

THE SPECTRAL THEOREM

Let A be an $n \times n$ symmetric real matrix. An amazing fact is that we can write A as

$$\left(\begin{array}{cccc|c} | & | & | & \cdots & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & | & & | \end{array} \right) \left(\begin{array}{cccc|c} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \lambda_3 & & \\ & & & \ddots & \\ & & & & \lambda_n \end{array} \right) \left(\begin{array}{c|c|c} - & \vec{v}_1 & - \\ - & \vec{v}_2 & - \\ - & \vec{v}_3 & - \\ - & \cdots & - \\ - & \vec{v}_n & - \end{array} \right)$$

with $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ **orthonormal vectors**.

Here is how we compute¹ the λ 's and the \vec{v} 's. The fact that this method works involves what seems to be some amazing pieces of good fortune.

Step 1 Compute the characteristic polynomial $\det(A - k\text{Id})$.

Factor this polynomial as $(\lambda_1 - k)(\lambda_2 - k) \cdots (\lambda_n - k)$. The Fundamental Theorem of Algebra² promises us that such a factorization is possible if we use complex numbers. However, it turns out in our case that life is much better than this:

Lucky Fact 1: All the roots of f are real.

Step 2 For each eigenvalue λ , compute an **orthonormal** basis for $\text{Ker}(A - \lambda\text{Id})$.

Putting all these bases together gives us a list of vectors: $\vec{v}_1, \vec{v}_2, \dots$

Lucky Fact 2: The geometric multiplicity of λ , meaning the dimension of this kernel, is equal to the number of times λ occurs as a root of f .

Thus, the total number of vectors in our list is equal to the number of roots of f , which is n .

Lucky Fact 3: These vectors form an orthonormal basis of \mathbb{R}^n .

In the rest of this note, we will explain why we got so lucky.

A KEY FACT

We will use the following key fact twice below:

Key Fact: Suppose that \vec{u} and \vec{v} are eigenvectors of A , with

$$A\vec{u} = \lambda\vec{u} \quad A\vec{v} = \mu\vec{v}.$$

Suppose that $\lambda \neq \mu$. Then \vec{u} and \vec{v} are orthogonal.

To see this, we compute $\vec{u}^T A\vec{v}$ in two ways. We have

$$\vec{u}^T A\vec{v} = \vec{u}^T (\lambda\vec{v}) = \lambda\vec{u}^T \vec{v}.$$

¹This is how to do the computation by hand. MATLAB uses a rather more sophisticated approach.

²The fundamental theorem of algebra is too hard for this course. The easiest exposition I know is at <http://www.math.binghamton.edu/loya/papers/LoyaFTA.pdf>.

But, also, $\vec{u}^T A \vec{v} = (\vec{v}^T A^T \vec{u})^T = (\vec{v}^T A \vec{u})^T$, where we have replaced A^T by A because A is symmetric. And $(\vec{v}^T A \vec{u})^T = (\vec{v}^T \mu \vec{u})^T = \mu (\vec{v}^T \vec{u})^T = \mu (\vec{u}^T \vec{v})$. Putting it all together,

$$\lambda \vec{u}^T \vec{v} = \mu \vec{u}^T \vec{v}.$$

Since $\lambda \neq \mu$, we get $\vec{u}^T \vec{v} = 0$. In other words, the dot product $\vec{u} \cdot \vec{v}$ is 0 or, in still other words, \vec{u} and \vec{v} are orthogonal.

We now start explaining the Lucky Facts.

LUCKY FACT 1

Suppose that f had a complex root, $\lambda = a + bi$. Let the corresponding eigenvector be $\vec{x} + i\vec{y}$, with x and y each real vectors. So

$$A(\vec{x} + i\vec{y}) = (a + bi)(\vec{x} + i\vec{y}).$$

Taking complex conjugates of everything, we also have

$$A(\vec{x} - i\vec{y}) = (a - bi)(\vec{x} - i\vec{y}).$$

So $\vec{x} - i\vec{y}$ is also an eigenvector, with eigenvalue $a - bi$.

Now, if λ is complex, then $b \neq 0$, and $a + bi \neq a - bi$. We can now use the Key Fact, with $\vec{u} = \vec{x} + i\vec{y}$ and $\vec{v} = \vec{x} - i\vec{y}$. (You probably expected us to use the Key Fact for real vectors, but there is nothing about our argument that needed the entries to be real.)

