Math 201 lecture for Friday, Week 9

Eigenvectors and eigenvalues

Definition. Let $f\colon V \to V$ be a linear transformation of a vector space V over F. A *nonzero* vector $v \in V$ is an *eigenvector* for f with *eigenvalue* $\lambda \in F$ if

$$f(v) = \lambda v.$$

If $A \in M_{n\times n}(F)$, a *nonzero* vector $v \in F^n$ is an *eigenvector* for A with *eigenvalue* $\lambda \in F$ if

$$Av = \lambda v.$$

Thus, eigenvectors and eigenvalues for A are the same as eigenvectors and eigenvalues for the associated linear function $f_A\colon F^n \to F^n$ (defined by $f_A(v) = Av$).

Here is why we like eigenvectors: Suppose that $\alpha = \langle v_1, \dots, v_n\rangle$ is an ordered basis of eigenvectors for $f\colon V \to V$ with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$, i.e., $f(v_i) = \lambda_i v_i$ for $i = 1, \dots, n$. Then the matrix $[f]_\alpha^\alpha$ representing f with respect to the basis α for the domain and codomain is the diagonal matrix $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

Example. Let

$$A = \begin{pmatrix} -1 & 2 \\ -6 & 6 \end{pmatrix}.$$

with corresponding linear function

$$\begin{aligned} f_A\colon \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x, y) &\mapsto (-x + 2y, -6x + 6y). \end{aligned}$$

It turns out that $(2, 3)$ and $(1, 2)$ are eigenvectors for f_A with eigenvalues 2 and 3, respectively:

$$\begin{pmatrix} -1 & 2 \\ -6 & 6 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix} = 2\begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

$$\begin{pmatrix} -1 & 2 \\ -6 & 6 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Find the matrix representing f_A with respect to the ordered basis

$$\alpha = \langle (2, 3), (1, 2)\rangle.$$

To do this we write the image of each vector in α as a linear combination of the vectors in α and pull off the coefficients to create columns:

$$\begin{aligned} f_A(2, 3) &= 2(2, 3) = 2 \cdot (2, 3) + 0 \cdot (1, 2) \\ f_A(1, 2) &= 3(1, 2) = 0 \cdot (2, 3) + 3 \cdot (1, 2). \end{aligned}$$

Hence,

$$[f_A]_\alpha^\alpha = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = \operatorname{diag}(2, 3).$$

That is the point: a basis of eigenvectors gives a matrix representative that is diagonal, which is the simplest type of matrix to think about. Let's think abstractly about what just happened. The matrix A is the matrix representing f_A with respect to the standard basis, and the matrix

$$D = \mathrm{diag}(2,3) = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

represents f_A with respect to the basis α. Let ϕ_α be the mapping that takes coordinates with respect to α. We get the commutative diagram:

$$\begin{tikzcd}
(1,2) \arrow[rrrrr, bend left=10] \arrow[ddddd, bend right=10] & & & & & 3(1,2) \arrow[ddddd, bend left=10] \\
& (2,3) \arrow[rrr, bend left=10] \arrow[ddd, bend right=10] & & & 2(2,3) \arrow[ddd, bend left=10] & \\
& & \mathbb{R}^2 \arrow[r, "A"] \arrow[d, "\phi_\alpha"', "\wr"] & \mathbb{R}^2 \arrow[d, "\phi_\alpha", "\wr"'] & & \\
& & \mathbb{R}^2 \arrow[r, "D"] & \mathbb{R}^2. & & \\
& (1,0) \arrow[rrr, bend right=10] & & & 2(1,0) & \\
(0,1) \arrow[rrrrr, bend right=10] & & & & & 3(0,1)
\end{tikzcd}$$

Reviewing something we talked about earlier in the semester: The matrix for ϕ_α would be a bit of a chore to write down. It's j-column would be the image of e_j. So we would have to write each e_j as a linear combination of the basis vectors in α. However, the inverse of ϕ_α is *easy* to write down. Take a look at the commutative diagram. By construction of ϕ_α, we have

$$\phi_\alpha^{-1}(1,0) = (2,3) \quad \text{and} \quad \phi_\alpha^{-1}(0,1) = (1,2).$$

So the matrix for ϕ_α^{-1} is

$$P = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}.$$

So the matrix for ϕ_α is P^{-1}. Therefore, another way to write the commutative diagram is

$$\begin{tikzcd}
\mathbb{R}^2 \arrow[r, "A"] \arrow[d, "P^{-1}"', "\wr"] & \mathbb{R}^2 \arrow[d, "P^{-1}", "\wr"'] \\
\mathbb{R}^2 \arrow[r, "D"] & \mathbb{R}^2.
\end{tikzcd}$$

From contemplating this diagram, we see that

$$D = P^{-1}AP.$$

Summary: having found eigenvectors $(2,3)$ and $(1,2)$, we place those eigenvectors as columns in a matrix P, and then $P^{-1}AP$ is a diagonal matrix with the corresponding eigenvalues on the diagonal.

