KEY FACTS ABOUT EIGENVECTORS

Let A be an $n \times n$ matrix. Let $f(k)$ be the characteristic polynomial $\det(A - k \cdot \mathrm{Id})$. In this note, we will establish the following facts:

Big Fact 1. *For a real number λ, there is a (nonzero) eigenvector of A with eigenvalue λ if and only if $f(\lambda) = 0$.*

Big Fact 2. *If A has r linearly independent eigenvectors $v_1, v_2, \dots, v_r$, all with eigenvalue λ, then $(k-\lambda)^r$ divides $f(k)$.*

Given a number λ, the ***geometric multiplicity*** of λ is defined to be the dimension of the space of vectors $\vec{v}$ for which $A\vec{v} = \lambda\vec{v}$. The ***algebraic multiplicity*** is the exponent of $k - \lambda$ which occurs in the factorization of $f(k)$.

So Big Fact 1 says that the algebraic multiplicity is positive if and only if the geometric multiplicity is positive. Big Fact 2 says that algebraic multiplicity is greater than or equal to geometric multiplicity.

Big Fact 3. *Let $\lambda_1, \lambda_2, \dots, \lambda_r$ be distinct zeroes of f, with $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r$ corresponding eigenvectors. Then $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r$ are linearly independent.*

We deduce some consequences about when there is a basis of eigenvectors.

Big Fact 4. *If there is a basis of eigenvectors of A, then $f(k)$ factors as $(\lambda_1 - k)(\lambda_2 - k)\cdots(\lambda_n - k)$. If k has such a factorization, and all the λ_i are distinct, then there is a basis of eigenvectors.*

Big Fact 1

For a nonzero vector $\vec{v}$, we have $A\vec{v} = \lambda\vec{v}$ if and only if $(A - \lambda \cdot \mathrm{Id})\vec{v} = 0$, if and only if $\vec{v}$ is in the kernel of $A - \lambda \cdot \mathrm{Id}$.

So there is a nonzero vector $\vec{v}$ with $A\vec{v} = \lambda\vec{v}$ if and only if the kernel of $A - \lambda \cdot \mathrm{Id}$ is nonzero. By our previous discussion of determinants, this happens precisely when $\det(A - \lambda \cdot \mathrm{Id}) = 0$.

Big Fact 2

Take the linearly independent vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r$ and find $n - r$ more vectors so that $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r, \vec{v}_{r+1}, \dots, \vec{v}_n$ is a basis of $\mathbb{R}^n$. (I should justify that you can always do this, but I'll gloss over that.)

For $1 \leq i \leq r$, we have $A\vec{v}_i = \lambda\vec{v}_i$. For $r + 1 \leq i \leq n$, we can write

$$A\vec{v}_i = \sum_{j=1}^{n} c_{ij}\vec{v}_j$$

for some coefficients c_{ij}. We can organize these facts into a matrix. For concreteness, we take $r = 2$ and $n = 5$.

$$\begin{pmatrix} | & | & | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 & \vec{v}_5 \\ | & | & | & | & | \end{pmatrix} \begin{pmatrix} \lambda & & c_{31} & c_{41} & c_{51} \\ & \lambda & c_{32} & c_{42} & c_{52} \\ & & c_{33} & c_{43} & c_{53} \\ & & c_{34} & c_{44} & c_{54} \\ & & c_{35} & c_{45} & c_{55} \end{pmatrix} = A \cdot \begin{pmatrix} | & | & | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 & \vec{v}_5 \\ | & | & | & | & | \end{pmatrix}$$

The blank entries in the matrix are zero.

Let

$$S = \begin{pmatrix} | & | & | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 & \vec{v}_5 \\ | & | & | & | & | \end{pmatrix}$$

So our equation above says

$$S \begin{pmatrix} \lambda & & c_{31} & c_{41} & c_{51} \\ & \lambda & c_{32} & c_{42} & c_{52} \\ & & c_{33} & c_{43} & c_{53} \\ & & c_{34} & c_{44} & c_{54} \\ & & c_{35} & c_{45} & c_{55} \end{pmatrix} = AS.$$

Since $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r, \vec{v}_{r+1}, \ldots, \vec{v}_n$ is a basis, the matrix S is invertible and we can write

$$S \begin{pmatrix} \lambda & & c_{31} & c_{41} & c_{51} \\ & \lambda & c_{32} & c_{42} & c_{52} \\ & & c_{33} & c_{43} & c_{53} \\ & & c_{34} & c_{44} & c_{54} \\ & & c_{35} & c_{45} & c_{55} \end{pmatrix} S^{-1} = A.$$

So

$$\det(A - k \cdot \mathrm{Id}) = \det(A - kSS^{-1}) = \det\ S \left(\begin{pmatrix} \lambda & & c_{31} & c_{41} & c_{51} \\ & \lambda & c_{32} & c_{42} & c_{52} \\ & & c_{33} & c_{43} & c_{53} \\ & & c_{34} & c_{44} & c_{54} \\ & & c_{35} & c_{45} & c_{55} \end{pmatrix} - k \cdot \mathrm{Id} \right) S^{-1}$$

$$= (\det S) \cdot \det \begin{pmatrix} \lambda - k & & c_{31} & c_{41} & c_{51} \\ & \lambda - k & c_{32} & c_{42} & c_{52} \\ & & c_{33} - k & c_{43} & c_{53} \\ & & c_{34} & c_{44} - k & c_{54} \\ & & c_{35} & c_{45} & c_{55} - k \end{pmatrix} \cdot (\det S^{-1})$$

The $\det S$ and $\det S^{-1}$ factors cancel so the characteristic polynomial is

$$\det \begin{pmatrix} \lambda - k & & c_{31} & c_{41} & c_{51} \\ & \lambda - k & c_{32} & c_{42} & c_{52} \\ & & c_{33} - k & c_{43} & c_{53} \\ & & c_{34} & c_{44} - k & c_{54} \\ & & c_{35} & c_{45} & c_{55} - k \end{pmatrix}$$

Using the zeroes in the left r columns, we see that every term of this determinant is divisible by $(\lambda - k)^r$.

