22. Diagonalizing Symmetric Matrices

Symmetric matrices have many applications. For example, if we consider the shortest distance between pairs of important cities, we might get a table like this:

	Davis	Seattle	San Francisco
Davis	0	2000	80
Seattle	2000	0	2010
San Francisco	80	2010	0

Encoded as a matrix, we obtain:

$$M = \begin{pmatrix} 0 & 2000 & 80 \\ 2000 & 0 & 2010 \\ 80 & 2010 & 0 \end{pmatrix} = M^T.$$

Definition A matrix is *symmetric* if it obeys

$$M = M^T.$$

One very nice property of symmetric matrices is that they always have real eigenvalues. The general proof is an exercise, but here's an example for 2×2 matrices.

Example For a general symmetric 2×2 matrix, we have:

$$\begin{aligned} P_\lambda \begin{pmatrix} a & b \\ b & d \end{pmatrix} &= \det \begin{pmatrix} \lambda - a & -b \\ -b & \lambda - d \end{pmatrix} \\ &= (\lambda - a)(\lambda - d) - b^2 \\ &= \lambda^2 - (a+d)\lambda - b^2 + ad \\ \Rightarrow \lambda &= \frac{a+d}{2} \pm \sqrt{b^2 + \left(\frac{(a-d)}{2}\right)^2}. \end{aligned}$$

Notice that the discriminant $4b^2 + (a-d)^2$ is always positive, so that the eigenvalues must be real.

Now, suppose a symmetric matrix M has two distinct eigenvalues $\lambda \neq \mu$ and eigenvectors x and y:

$$Mx = \lambda x, \qquad My = \lambda y.$$

Consider the dot product $x \cdot y = x^T y = y^T x$. And now calculate:

$$\begin{aligned} x^T M y &= x^T \mu y = \mu x \cdot y, \text{ and} \\ x^T M y &= (y^T M x)^T \text{ (by transposing a } 1 \times 1 \text{ matrix)} \\ &= x^T M^T y \\ &= x^T M y \\ &= x^T \lambda y \\ &= \lambda x \cdot y. \end{aligned}$$

Subtracting these two results tells us that:

$$0 = x^T M y - x^T M y = (\mu - \lambda)\, x \cdot y.$$

Since μ and λ were assumed to be distinct eigenvalues, $\lambda - \mu$ is non-zero, and so $x \cdot y = 0$. Then we have proved the following theorem.

Theorem. *Eigenvectors of a symmetric matrix with distinct eigenvalues are orthogonal.*

Example The matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ has eigenvalues determined by

$$\det(M - \lambda) = (2 - \lambda)^2 - 1 = 0.$$

Then the eigenvalues of M are 3 and 1, and the associated eigenvectors turn out to be $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. It is easily seen that these eigenvectors are orthogonal:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0$$

Last lecture we saw that the matrix P built from orthonormal basis vectors $\{v_1, \ldots, v_n\}$

$$P = \begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix}$$

was an orthogonal matrix:

$$P^{-1} = P^T, \text{ or } PP^T = I = P^T P.$$

Moreover, given any (unit) vector x_1, one can always find vectors x_2, …, x_n such that $\{x_1, \ldots, x_n\}$ is an orthonormal basis. (Such a basis can be obtained using the "Gram-Schmidt" procedure, which we will present later.) Hyperlink to notes24!

