Chapter 6: Eigenvalues and Eigenvectors

6.1. Introduction to Eigenvalues

In this unit (as in the last one), all matrices are square. Suppose A is an $n \times n$ matrix, so that premultiplication by it takes n-entry vectors to other n-entry vectors. For at least some matrices, some vectors are special, i.e., multiplication by A just takes them to scalar multiples of themselves:

Definition. Given the $n \times n$ matrix A, if $A\boldsymbol{v} = \lambda\boldsymbol{v}$ for some scalar λ and nonzero vector $\boldsymbol{v}$, then λ is an *eigenvalue* of A corresponding to the *eigenvector* $\boldsymbol{v}$. It is easy to verify that, for a fixed λ, the set of all vectors $\boldsymbol{v}$ in $\mathbb{R}^n$ that $A\boldsymbol{v} = \lambda\boldsymbol{v}$ (i.e., all the eigenvectors corresponding to λ and the zero vector) constitute a subspace of $\mathbb{R}^n$, the *eigenspace* corresponding to λ.

Example. Suppose $A = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$, a diagonal matrix. Then the eigenvalues of A are -1 and 2, with corresponding eigenvectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ respectively. The corresponding eigenspaces are the x_1- and x_2-axes respectively.

Example. Take a plane through the origin in $\mathbb{R}^3$, and let P be the orthogonal projection matrix onto that plane. Then the plane is the eigenspace of P corresponding to the eigenvalue 1 — vectors in the plane are unchanged by P — and the orthogonal complement, i.e., the line through the origin perpendicular to the plane, is the eigenspace of P corresponding to the eigenvalue 0 — vectors in that line are taken to the zero vector, i.e., 0 times themselves.

Here is the fact that gives us a way to find eigenvalues:

$$\begin{aligned} \lambda \text{ is an eigenvalue of } A &\iff A\boldsymbol{v} = \lambda\boldsymbol{v} \text{ for some } \boldsymbol{v} \neq \boldsymbol{0} \\ &\iff (A - \lambda I)\boldsymbol{v} = \boldsymbol{0} \text{ for some } \boldsymbol{v} \neq \boldsymbol{0} \\ &\iff A - \lambda I \text{ is singular} \\ &\iff \det(A - \lambda I) = 0 \end{aligned}$$

So the roots of $\det(A - \lambda I)$, a polynomial of degree n called the *characteristic polynomial* of A, are the eigenvalues of A. To find the corresponding eigenvectors, we solve the equations $(A - \lambda I)\boldsymbol{v} = \boldsymbol{0}$ for the values of λ that we just found. Because the constant terms are all zeros, we won't bother to write them; we will just write the coefficient matrix and do elementary row operations to it. We know that there will be at least one row of zeros when we get to row echelon form, so it turns out to be less work than it might have been:

Example.

$$A = \begin{bmatrix} -7 & 1 & 0 \\ 1 & -3 & 2 \\ 0 & 2 & -7 \end{bmatrix} : \quad \begin{vmatrix} -7-\lambda & 1 & 0 \\ 1 & -3-\lambda & 2 \\ 0 & 2 & -7-\lambda \end{vmatrix}$$

$$\begin{aligned} &= (-7-\lambda)^2(-3-\lambda) + 0 + 0 - 0 - 4(-7-\lambda) - 1(-7-\lambda) \\ &= (-7-\lambda)[21 + 10\lambda + \lambda^2 - 5] \\ &= (-7-\lambda)(\lambda+2)(\lambda+8) \\ &= 0 \text{ when } \lambda = -7, -2, -8 \end{aligned}$$

$$\lambda = -7 \;:\quad \begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 2 \\ 0 & 2 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} : \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

$$\lambda = -2 \ : \quad \begin{bmatrix} -5 & 1 & 0 \\ 1 & -1 & 2 \\ 0 & 2 & -5 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1/2 \\ 0 & 1 & -5/2 \\ 0 & 0 & 0 \end{bmatrix} \ : \ \begin{bmatrix} 1/2 \\ 5/2 \\ 1 \end{bmatrix}$$

$$\lambda = -8 \ : \quad \begin{bmatrix} 1 & 1 & 0 \\ 1 & 5 & 2 \\ 0 & 2 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1/2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 0 \end{bmatrix} \ : \ \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \end{bmatrix}$$

Now I claim that the three resulting eigenvectors are linearly independent. How can I be sure? Because of the following fact:

