^n \mathbf{v} = b^n \mathbf{v}.$$

Switching back to the standard basis, we have

$$\begin{aligned} M^n \mathbf{e}_1 &= M^n \mathbf{u} = a^n \mathbf{e}_1, \\ M^n \mathbf{e}_2 &= M^n (\mathbf{v} - \mathbf{u}) = b^n \mathbf{v} - a^n \mathbf{u} = (b^n - a^n) \mathbf{e}_1 + b^n \mathbf{e}_2. \end{aligned}$$

The preceding relations, expressed in matrix form, are

$$M^n = \begin{bmatrix} a^n & b^n - a^n \\ 0 & b^n \end{bmatrix},$$

thereby explaining the above phenomenon.

Quantum mechanical application

In quantum mechanical and quantum chemical computations matrix diagonalization is one of the most frequently applied numerical processes. The basic reason is that the time-independent Schrödinger equation is an eigenvalue equation, albeit in most of the physical situations on an infinite dimensional Hilbert space.

A very common approximation is to truncate Hilbert space to finite dimension, after which the Schrödinger equation can be formulated as an eigenvalue problem of a real symmetric, or complex Hermitian matrix. Formally this approximation is founded on the variational principle, valid for Hamiltonians that are bounded from below.

First-order perturbation theory also leads to matrix eigenvalue problem for degenerate states.

See also

- Defective matrix
- Scaling (geometry)
- Triangular matrix
- Semisimple operator
- Diagonalizable group
- Jordan normal form
- Weight module – associative algebra generalization

- Orthogonal diagonalization

Notes

References

1. Horn, Roger A.; Johnson, Charles R. (2013). *Matrix Analysis, second edition*. Cambridge University Press. ISBN 9780521839402.
 2. Anton, H.; Rorres, C. (22 Feb 2000). *Elementary Linear Algebra (Applications Version)* (<http://archive.org/details/studentsolutions00grob>) (8th ed.). John Wiley & Sons. ISBN 978-0-471-17052-5.
-

Retrieved from "https://en.wikipedia.org/w/index.php?title=Diagonalizable_matrix&oldid=1190204176"

▪