Eigenvalues and Eigenvectors

5.2 Diagonalization of a Matrix

5.2 Diagonalization of a Matrix

5C Suppose the n by n matrix A has n linearly independent eigenvectors. If these eigenvectors are the columns of a matrix S, then $S^{-1}AS$ is a diagonal matrix Λ. The eigenvalues of A are on the diagonal of Λ:

$$\textbf{Diagonalization} \qquad S^{-1}AS = \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}. \qquad (1)$$

We call S the "eigenvector matrix" and Λ the "eigenvalue matrix"—using a capital lambda because of the small lambdas for the eigenvalues on its diagonal.

Proof. Put the eigenvectors x_i in the columns of S, and compute AS by columns:

$$AS = A\begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \\ | & | & & | \end{bmatrix}.$$

Then the trick is to split this last matrix into a quite different product $S\Lambda$:

$$\begin{bmatrix} \\ \lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \\ \\ \end{bmatrix} = \begin{bmatrix} \\ x_1 & x_2 & \cdots & x_n \\ \\ \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

It is crucial to keep these matrices in the right order. If Λ came before S (instead of after), then λ_1 would multiply the entries in the first row. We want λ_1 to appear in the first column. As it is, $S\Lambda$ is correct. Therefore,

$$AS = S\Lambda, \quad \text{or} \quad S^{-1}AS = \Lambda, \quad \text{or} \quad A = S\Lambda S^{-1}. \tag{2}$$

S is invertible, because its columns (the eigenvectors) were assumed to be independent.

We add four remarks before giving any examples or applications. $\square$

Remark* 1.** If the matrix A has no repeated eigenvalues—the numbers $\lambda_1, \dots, \lambda_n$ are distinct—then its n eigenvectors are automatically independent (see 5D below). Therefore ***any matrix with distinct eigenvalues can be diagonalized.

***Remark* 2.** The diagonalizing matrix S is *not unique*. An eigenvector x can be multiplied by a constant, and remains an eigenvector. We can multiply the columns of S by any nonzero constants, and produce a new diagonalizing S. Repeated eigenvalues leave even more freedom in S. For the trivial example $A = I$, any invertible S will do: $S^{-1}IS$ is is always diagonal (Λ is just I). All vectors are eigenvectors of the identity.

***Remark* 3.** *Other matrices S will not produce a diagonal* Λ. Suppose the first column of S is y. Then the first column of $S\Lambda$ is $\lambda_1 y$. If this is to agree with the first column of AS, which by matrix multiplication is Ay, then y must be an eigenvector: $Ay = \lambda_1 y$. The *order* of the eigenvectors in S and the eigenvalues in Λ is automatically the same.

Remark 4. Not all matrices possess n linearly independent eigenvectors, so ***not all matrices are diagonalizable***. The standard example of a "defective matrix" is

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

Its eigenvalues are $\lambda_1 = \lambda_2 = 0$, since it is triangular with zeros on the diagonal:

$$\det(A - \lambda I) = \det \begin{bmatrix} -\lambda & 1 \\ 0 & -\lambda \end{bmatrix} = \lambda^2.$$

All eigenvectors of this A are multiples of the vector $(1,0)$:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \text{or} \qquad x = \begin{bmatrix} c \\ 0 \end{bmatrix}.$$

$\lambda = 0$ is a double eigenvalue—its *algebraic multiplicity* is 2. But the *geometric multiplicity* is 1—there is only one independent eigenvector. We can't construct S.

Here is a more direct proof that this A is not diagonalizable. Since $\lambda_1 = \lambda_2 = 0$, Λ would have to be the zero matrix, But if $\Lambda = S^{-1}AS = 0$, then we premultiply by S and postmultiply by S^{-1}, to deduce falsely that $A = 0$. There is no invertible S.

That failure of diagonalization was ***not*** a result of $\lambda = 0$. It came from $\lambda_1 = \lambda_2$:

Repeated eigenvalues $$A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}.$$

Their eigenvalues are 3, 3 and 1, 1. They are not singular! The problem is the shortage of eigenvectors—which are needed for S. That needs to be emphasized:

Diagonalizability of** A **depends on enough eigenvectors.
Invertibility of** A **depends on nonzero eigenvalues.

Diagonalization can fail only if there are repeated eigenvalues. Even then, it does not always fail. $A = I$ has repeated eigenvalues $1, 1, \ldots, 1$ but it is already diagonal! There is no shortage of eigenvectors in that case.

5D If eigenvectors $x_1, \ldots, x_k$ correspond to *different eigenvalues* $\lambda_1, \ldots, \lambda_k$, then those eigenvectors are linearly independent.

