

Diagonalizable matrix

From Wikipedia, the free encyclopedia

(Redirected from Matrix diagonalization)

In linear algebra, 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 $T: V \rightarrow V$ 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.^[1] A square matrix which is not diagonalizable is called *defective*.

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. Geometrically, a diagonalizable matrix is an **inhomogeneous dilation** (or *anisotropic scaling*) — it scales the space, as does a *homogeneous dilation*, but by a different factor in each direction, determined by the scale factors on each axis (diagonal entries).

Contents

- 1 Characterisation
- 2 Diagonalization
- 3 Simultaneous diagonalization
- 4 Examples
 - 4.1 Diagonalizable matrices
 - 4.2 Matrices that are not diagonalizable
 - 4.3 How to diagonalize a matrix
 - 4.3.1 Alternative Method
- 5 An application
 - 5.1 Particular application
- 6 Quantum mechanical application
- 7 See also
- 8 Notes
- 9 References
- 10 External links

Characterisation

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

- 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 .
- 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 .

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 . (Put in another way, a matrix is diagonalizable if and only if all of its elementary divisors are linear.)

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

- An n -by- 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. Let us 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

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

- 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 .

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

Diagonalization

If a matrix A can be diagonalized, that is,

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n \end{pmatrix},$$

then:

$$AP = P \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n \end{pmatrix}.$$

Writing P as a block matrix of its column vectors

$$P = (\vec{\alpha}_1 \quad \vec{\alpha}_2 \quad \cdots \quad \vec{\alpha}_n),$$

the above equation can be rewritten as

$$A\vec{\alpha}_i = \lambda_i\vec{\alpha}_i \quad (i = 1, 2, \dots, n).$$

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 the basis of 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 the matrix A is a Hermitian matrix (resp. symmetric matrix), eigenvectors of A can be chosen to form an orthonormal basis of \mathbf{C}^n (resp. \mathbf{R}^n). Under such circumstance P will be a unitary matrix (resp. orthogonal matrix) and P^{-1} equals the conjugate transpose (resp. transpose) of P .

Simultaneous diagonalization

See also: Simultaneous triangularisability and Weight (representation theory)

A set of matrices are said to be *simultaneously diagonalisable* 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 characterises simultaneously diagonalisable matrices: A set of diagonalizable matrices commutes if and only if the set is simultaneously diagonalisable.^[2]

The set of all n -by- n diagonalisable matrices (over \mathbf{C}) with $n > 1$ is not simultaneously diagonalisable. 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 diagonalisable 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 diagonalisable matrices generate a toral Lie algebra.

Examples

Diagonalizable matrices

- Involutions are diagonalisable over the reals (and indeed any field of characteristic not 2), with 1's and -1's on the diagonal
- Finite order endomorphisms are diagonalisable over the complex numbers (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 0's and 1's 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.

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.

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

How to diagonalize a matrix

Consider a matrix

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 0 \\ 2 & -4 & 2 \end{bmatrix}.$$

This matrix has eigenvalues

$$\lambda_1 = 3, \quad \lambda_2 = 2, \quad \lambda_3 = 1.$$

A is a 3-by-3 matrix with 3 different eigenvalues; therefore, it is diagonalizable. Note that if there are exactly n distinct eigenvalues in an $n \times n$ matrix then this matrix is diagonalizable.

These eigenvalues are the values that will appear in the diagonalized form of matrix A , so by finding the eigenvalues of A we have diagonalized it. We could stop here, but it is a good check to use the eigenvectors to diagonalize A .

The eigenvectors of A are

$$v_1 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad v_3 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}.$$

One can easily check that $Av_k = \lambda_k v_k$.

Now, let P be the matrix with these eigenvectors as its columns:

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

Note 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 .^[3]

Then P diagonalizes A , as a simple computation confirms, having calculated P^{-1} using any suitable method:

$$P^{-1}AP = \begin{bmatrix} 0 & -1 & 0 \\ 2 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 0 \\ 2 & -4 & 2 \end{bmatrix} \begin{bmatrix} -1 & 0 & -1 \\ -1 & 0 & 0 \\ 2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Note that the eigenvalues λ_k appear in the diagonal matrix.

Alternative Method

Starting with: $PD = AP$, where $P = [p_1, p_2]$, and the Diagonalization matrix D is:

$$D = \begin{bmatrix} d_{11} & 0 \\ 0 & d_{22} \end{bmatrix}.$$

Distribute A into the column vectors of P .

$$PD = [Ap_1, Ap_2].$$

Then D can be broken down to its column vectors as follows:

$$P[d_{11}\vec{e}_1, d_{22}\vec{e}_2] = [A\vec{p}_1, A\vec{p}_2].$$

Multiplying P on the left side of the equation gives:

$$[d_{11}\vec{p}_1, d_{22}\vec{p}_2] = [A\vec{p}_1, A\vec{p}_2].$$

Setting each entry of the matrix to its corresponding entry:

$$\begin{aligned} d_{11}\vec{p}_1 &= A\vec{p}_1 \\ d_{22}\vec{p}_2 &= A\vec{p}_2. \end{aligned}$$

Then the equations can be solved as follows, using the same process for both:

$$\begin{aligned} (d_{11}I)\vec{p}_1 &= A\vec{p}_1 \\ (d_{11}I - A)\vec{p}_1 &= 0 \end{aligned}$$

and it solves for d_{11} , which is the first entry in the diagonal matrix, and also the first eigenvalue.

An application

Diagonalization can be used to compute the powers of a matrix A efficiently, provided the matrix is diagonalizable. Suppose we have found that

$$P^{-1}AP = D$$

is a diagonal matrix. Then, as the matrix product is associative,

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

and the latter is easy to calculate since it only involves the powers of a diagonal matrix. This approach can be generalized to matrix exponential and other matrix functions since they can be defined as power series.

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 \mathbf{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 space (a 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. But also first-order perturbation theory for degenerate states leads to a matrix eigenvalue problem.

See also

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

Notes

- [^] Horn & Johnson 1985
- [^] Horn & Johnson 1985, pp. 51–53
- [^] Anton, H.; Rorres, C. (22 Feb 2000). *Elementary Linear Algebra (Applications Version)* (8th ed.). John Wiley & Sons. ISBN 978-0-471-17052-5.

References

- Horn, Roger A.; Johnson, Charles R. (1985). *Matrix Analysis*. Cambridge University Press. ISBN 978-0-521-38632-6.

External links

- Diagonalization (<http://planetmath.org/?op=getobj&from=objects&id=1960>), PlanetMath.org.

Retrieved from "http://en.wikipedia.org/w/index.php?title=Diagonalizable_matrix&oldid=549167038#Diagonalization"

Categories: Matrices

-
- This page was last modified on 7 April 2013 at 15:24.
 - Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy.
Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.