Proposition. *Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ are eigenvectors of the (square) matrix A corresponding to different eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$. Then the $\boldsymbol{v}$'s are independent.*

Proof. We assume they are dependent, and we reach a contradiction. (In theory courses we would just say, "Assume by way of contradiction — BWOC for short — that they are dependent.") We may suppose without loss of generality (WLOG for short) that $\boldsymbol{v}_k$ is the first one that is a linear combination of the earlier ones. (Because $\boldsymbol{v}_1$ is an eigenvector, it is not zero, so $k > 1$.) So there are scalars c_i for which

$$\boldsymbol{v}_k = c_1\boldsymbol{v}_1 + \cdots + c_{k-1}\boldsymbol{v}_{k-1} \ .$$

Let's first multiply that equation on both sides by the scalar λ_k, and then by the matrix A, and then subtract the results. We get:

$$\begin{aligned} \lambda_k\boldsymbol{v}_k &= c_1\lambda_k\boldsymbol{v}_1 + \cdots + c_{k-1}\lambda_k\boldsymbol{v}_{k-1} \\ \lambda_k\boldsymbol{v}_k &= c_1\lambda_1\boldsymbol{v}_1 + \cdots + c_{k-1}\lambda_{k-1}\boldsymbol{v}_{k-1} \\ \boldsymbol{0} &= c_1(\lambda_k - \lambda_1)\boldsymbol{v}_1 + \cdots + c_{k-1}(\lambda_k - \lambda_{k-1})\boldsymbol{v}_{k-1} \end{aligned}$$

Now because $\boldsymbol{v}_k$ was nonzero, at least one of the c_i's is nonzero, and all the $(\lambda_k - \lambda_i)$'s are nonzero; so if, say, c_{k-1} was the last nonzero c_i, then we could divide by $c_{k-1}(\lambda_k - \lambda_{k-1})$ and write $\boldsymbol{v}_{k-1}$ in terms of the earlier $\boldsymbol{v}_i$'s. But that contradicts our choice of $\boldsymbol{v}_k$ as the first one for which we could do that. So the assumption must be false: The $\boldsymbol{v}$'s must be independent. □

That's how I know the three vectors in the example above are independent. And because there are three of them in $\mathbb{R}^3$, they form a basis for $\mathbb{R}^3$. And why is that good? Well, diagonal matrices are particularly easy to work with (think of finding powers of them, for example), and this result means we can "make A into a diagonal matrix":

Proposition. *Given an $n \times n$ matrix A, suppose there is a basis $\mathcal{B}$ of $\mathbb{R}^n$ consisting of eigenvectors of A. Form the matrix S with columns the vectors in $\mathcal{B}$, and the diagonal matrix Λ with main diagonal entries the eigenvalues of A corresponding to the vectors in $\mathcal{B}$ (in order). Then $A = S\Lambda S^{-1}$.*

Proof. Because the columns of S are a basis for $\mathbb{R}^n$, S is invertible; and we actually want to show $\Lambda = S^{-1}AS$. In other words, we want to show that the columns of $S^{-1}AS$ are the columns $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ of the identity matrix times the respective eigenvalues $\lambda_1, \ldots, \lambda_n$. Now the i-th column of $S^{-1}AS$ is $S^{-1}AS\boldsymbol{e}_i$. To evaluate this, start on the right: $S\boldsymbol{e}_i$ is the i-th column in S, i.e., the eigenvector in $\mathcal{B}$ corresponding to the eigenvalue λ_i. So $AS\boldsymbol{e}_i = \lambda_i S\boldsymbol{e}_i$; and then

$$S^{-1}AS\boldsymbol{e}_i = S^{-1}(\lambda_i S\boldsymbol{e}_i) = \lambda_i(S^{-1}S\boldsymbol{e}_i) = \lambda_i e_i \ .$$

So we really do have $S^{-1}AS = \Lambda$, as we claimed. □

Example. (continued)

$$\begin{bmatrix} -7 & 1 & 0 \\ 1 & -3 & 2 \\ 0 & 2 & -7 \end{bmatrix} = \begin{bmatrix} -2 & 1/2 & 1/2 \\ 0 & 5/2 & -1/2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -7 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -8 \end{bmatrix} \begin{bmatrix} -2 & 1/2 & 1/2 \\ 0 & 5/2 & -1/2 \\ 1 & 1 & 1 \end{bmatrix}^{-1} .$$