Diagonalization of a Matrix (矩阵的对角化)

Lecture 22

Dept. of Math.

2023.11

Diagonalization of a Matrix

Introduction

We start right off with the one essential computation. It is perfectly simple and will be used in every section of this chapter. The eigenvectors diagonalize a matrix:

Theorem

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:

$$S^{-1}AS = \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

Diagonalization of a Matrix

We call S the "eigenvector matrix" and Λ the "eigenvalue matrix".

There are four remarks before giving any examples or applications.

1: Any matrix with distinct eigenvalues can be diagonalized. If the matrix A has no repeated eigenvalues–the numbers $\lambda_1, \lambda_2, \cdots, \lambda_n$ are distinct–then its n eigenvectors are automatically independent. Therefore any matrix with distinct eigenvalues can be diagonalized.

2: The diagonalizing matrix is not unique. An eigenvector x can be multiplied by a constant, and remains an eigenvector.

Remarks

3: Other matrices S will not produce a diagonal Λ.

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

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

$\lambda = 0$ is a double eigenvalue–its algebraic multiplicity is 2. But the geometric multiplicity is 1–there is only independent eigenvector. We can not construct S.

Diagonalizability and Invertibility

There is **no** connection between diagonalizability and invertibility. Diagonalization can fail **only if** there are **repeated** eigenvalues. The problem is the shortage of eigenvectors–which are needed for S. That needs to be emphasized:

Proposition

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

Algebraic multiplicity and geometric multiplicity.

Theorem

Theorem

If eigenvectors $x_1, x_2, \cdots, x_k$ correspond to different eigenvalues $\lambda_1, \cdots, \lambda_k$, then those eigenvectors are linearly independent.

Proof.

Suppose first that $k = 2$, and that some combination of x_1 and x_2 produces zero: $c_1 x_1 + c_2 x_2 = 0$. Multiplying by A, we find $c_1 \lambda_1 x_1 + c_2 \lambda_2 x_2 = 0$. Subtracting λ_2 times the previous equation, the vector x_2 disappears: $c_1(\lambda_1 - \lambda_2) x_1 = 0$. Since $\lambda_1 \neq \lambda_2$ and $x_1 \neq 0$, we are forced into $c_1 = 0$. Similarly $c_2 = 0$, and the two vectors are independent; only the trivial combination gives zero. The same argument extends to any number of vectors. $\square$

Examples of Diagonalization

The main point of this section is $S^{-1}AS = \Lambda$. 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 $\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

Example 2

Example 2. The eigenvalues and eigenvectors of a rotation matrix. 90° rotation:

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

How can a vector be rotated and still have its direction unchanged?

Imaginary numbers? Complex numbers are needed even for real matrices! See section 5.5 for more!

Powers and Products: A^k and AB

The eigenvalues of A^2 are exactly $\lambda_1^2, \cdots, \lambda_n^2$, and every eigenvector of A is also an eigenvector of A^2.

Theorem

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

$$\Lambda^k = S^{-1}A^kS$$

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 without diagonalizing:

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

Example 3

Example 3. If K is rotation through 90°, then K^2 is rotation through 180°(which means $-I$) and K^{-1} is rotation through -90°:

$$K = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},\ K^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \text{and } K^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

The eigenvalues of K are i and $-i$; their squares are -1 and -1; their reciprocals are $1/i = -i$ and $1/(-i) = i$. Then K^4 is a complete rotation through 360°.

$$K^4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and also } \Lambda^4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Notice also that complex numbers are needed even for real matrices.

Product

- For a product of two matrices, we can ask about the eigenvalues of AB–but we won't get a good answer.
- In general, if A has an eigenvalue λ and B has an eigenvalue μ, AB does not have $\lambda\mu$ as its eigenvalue. For instance, we have two matrices with zero eigenvalues, while AB has $\lambda = 1$:

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

  The eigenvectors of this A and B are completely different, which is typical.
- Similarly, the eigenvalues of $A+B$ generally have nothing to do with $\lambda+\mu$.

Theorem

If the eigenvector is the same for A and B, then the eigenvalues multiply and AB has the eigenvalue $\mu\lambda$. But there is something more important.

There is an easy way to recognize when A and B share a full set of eigenvectors, and that is a key question in quantum mechanics:

Theorem

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

Proof

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:

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

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

Proof

In the opposite direction, suppose $AB = BA$. Starting from $Ax = \lambda x$, we have

$$ABx = BAx = B\lambda x = B(\lambda x) = \lambda Bx.$$

Thus x and Bx are both eigenvectors of A, sharing the same λ. If we assume for convenience that the eigenvalues of A are distinct–the eigenspaces are all one-dimensional–then Bx must be a multiple of x. In other words, x is an eigenvector of B as well as A. The proof with repeated eigenvalues is a little longer(left as an exercise).

Heisenberg's uncertainty principle

Heisenberg's uncertainty principle comes from non-commuting matrices, like position P and momentum Q. Position is symmetric, momentum is skew-symmetric, and together they satisfy:

$$QP - PQ = I.$$

The uncertainty principle follows directly from the Schwarz inequality $(Qx)^T(Px) \leq ||Qx||||Px||$ of Section 3.2:

$$||x||^2 = x^T x = x^T(QP - PQ)x \leq 2||Qx||||Px||$$

The product of $||Qx||/||x||$ and $||Px||/||x||$– momentum and position errors, when the wave function is x–is at least $\frac{1}{2}$. It is impossible to get both errors small, because when you try to measure the position of a particle you change its momentum.

Final Note

At the end we come back to $A = S\Lambda S^{-1}$.

- That factorization is particularly suited to take powers of A, and the simplest case A^2 makes the point.
- The LU factorization is hopeless when squared, but $S\Lambda S^{-1}$ is perfect. The square is $S\Lambda^2 S^{-1}$, and the eigenvectors are unchanged.
- By following those eigenvectors we will solve difference equations and differential equations.

Homework Assignment 22

5.2: 1, 2, 7, 8, 12, 13, 19, 23, 32, 40.