Linear Algebra

Diagonalisation and Symmetric Matrices

Introduction to Linear Algebra by Gilbert Strang, http://pi.math.cornell.edu/~jerison/math2940/real-eigenvalues.pdf

When $\boldsymbol{x}$ is an eigenvector, multiplication by A is just multiplication by a single number: $A\boldsymbol{x} = \lambda\boldsymbol{x}$. All the difficulties of matrices are swept away.

6D Diagonalization Suppose the n by n matrix A has n linearly independent eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. Put them into the columns of an ***eigenvector matrix*** S. Then $S^{-1}AS$ is the ***eigenvalue matrix*** Λ:

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

Proof Multiply A times its eigenvectors, which are the columns of S. The first column of AS is $A\boldsymbol{x}_1$. That is $\lambda_1\boldsymbol{x}_1$. Each column of S is multiplied by its eigenvalue:

$$AS = A\begin{bmatrix} & & \\ \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_n \\ & & \end{bmatrix} = \begin{bmatrix} & & \\ \lambda_1\boldsymbol{x}_1 & \cdots & \lambda_n\boldsymbol{x}_n \\ & & \end{bmatrix}.$$

The trick is to split this matrix AS into S times Λ:

$$\begin{bmatrix} & & \\ \lambda_1\boldsymbol{x}_1 & \cdots & \lambda_n\boldsymbol{x}_n \\ & & \end{bmatrix} = \begin{bmatrix} & & \\ \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_n \\ & & \end{bmatrix}\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = S\Lambda.$$

Keep those matrices in the right order! Then λ_1 multiplies the first column $\boldsymbol{x}_1$, as shown. The diagonalization is complete, and we can write $AS = S\Lambda$ in two good ways:

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

The matrix S has an inverse, because its columns (the eigenvectors of A) were assumed to be linearly independent. *Without n independent eigenvectors, we can't diagonalize.*

Example 1 The projection matrix $P = \begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix}$ has $\lambda = 1$ and 0. Put the eigenvectors (1, 1) and (−1, 1) into S. Then $S^{-1}PS$ is the eigenvalue matrix Λ:

$$\underset{S^{-1}}{\begin{bmatrix} .5 & .5 \\ -.5 & .5 \end{bmatrix}} \underset{P}{\begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix}} \underset{S}{\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}} \underset{=}{=} \underset{\Lambda}{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}$$

Remark 1 Suppose the numbers $\lambda_1, \ldots, \lambda_n$ are all different. Then it is automatic that the eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are independent. Therefore *any matrix that has no repeated eigenvalues can be diagonalized.*

Remark 2 The eigenvector matrix S is not unique. ***We can multiply eigenvectors by any nonzero constants.*** Suppose we multiply the columns of S by 5 and −1. Divide the rows of S^{-1} by 5 and −1 to find the new inverse:

$$S_{\text{new}}^{-1} P S_{\text{new}} = \begin{bmatrix} .1 & .1 \\ .5 & -.5 \end{bmatrix} \begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix} \begin{bmatrix} 5 & 1 \\ 5 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \text{ same } \Lambda.$$

The extreme case is $A = I$, when every vector is an eigenvector. Any invertible matrix S can be the eigenvector matrix. Then $S^{-1}IS = I$ (which is Λ).

Remark 4 (repeated warning for repeated eigenvalues) Some matrices have too few eigenvectors. *Those matrices are not diagonalizable.* Here are two examples:

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

Their eigenvalues happen to be 0 and 0. Nothing is special about $\lambda = 0$—it is the repetition of λ that counts. All eigenvectors of the second matrix are multiples of $(1, 0)$!

$$A\boldsymbol{x} = 0\boldsymbol{x} \quad \text{means} \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{x} = c \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

There is no second eigenvector, so the unusual matrix A cannot be diagonalized.

- ***Invertibility*** is concerned with the ***eigenvalues*** (zero or not).
- ***Diagonalizability*** is concerned with the ***eigenvectors*** (too few or enough).

6E (Independent x from different λ) Eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j$ that correspond to distinct (all different) eigenvalues are linearly independent. An n by n matrix that has n different eigenvalues (no repeated λ's) must be diagonalizable.

Example 2 The Markov matrix $A = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}$ in the last section had $\lambda_1 = 1$ and $\lambda_2 = .5$. Here is $A = S\Lambda S^{-1}$ with those eigenvalues in Λ:

$$\begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} = \begin{bmatrix} .6 & 1 \\ .4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & .5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ .4 & -.6 \end{bmatrix} = S\Lambda S^{-1}.$$