Now suppose M is a symmetric $n \times n$ matrix and λ_1 is an eigenvalue with eigenvector x_1. Let the square matrix of column vectors P be the following:

$$P = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix},$$

where x_1 through x_n are orthonormal, and x_1 is an eigenvector for M, but the others are not necessarily eigenvectors for M. Then

$$MP = \begin{pmatrix} \lambda_1 x_1 & M x_2 & \dots & M x_n \end{pmatrix}.$$

But P is an orthogonal matrix, so $P^{-1} = P^T$. Then:

$$\begin{aligned}
P^{-1} = P^T &= \begin{pmatrix} x_1^T \\ \vdots \\ x_n^T \end{pmatrix} \\
\Rightarrow P^T M P &= \begin{pmatrix} x_1^T \lambda_1 x_1 & * & \dots & * \\ x_2^T \lambda_1 x_1 & * & \dots & * \\ \vdots & & & \vdots \\ x_n^T \lambda_1 x_1 & * & \dots & * \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & * & & \vdots \\ 0 & * & \dots & * \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \hat{M} & \\ 0 & & & \end{pmatrix}
\end{aligned}$$

The last equality follows since $P^T M P$ is symmetric. The asterisks in the matrix are where "stuff" happens; this extra information is denoted by $\hat{M}$ in the final equation. We know nothing about $\hat{M}$ except that it is an $(n-1) \times (n-1)$ matrix and that it is symmetric. But then, by finding an (unit) eigenvector for $\hat{M}$, we could repeat this procedure successively. The end result would be a diagonal matrix with eigenvalues of M on the diagonal. Then we have proved a theorem.

Theorem. *Every symmetric matrix is similar to a diagonal matrix of its eigenvalues. In other words,*

$$M = M^T \Rightarrow M = PDP^T$$

where P is an orthogonal matrix and D is a diagonal matrix whose entries are the eigenvalues of M.

To diagonalize a real symmetric matrix, begin by building an orthogonal matrix from an orthonormal basis of eigenvectors.

Example The symmetric matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ has eigenvalues 3 and 1 with eigenvectors $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ respectively. From these eigenvectors, we normalize and build the orthogonal matrix:

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}$$

Notice that $P^T P = I_2$. Then:

$$MP = \begin{pmatrix} \frac{3}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{3}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}.$$

In short, $MP = DP$, so $D = P^T MP$. Then D is the diagonalized form of M and P the associated change-of-basis matrix from the standard basis to the basis of eigenvectors.

References

- Hefferon, Chapter Three, Section V: Change of Basis

Wikipedia:

- Symmetric Matrix
- Diagonalizable Matrix
- Similar Matrix

Review Questions

1. (On Reality of Eigenvectors)

   i. Suppose $z = x + iy$ where $x, y \in \mathbb{R}, i = \sqrt{-1}$, and $\overline{z} = x - iy$. Compute $\overline{\overline{z}}$. What can you say about $z\overline{z}$ and $\overline{z}z$? This operation is called *complex conjugation*.

   ii. What can you say about complex numbers λ that obey $\lambda = \overline{\lambda}$?

   iii. Let $x = \begin{pmatrix} z^1 \\ \vdots \\ z^n \end{pmatrix} \in \mathbb{C}^n$. Let $x^\dagger = \begin{pmatrix} \overline{z^1} & \dots & \overline{z^n} \end{pmatrix} \in \mathbb{C}^n$. Compute $x^\dagger x$. What can you say about the result?

   iv. Suppose $M = M^T$ is an $n \times n$ symmetric matrix with real entries. Let λ be an eigenvalue of M with eigenvector x, so $Mx = \lambda x$. Compute:
   $$\frac{x^\dagger M x}{x^\dagger x}$$

   v. Suppose Λ is a 1×1 matrix. What is Λ^T?

   vi. What is the size of the matrix $x^\dagger M x$?

   vii. For any matrix (or vector) N, we can compute $\overline{N}$ by applying complex conjugation to each entry of N. Compute $\overline{(x^\dagger)^T}$. Then compute $\overline{(x^\dagger M x)^T}$.

   viii. Show that $\lambda = \overline{\lambda}$. What does this say about λ?

2. Let $x_1 = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, where $a^2 + b^2 + c^2 = 1$. Find vectors x_2 and x_3 such that $\{x_1, x_2, x_3\}$ is an orthonormal basis for $\mathbb{R}^3$.

3. What can you say about the sum of the dimensions of the eigenspaces of a real symmetric matrix?