A matrix with n distinct eigenvalues can be diagonalized. This is the typical case.

Examples of Diagonalization

The main point of this section is $S^{-1}AS = A$. The eigenvector matrix S converts A into its eigenvalue matrix Λ (diagonal). We see this for projections and rotations.

Example 1. The projection $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ has eigenvalue matrix $\Lambda = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right]$. The eigenvectors go into the columns of S:

$$S = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad \text{and} \qquad AS = S\Lambda = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

That last equation can be verified at a glance. Therefore $S^{-1}AS = \Lambda$.

Powers and Products: A^k and AB

There is one more situation in which the calculations are easy. ***The eigenvalue of A^2 are exactly $\lambda_1^2, \ldots, \lambda_n^2$, and every eigenvector of A is also an eigenvector of A^2.*** We start from $Ax = \lambda x$, and multiply again by A:

$$A^2x = A\lambda x = \lambda Ax = \lambda^2 x. \tag{3}$$

Thus λ^2 is an eigenvalue of A^2, with the same eigenvector x. If the first multiplication by A leaves the direction of x unchanged, then so does the second.

The same result comes from diagonalization, by squaring $S^{-1}AS = \Lambda$:

$$\textbf{Eigenvalues of } A^2 \qquad (S^{-1}AS)(S^{-1}AS) = \Lambda^2 \quad \text{or} \quad S^{-1}A^2S = \Lambda^2.$$

The matrix A^2 is diagonalized by the same S, so the eigenvectors are unchanged. The eigenvalues are squared. This continues to hold for any power of A:

5E The eigenvalues of A^k are $\lambda_1^k, \ldots, \lambda_n^k$, and each eigenvector of A is still an eigenvector of A^k. When S diagonalizes A, it also diagonalizes A^k:

$$\Lambda^k = (S^{-1}AS)(S^{-1}AS)\cdots(S^{-1}AS) = S^{-1}A^kS. \tag{4}$$

Each S^{-1} cancels an S, except for the first S^{-1} and the last S.

If A is invertible this rule also applies to its inverse (the power $k = -1$). ***The eigenvalues of*** A^{-1} ***are*** $1/\lambda_i$. That can be seen even without diagonalizing:

$$\textit{if} \quad Ax = \lambda x \quad \textit{then} \quad x = \lambda A^{-1}x \quad \textit{and} \quad \frac{1}{\lambda}x = A^{-1}x.$$

5F Diagonalizable matrices share the same eigenvector matrix S if and only if $AB = BA$.

Proof. If the same S diagonalizes both $A = S\Lambda_1 S^{-1}$ and $B = S\Lambda_2 S^{-1}$, we can multiply in either order:

$$AB = S\Lambda_1 S^{-1} S\Lambda_2 S^{-1} = S\Lambda_1\Lambda_2 S^{-1} \quad \text{and} \quad BA = S\Lambda_2 S^{-1} S\Lambda_1 S^{-1} = S\Lambda_2\Lambda_1 S^{-1}.$$

Since $\Lambda_1\Lambda_2 = \Lambda_2\Lambda_1$ (diagonal matrices always commute) we have $AB = BA$.

Problem Set 5.2

1. Factor the following matrices into $S\Lambda S^{-1}$:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \qquad \text{and} \qquad A = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}.$$

2. Find the matrix A whose eigenvalues are 1 and 4, and whose eigenvectors are $\left[\begin{smallmatrix}3\\1\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}2\\1\end{smallmatrix}\right]$, respectively. (*Hint*: $A = S\Lambda S^{-1}$.)

4. If a 3 by 3 upper triangular matrix has diagonal entries 1, 2, 7, how do you know it can be diagonalized? What is Λ?

5. Which of these matrices cannot be diagonalized?

$$A_1 = \begin{bmatrix} 2 & -2 \\ 2 & -2 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 2 & 0 \\ 2 & -2 \end{bmatrix} \qquad A_3 = \begin{bmatrix} 2 & 0 \\ 2 & 2 \end{bmatrix}.$$

17. If A has $\lambda_1 = 2$ with eigenvector $x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\lambda_2 = 5$ with $x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, use $S\Lambda S^{-1}$ to find A. No other matrix has the same λ's and x's.

19. True or false: If the n columns of S (eigenvectors of A) are independent, then

(a) A is invertible.

(b) A is diagonalizable.

(c) S is invertible.

(d) S is diagonalizable.

25. True or false: If the eigenvalues of A are 2, 2, 5, then the matrix is certainly

(a) invertible.

(b) diagonalizable.

(c) not diagonalizable.