Big Fact 3

The proof of Big Fact 3 is a bit challenging, and I imagine I won't cover it in class. But it is a nice result!

Our proof is by induction on r. The base case $r = 1$ is trivial. To build confidence, let's also check $r = 2$. We have $A\vec{v}_1 = \lambda_1 \vec{v}_1$ and $A\vec{v}_2 = \lambda_2 \vec{v}_2$. If $\vec{v}_1$ were proportional to $\vec{v}_2$, then A would stretch both $\vec{v}_1$ and $\vec{v}_2$ by the same factor. But it stretches $\vec{v}_1$ by λ_1 and $\vec{v}_2$ by λ_2, and we said that $\lambda_1 \neq \lambda_2$. So, in fact, $\vec{v}_1$ and $\vec{v}$ are not proportional after all, which is what we wanted.

Let's now do the case of larger r. Suppose we already know that $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{r-1}$ are linearly independent. Suppose that we have a nontrivial linear relation

$$c_1 \vec{v}_1 + c_2 \vec{v}_2 + \cdots c_{r-1} \vec{v}_{r-1} + c_r \vec{v}_r = 0. \qquad (*)$$

Multiplying equation $(*)$ by A, we get

$$c_1 A\vec{v}_1 + c_2 A\vec{v}_2 + \cdots c_{r-1} A\vec{v}_{r-1} + c_r A\vec{v}_r = 0$$

or

$$c_1 \lambda_1 \vec{v}_1 + c_2 \lambda_2 \vec{v}_2 + \cdots c_{r-1} \lambda_{r-1} \vec{v}_{r-1} + c_r \lambda_r \vec{v}_r = 0$$

Now subtract off λ_r times equation $(*)$:

$$(c_1 \lambda_1 \vec{v}_1 + c_2 \lambda_2 \vec{v}_2 + \cdots c_{r-1} \lambda_{r-1} \vec{v}_{r-1} + c_r \lambda_r \vec{v}_r) - \lambda_r \left(c_1 \vec{v}_1 + c_2 \vec{v}_2 + \cdots c_{r-1} \vec{v}_{r-1} + c_r \vec{v}_r \right) = 0$$

or

$$c_1 (\lambda_1 - \lambda_r) \vec{v}_1 + c_2 (\lambda_2 - \lambda_r) \vec{v}_2 + \cdots + c_{r-1} (\lambda_1 - \lambda_{r-1}) \vec{v}_{r-1} = 0. \qquad (\dagger)$$

Equation (†) is a linear relation among $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{r-1}$. Our induction assumption was that we had already showed this couldn't happen. At this point, we have basically reached a contradiction, and the rest is just mopping up.

The only way we might not have a contradiction is if all the coefficients in (†) were zero. In other words, if $c_1(\lambda_1 - \lambda_r) = c_2(\lambda_2 - \lambda_r) = \cdots = c_{r-1}(\lambda_1 - \lambda_r) = 0$. Since λ_r is not equal to any of the other λ's, this would mean $c_1 = c_2 = \cdots = c_{r-1} = 0$.

Plugging into equation $(*)$, we get $c_r \vec{v}_r = 0$. But $\vec{v}_r$ is not zero, so this means $c_r = 0$. We have now shown that all the coefficients in equation $(*)$ are 0. In other words, the only linear relation between the $\vec{v}_i$ is a trivial relation.

Now, we have a contradiction, and the proof is complete.

Big Fact 4

First, suppose that there is a basis $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ of eigenvectors. Then, as in the argument for Big Fact 4, we deduce that

$$\begin{pmatrix} | & | & | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 & \vec{v}_5 \\ | & | & | & | & | \end{pmatrix} \begin{pmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \lambda_3 & & \\ & & & \lambda_4 & \\ & & & & \lambda_5 \end{pmatrix} = A \cdot \begin{pmatrix} | & | & | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 & \vec{v}_5 \\ | & | & | & | & | \end{pmatrix}$$

As in that argument, we deduce that

$$\det(A - k \cdot \mathrm{Id}) = \det \begin{pmatrix} \lambda_1 - k & & & & \\ & \lambda_2 - k & & & \\ & & \lambda_3 - k & & \\ & & & \lambda_4 - k & \\ & & & & \lambda_5 - k \end{pmatrix}.$$

The right hand side is clearly $(\lambda_1 - k)(\lambda_2 - k) \cdots (\lambda_n - k)$.

We now make the reverse argument. Suppose that $f(k) = (\lambda_1 - k)(\lambda_2 - k) \cdots (\lambda_n - k)$ and that the λ_i are distinct. From Big Fact 1, for each λ_i, there is an eigenvector $\vec{v}_i$. By Big Fact 3, they are linearly independent. Since there are n of them, and they are linearly independent, they are a basis.

So what if f has repeated roots?

If $f(k)$ does factor as $(\lambda_1 - k)(\lambda_2 - k) \cdots (\lambda_n - k)$, but with some repeated λ_i, then you can't predict whether or not A has a basis of eigenvectors until you actually do the computation and see. Consider

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Then A and B both have characteristic polynomial $(2-k)^2$. The matrix A only has a one dimensional space of eigenvectors, so we can't find a basis of eigenvectors for A. For B, every vector is an eigenvector, so any basis of $\mathbb{R}^2$ is a basis of eigenvectors.