To generalize:

Let $A \in M_{n\times n}(F)$ with corresponding linear function $f_A\colon F^n \to F^n$. Suppose $\alpha = \langle v_1, \ldots, v_n\rangle$ is an ordered basis of eigenvectors with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$, i.e., $Av_i = \lambda_i v_i$ for $i = 1, \ldots, n$. Let P be the matrix whose columns are $v_1, \ldots, v_n$. Then

$$P^{-1}AP = D,$$

where $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, and we have a commutative diagram

$$\begin{tikzcd} F^n \arrow[r, "A"] \arrow[d, "P^{-1}"', "\wr"] & F^n \arrow[d, "P^{-1}", "\wr"'] \\ F^n \arrow[r, "D"] & F^n. \end{tikzcd}$$

How does one find eigenvectors and eigenvalues? Let $A \in M_{n\times n}(F)$ with corresponding function $f_A\colon F^n \to F^n$ (so $f_A(v) := Av$). We are looking for a nonzero vector $v \in F^n$ and a scalar λ such that $Av = \lambda v$. To achieve that, the following argument is of **central importance**:

$$Av = \lambda v \quad \Leftrightarrow \quad (A - \lambda I_n)v = 0 \quad \Leftrightarrow \quad v \in \ker(A - \lambda v).$$

This says that:

$\lambda \in F$ **is an eigenvalue for** A **if and only if** $\ker(A - \lambda I_n) \neq \{0\}$.

So we would like to determine those λ for which the kernel of $A - \lambda I_n$ is nontrivial, for which the following is key:

$$\ker(A - \lambda I_n) \neq \{0\} \quad \Leftrightarrow \quad \operatorname{rank}(A - \lambda I_n) < n \quad \Leftrightarrow \quad \det(A - \lambda I_n) = 0.$$

Let's apply this to the matrix A in our example:

$$\begin{aligned} \det\left(\begin{pmatrix} -1 & 2 \\ -6 & 6 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) &= \det\left(\begin{pmatrix} -1 & 2 \\ -6 & 6 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right) \\ &= \det\begin{pmatrix} -1-\lambda & 2 \\ -6 & 6-\lambda \end{pmatrix} \\ &= (-1-\lambda)(6-\lambda) - 2(-6) \\ &= \lambda^2 - 5\lambda + 6 \\ &= (\lambda - 2)(\lambda - 3). \end{aligned}$$

Thus, $\ker(A - \lambda I_n) \neq \{0\}$ if and only if $\lambda = 2, 3$. So the eigenvalues for A are 2 and 3.

Having found the eigenvalues, **how do we go about finding corresponding eigenvalues**? For each eigenvalue λ, there are nonzero elements $\ker(A - \lambda I_n)$. So we just apply our algorithm for finding the kernel of a matrix:

$\boxed{\lambda = 2}$

$$A - 2I_2 = \begin{pmatrix} -1 & 2 \\ -6 & 6 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ -6 & 4 \end{pmatrix}.$$

So we need to find $(x, y) \in \mathbb{R}^2$ satisfying

$$\begin{pmatrix} -3 & 2 \\ -6 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Therefore, we perform Gaussian elimination:

$$\begin{pmatrix} -3 & 2 \\ -6 & 4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -\frac{2}{3} \\ 0 & 0 \end{pmatrix}.$$

Hence,

$$\ker(A - 2I_2) = \left\{ \left(\frac{2}{3} y, y \right) : y \in \mathbb{R} \right\}.$$

For a basis we could take $\left(\frac{2}{3}, 1\right)$, or easier, $(2, 3)$.

Similarly for the other eigenvalue:

$\boxed{\lambda = 3}$

$$\begin{aligned} A - 3I_2 &= \begin{pmatrix} -1 & 2 \\ -6 & 6 \end{pmatrix} - \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \\ &= \begin{pmatrix} -4 & 2 \\ -6 & 3 \end{pmatrix} \\ &\rightsquigarrow \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Hence,

$$\ker(A - 3I_2) = \left\{ \left(\frac{1}{2} y, y \right) : y \in \mathbb{R} \right\}.$$

For a basis we could take $\left(\frac{1}{2}, 1\right)$, or easier, $(1, 2)$.

Let $A \in M_{n \times n}(F)$. The eigenvalues for A are exactly the solutions λ to

$$\det(A - \lambda I_n) = 0.$$

If $\lambda \in F$ is an eigenvalue, it corresponding eigenvectors are the nonzero vectors in

$$\ker(A - \lambda I_n).$$

Use our algorithm to find a basis for the matrix $A - \lambda I_n$.