26. If the eigenvalues of A are 1 and 0, write everything you know about the matrices A and A^2.

27. Complete these matrices so that $\det A = 25$. Then trace $= 10$, and $\lambda = 5$ is repeated! Find an eigcnvector with $Ax = 5x$. These matrices will nothe diagonalizabie because there is no second line of eigenvectors.

$$A = \begin{bmatrix} 8 & \\ & 2 \end{bmatrix}, \qquad A = \begin{bmatrix} 9 & 4 \\ & 1 \end{bmatrix}, \quad \text{and} \quad A = \begin{bmatrix} 10 & 5 \\ -5 & \end{bmatrix}.$$

29. $A^k = S\Lambda^k S^{-1}$ approaches the zero matrix as $k \to \infty$ if and only if every λ has absolute value less than ____. Does $A^k \to 0$ or $B^k \to 0$?

$$A = \begin{bmatrix} .6 & .4 \\ .4 & .6 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} .6 & .9 \\ .1 & .6 \end{bmatrix}.$$

31. Find Λ and S to diagonalize B in Problem 29. What is $B^{10}u_0$ for these u_0?

$$u_0 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \qquad u_0 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \quad \text{and} \quad u_0 = \begin{bmatrix} 6 \\ 0 \end{bmatrix}.$$

Positive Definite Matrices

6.2 Tests for Positive Definiteness

6.2 Tests for Positive Definiteness

Which symmetric matrices have the property that $x^{\mathrm{T}}Ax > 0$ for all nonzero vectors x? There are four or five different ways to answer this question, and we hope to find all of them. The previous section began with some hints about the signs of eigenvalues. but that gave place to the tests on a, b, c:

$$b = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \quad \text{is positive definite when} \quad a > 0 \quad \text{and} \quad ac - b^2 > 0.$$

From those conditions, ***both eigenvalues are positive***. Their product $\lambda_1\lambda_2$ is determinant $ac - b^2 > 0$, so the eigenvalues are either both positive or both negative. They must be positive because their sum is the trace $a + c > 0$.

Looking at a and $ac-b^2$, it is even possible to spot the appearance of the ***pivots***. They turned up when we decomposed $x^{\mathrm{T}}Ax$ into a sum of squares:

$$\textbf{Sum of squares} \qquad ax^2+2bxy+cy^2 = a\left(x+\frac{b}{a}y\right)^2+\frac{ac-b^2}{a}y^2. \tag{1}$$

Those coefficients a and $(ac-b^2)/a$ are the pivots for a 2 by 2 matrix. For larger matrices the pivots still give a simple test for positive definiteness: $x^{\mathrm{T}}Ax$ stays positive when n independent squares are multiplied by ***positive pivots***.

One more preliminary remark. The two parts of this hook were linked by the chapter on determinants. Therefore we ask what part determinants play. ***It is not enough to require that the determinant of A is positive***. If $a=c=-1$ and $b=0$. then $\det A=1$ but $A=-I=$ negative definite. The determinant test is applied not only to A itself, giving $ac-b^2>0$, but also to the 1 by 1 submatrix a in the upper left-hand corner.

The natural generalization will involve all n of the *upper left submatrices* of A:

$$A_1 = \begin{bmatrix} a_{11} \end{bmatrix}, \qquad A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \qquad A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \cdots, \qquad A_n = A.$$

Here is the main theorem on positive definiteness, and a reasonably detailed proof:

6B Each of the following tests is a necessary and sufficient condition for the real symmetric matrix A to be ***positive definite***:

(I) $x^{\mathrm{T}}kx > 0$ for all nonzero real vectors x.

(II) All the eigenvalues of A satisfy $\lambda_i > 0$.

(III) All the upper left submatrices A_k have positive determinants.

(IV) All the pivots (without row exchanges) satisfy $d_k > 0$.

elimination on a symmetric matrix: $A = LDL^{\mathrm{T}}$.

Example 1. Positive pivots 2, $\frac{3}{2}$, and $\frac{4}{3}$:

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{bmatrix} \begin{bmatrix} 2 & & \\ & \frac{3}{2} & \\ & & \frac{4}{3} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{bmatrix} = LDL^{\mathrm{T}}.$$

$$\textbf{Determinant test} \qquad \det A_1 = 2, \quad \det A_2 = 3, \quad \det A_3 = \det A = 4.$$

The pivots are the ratios $d_1 = 2$, $d_2 = \frac{3}{2}$, $d_3 = \frac{4}{3}$. Ordinarily the eigenvalue test is the longest computation. For this A we know the λ's are all positive:

$$\textbf{Eigenvalue test} \qquad \lambda_1 = 2 - \sqrt{2}, \quad \lambda_2 = 2, \quad \lambda_3 = 2 + \sqrt{2}.$$

Each test is enough by itself.