The eigenvectors $(.6, .4)$ and $(1, -1)$ are in the columns of S. They are also the eigenvectors of A^2, because $A^2\boldsymbol{x} = A\lambda\boldsymbol{x} = \lambda^2\boldsymbol{x}$. Then A^2 has the same S, and ***the eigenvalue matrix of A^2 is Λ^2***:

$$\boxed{A^2 = S\Lambda S^{-1}S\Lambda S^{-1} = S\Lambda^2 S^{-1}.}$$

Just keep going, and you see why the high powers A^k approach a "steady state":

$$A^k = S\Lambda^k S^{-1} = \begin{bmatrix} .6 & 1 \\ .4 & -1 \end{bmatrix} \begin{bmatrix} 1^k & 0 \\ 0 & (.5)^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ .4 & -.6 \end{bmatrix}.$$

As k gets larger, $(.5)^k$ gets smaller. In the limit it disappears completely. That limit is

$$A^\infty = \begin{bmatrix} .6 & 1 \\ .4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ .4 & -.6 \end{bmatrix} = \begin{bmatrix} .6 & .6 \\ .4 & .4 \end{bmatrix}.$$

The limit has the eigenvector $\boldsymbol{x}_1$ in both columns. We saw this A^∞ on the very first page of the chapter. Now we see it more quickly from powers like $A^{100} = S\Lambda^{100}S^{-1}$.

Symmetric Matrices

What is special about $Ax = \lambda x$ when A is symmetric? We are looking for special properties of the eigenvalues λ and the eigenvectors $\boldsymbol{x}$ when $A = A^{\mathrm{T}}$.

The diagonalization $A = S\Lambda S^{-1}$ will reflect the symmetry of A. We get some hint by transposing to $A^{\mathrm{T}} = (S^{-1})^{\mathrm{T}}\Lambda S^{\mathrm{T}}$. Those are the same since $A = A^{\mathrm{T}}$. Possibly S^{-1} in the first form equals S^{T} in the second form. Then $S^{\mathrm{T}}S = I$. That makes each eigenvector in S orthogonal to the other eigenvectors. The key facts get first place in the Table at the end of this chapter, and here they are:

1. A symmetric matrix has only ***real eigenvalues***.
2. The ***eigenvectors*** can be chosen ***orthonormal***.

Those orthonormal eigenvectors go into the columns of S. There are n of them (independent because they are orthonormal). Every symmetric matrix can be diagonalized.

Its eigenvector matrix S becomes an orthogonal matrix Q. Orthogonal matrices have $Q^{-1} = Q^{\mathrm{T}}$—what we suspected about S is true. To remember it we write $S = Q$, when we choose orthonormal eigenvectors.

Why do we use the word "choose"? Because the eigenvectors do not *have* to be unit vectors. Their lengths are at our disposal. We will choose unit vectors—eigenvectors of length one, which are orthonormal and not just orthogonal. Then $A = S\Lambda S^{-1}$ is in its special and particular form $Q\Lambda Q^{\mathrm{T}}$ for symmetric matrices:

6H (Spectral Theorem) Every symmetric matrix has the factorization $A = Q\Lambda Q^{\mathrm{T}}$ with real eigenvalues in Λ and orthonormal eigenvectors in Q:

$$A = Q\Lambda Q^{-1} = Q\Lambda Q^{\mathrm{T}} \quad \text{with} \quad Q^{-1} = Q^{\mathrm{T}}.$$

6J Orthogonal Eigenvectors Eigenvectors of a real symmetric matrix (when they correspond to different λ's) are always perpendicular.

A has real eigenvalues and n real orthogonal eigenvectors if and only if $A = A^{\mathrm{T}}$.

Proof Suppose $A\boldsymbol{x} = \lambda_1\boldsymbol{x}$ and $A\boldsymbol{y} = \lambda_2\boldsymbol{y}$ and $A = A^{\mathrm{T}}$. Take dot products of the first equation with $\boldsymbol{y}$ and the second with $\boldsymbol{x}$:

$$(\lambda_1\boldsymbol{x})^{\mathrm{T}}\boldsymbol{y} = (A\boldsymbol{x})^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}A\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}\lambda_2\boldsymbol{y}. \tag{5}$$

The left side is $\boldsymbol{x}^{\mathrm{T}}\lambda_1\boldsymbol{y}$, the right side is $\boldsymbol{x}^{\mathrm{T}}\lambda_2\boldsymbol{y}$. Since $\lambda_1 \neq \lambda_2$, this proves that $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} = 0$. The eigenvector $\boldsymbol{x}$ (for λ_1) is perpendicular to the eigenvector $\boldsymbol{y}$ (for λ_2).

Note: If a **symmetric matrix** has any **repeated eigenvalues**, it is still possible to determine a full set of mutually orthogonal eigenvectors, but not every full set of eigenvectors will have the orthogonality property.

Symmetric matrices have real eigenvalues

Recall that if $z = a + bi$ is a complex number, its complex conjugate is defined by $\overline{z} = a - bi$. We have $z\overline{z} = (a + bi)(a - bi) = a^2 + b^2$, so $z\overline{z}$ is always a nonnegative real number (and equals 0 only when $z = 0$). It is also true that if w, z are complex numbers, then $\overline{wz} = \overline{w}\,\overline{z}$.