The rectangular matrix will be R

as we now show—provided that the n columns of R are linearly independent:

6C The symmetric matrix A is positive definite if and only if

(V) There is a matrix R with independent columns such that $A = R^{\mathrm{T}}R$.

Semidefinite Matrices

The tests for semidefiniteness will relax $x^{\mathrm{T}}Ax > 0$, $\lambda > 0$, $d > 0$, and $\det > 0$, to allow zeros to appear. The main point is to see the analogies with the positive definite case.

6D Each of the following tests is a necessary and sufficient condition for a symmetric matrix A to be ***positive semidefinite***:

(I′) $x^{\mathrm{T}}Ax \geq 0$ for all vectors x (this defines positive semidefinite).

(II′) All the eigenvalues of A satisfy $\lambda_i \geq 0$.

(III′) No principal submatrices have negative determinants.

(IV′) No pivots are negative.

(V′) There is a matrix R, possibly with dependent columns, such that $A = R^{\mathrm{T}}R$.

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \text{ is positive semidefinite, and } \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \text{ is negative semidefinite.}$$

A row exchange comes with the same column exchange to maintain symmetry.

Example 2.

$$A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \quad \text{is positive } \textit{semi}\text{definite, by all five tests:}$$

(I′) $x^{\mathrm{T}}Ax = (x_1 - x_2)^2 + (x_1 - x_3)^2 + (x_2 - x_3)^2 \geq 0$ (zero if $x_1 = x_2 = x_3$).

(II′) The eigenvalues are $\lambda_1 = 0$, $\lambda_2 = \lambda_3 = 3$ (a zero eigenvalue).

(III′) $\det A = 0$ and smaller determinants are positive.

(IV′) $$A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{3}{2} & -\frac{3}{2} \\ 0 & -\frac{3}{2} & \frac{3}{2} \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 0 & \mathbf{0} \end{bmatrix} \quad \text{(missing pivot).}$$

(V′) $A = R^{\mathrm{T}}R$ with dependent columns in R:

$$\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \quad (1,1,1) \text{ in the nullspace.}$$

For $A = \begin{bmatrix} 4 & & \\ & 1 & \\ & & \frac{1}{9} \end{bmatrix}$, the equation is $x^{\mathrm{T}}Ax = 4x_1^2 + x_2^2 + \frac{1}{9}x_3^2$

$A = \left[\begin{smallmatrix} 5 & 4 \\ 4 & 5 \end{smallmatrix}\right]$ and $x^{\mathrm{T}}Ax = 5u^2 + 8uv + 5v^2$

6G For any symmetric matrix A, ***the signs of the pivots agree with the signs of the eigenvalues***. The eigenvalue matrix Λ and the pivot matrix D have the same number of positive entries, negative entries, and zero entries.

Problem Set 6.2

1. For what range of numbers a and b are the matrices A and B positive definite?

$$A = \begin{bmatrix} a & 2 & 2 \\ 2 & a & 2 \\ 2 & 2 & a \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 & 4 \\ 2 & b & 8 \\ 4 & 8 & 7 \end{bmatrix}.$$

2. Decide for or against the positive definiteness of

$$A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}^2 .$$

4. Show from the eigenvalues that if A is positive definite, so is A^2 and so is A^{-1}.

7. If $A = Q\Lambda Q^{\mathrm{T}}$ is symmetric positive definite, then $R = Q\sqrt{\Lambda}Q^{\mathrm{T}}$ is its *symmetric positive definite square root*. Why does R have positive eigenvalues? Compute R and verify $R^2 = A$ for

$$A = \begin{bmatrix} 10 & 6 \\ 6 & 10 \end{bmatrix} \qquad \text{and} \qquad A = \begin{bmatrix} 10 & -6 \\ -6 & 10 \end{bmatrix}.$$

19. Which 3 by 3 symmetric matrices A produce these functions $f = x^{\mathrm{T}}Ax$? Why is the first matrix positive definite but not the second one?

(a) $f = 2(x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_2x_3)$.

(b) $f = 2(x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_1x_3 - x_2x_3)$.

23. Give a quick reason why each of these statements is true:

(a) Every positive definite matrix is invertible.

(b) The only positive definite projection matrix is $P = I$.

(c) A diagonal matrix with positive diagonal entries is positive definite.

(d) A symmetric matrix with a positive determinant might not be positive definite!

30. Without multiplying $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$, find

(a) the determinant of A.
(b) the eigenvalues of A.
(c) the eigenvectors of A.
(d) a reason why A is symmetric positive definite.