Let $\mathbf{v}$ be a vector whose entries are allowed to be complex. It is no longer true that $\mathbf{v} \cdot \mathbf{v} \geq 0$ with equality only when $\mathbf{v} = \mathbf{0}$. For example,

$$\begin{bmatrix} 1 \\ i \end{bmatrix} \cdot \begin{bmatrix} 1 \\ i \end{bmatrix} = 1 + i^2 = 0.$$

However, if $\overline{\mathbf{v}}$ is the complex conjugate of $\mathbf{v}$, it is true that $\overline{\mathbf{v}} \cdot \mathbf{v} \geq 0$ with equality only when $\mathbf{v} = \mathbf{0}$. Indeed,

$$\begin{bmatrix} a_1 - b_1 i \\ a_2 - b_2 i \\ \vdots \\ a_n - b_n i \end{bmatrix} \cdot \begin{bmatrix} a_1 + b_1 i \\ a_2 + b_2 i \\ \vdots \\ a_n + b_n i \end{bmatrix} = (a_1^2 + b_1^2) + (a_2^2 + b_2^2) + \cdots + (a_n^2 + b_n^2)$$

which is always nonnegative and equals zero only when all the entries a_i and b_i are zero.

With this in mind, suppose that λ is a (possibly complex) eigenvalue of the real symmetric matrix A. Thus there is a nonzero vector $\mathbf{v}$, also with complex entries, such that $A\mathbf{v} = \lambda\mathbf{v}$. By taking the complex conjugate of both sides, and noting that $\overline{A} = A$ since A has real entries, we get $\overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}} \Rightarrow A\overline{\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}$. Then, using that $A^T = A$,

$$\overline{\mathbf{v}}^T A\mathbf{v} = \overline{\mathbf{v}}^T (A\mathbf{v}) = \overline{\mathbf{v}}^T (\lambda\mathbf{v}) = \lambda(\overline{\mathbf{v}} \cdot \mathbf{v}),$$
$$\overline{\mathbf{v}}^T A\mathbf{v} = (A\overline{\mathbf{v}})^T \mathbf{v} = (\overline{\lambda}\overline{\mathbf{v}})^T \mathbf{v} = \overline{\lambda}(\overline{\mathbf{v}} \cdot \mathbf{v}).$$

Since $\mathbf{v} \neq \mathbf{0}$, we have $\overline{\mathbf{v}} \cdot \mathbf{v} \neq 0$. Thus $\lambda = \overline{\lambda}$, which means $\lambda \in \mathbf{R}$.

Example 1 Find the λ's and $\boldsymbol{x}$'s when $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ and $A - \lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix}$.

Solution The equation $\det(A - \lambda I) = 0$ is $\lambda^2 - 5\lambda = 0$. The eigenvalues are 0 and 5 (*both real*). We can see them directly: $\lambda = 0$ is an eigenvalue because A is singular, and $\lambda = 5$ is the other eigenvalue so that $0 + 5$ agrees with $1 + 4$. This is the *trace* down the diagonal of A.

Two eigenvectors are $(2, -1)$ and $(1, 2)$—orthogonal but not yet orthonormal. The

These eigenvectors have length $\sqrt{5}$. Divide them by $\sqrt{5}$ to get unit vectors. Put those into the columns of S (which is Q). Then $Q^{-1}AQ$ is Λ and $Q^{-1} = Q^{\mathrm{T}}$:

$$Q^{-1}AQ = \frac{\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \frac{\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}}{\sqrt{5}} = \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} = \Lambda.$$

Example 2 Find the λ's and $\boldsymbol{x}$'s for this symmetric matrix with trace zero:

$$A = \begin{bmatrix} -3 & 4 \\ 4 & 3 \end{bmatrix} \quad \text{has} \quad \det(A - \lambda I) = \begin{vmatrix} -3-\lambda & 4 \\ 4 & 3-\lambda \end{vmatrix} = \lambda^2 - 25.$$

The roots of $\lambda^2 - 25 = 0$ are $\lambda_1 = 5$ and $\lambda_2 = -5$ (both real). The eigenvectors $\boldsymbol{x}_1 = (1, 2)$ and $\boldsymbol{x}_2 = (-2, 1)$ are perpendicular. To make them into unit vectors, divide by their lengths $\sqrt{5}$. The new $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are the columns of Q, and Q^{-1} equals Q^{T}:

$$A = Q\Lambda Q^T = \frac{\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}}{\sqrt{5}} \begin{bmatrix} 5 & 0 \\ 0 & -5 \end{bmatrix} \frac{\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}}{\sqrt{5}}.$$

$A^T = A$

$$A = Q \Lambda Q^T$$

$$= \begin{bmatrix} | & | & \\ q_1 & q_2 & \cdots \\ | & | & \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \end{bmatrix} \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \end{bmatrix} = \lambda_1 q_1 q_1^T + \lambda_2 q_2 q_2^T + \cdots$$

Every symm matrix is
a comb of perp. projection matrices