So we deduce that $\vec{u} \cdot \vec{v} = 0$. Plugging in the formulas for \vec{u} and \vec{v} , that $(\vec{x} + i\vec{y}) \cdot (\vec{x} - i\vec{y}) = 0$.

$$\text{But } (\vec{x} + i\vec{y}) \cdot (\vec{x} - i\vec{y}) = \vec{x} \cdot \vec{x} - i\vec{x} \cdot \vec{y} + i\vec{y} \cdot \vec{x} + \vec{y} \cdot \vec{y} = \vec{x} \cdot \vec{x} + \vec{y} \cdot \vec{y} = |\vec{x}|^2 + |\vec{y}|^2.$$

There is no way this is zero. (Remember that \vec{x} and \vec{y} have real entries.) More precisely, the only way it could be zero is if $\vec{x} = \vec{y} = 0$, but then $\vec{x} + i\vec{y}$ is 0 and is not an example of a nontrivial eigenvector.

We started out supposing that f has a complex root, and reached the absurd conclusion that $|\vec{x}|^2 + |\vec{y}|^2 = 0$. The resolution is that, in fact, f doesn't have any complex roots, and all its roots are real. (Here we are using the Fundamental Theorem of Algebra to know that there are n roots total – real and complex.) Lucky Fact 1 explained!

LUCKY FACT 2

Let λ be a root of $f(k)$, so $\det(A - \lambda \text{Id}) = 0$. Let r be the geometric multiplicity of λ , meaning that there is an r dimensional space of solutions to the equation $A\vec{v} = \lambda\vec{v}$.

Let $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r$ be an orthonormal basis for the space of solutions to $A\vec{v} = \lambda\vec{v}$. Find additional vectors $\vec{v}_{r+1}, \dots, \vec{v}_n$ so that $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r, \vec{v}_{r+1}, \dots, \vec{v}_n$ is an orthonormal basis for \mathbb{R}^n . 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 equations into a matrix. For concreteness, we take $r = 2$ and $n = 5$.

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

You may recognize this argument from the November 13 notes.

$$\text{Let } S = \begin{pmatrix} | & | & | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 & \vec{v}_5 \\ | & | & | & | & | \end{pmatrix}. \quad \text{So } SAS^{-1} = \begin{pmatrix} \lambda & & & & \\ & \lambda & & & \\ c_{31} & c_{32} & c_{33} & c_{34} & c_{35} \\ c_{41} & c_{42} & c_{43} & c_{44} & c_{45} \\ c_{51} & c_{52} & c_{53} & c_{54} & c_{55} \end{pmatrix}.$$

But $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ is an orthonormal basis, so S is orthogonal and $S^{-1} = S^T$. So we have

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

Notice that $(S^T AS)^T = S^T A^T S = S^T AS$, so the above matrix is symmetric. In particular, that lower left block is entirely zero and we have

$$S^{-1}AS = \left(\begin{array}{c|c} \lambda & \\ \hline & \lambda \\ \hline & \mathbf{C} \end{array} \right)$$

for some symmetric $(n-r) \times (n-r)$ matrix C . As in the November 13 notes, we see that

$$\begin{aligned} \det(A - k \cdot \text{Id}_n) &= \det \left(\left(\begin{array}{c|c} \lambda & \\ \hline & \lambda \\ \hline & \mathbf{C} \end{array} \right) - k \cdot \text{Id}_n \right) = \det \left(\begin{array}{c|c} \lambda - k & \\ \hline & \lambda - k \\ \hline & \mathbf{C} - k \cdot \text{Id}_{n-r} \end{array} \right) \\ &= (\lambda - k)^r \det(\mathbf{C} - k \cdot \text{Id}_{n-r}). \end{aligned}$$

Our goal is to show that $\lambda - k$ divides $\det(A - k \cdot \text{Id}_n)$ exactly r times. So we want to know that $\lambda - k$ does **not** divide $\det(\mathbf{C} - k \text{Id}_{n-r})$. In other words, we want to know that λ is **not** an eigenvalue of C .

Suppose, to the contrary that $C\vec{u} = \lambda\vec{u}$. Then

$$\left(\begin{array}{c|c} \lambda & \\ \hline & \lambda \\ \hline & \mathbf{C} \end{array} \right) \begin{pmatrix} 0 \\ 0 \\ \vec{u} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \lambda\vec{u} \end{pmatrix}.$$

So $\begin{pmatrix} 0 \\ 0 \\ \vec{u} \end{pmatrix}$ is another λ eigenvector for A , which should have been listed when we found the v_i .

LUCKY FACT 3

We need to explain two things: Why the \vec{v}_i are orthonormal, and why they are a basis.

If \vec{v}_i and \vec{v}_j both come from the same eigenvalue λ , then $\vec{v}_i \cdot \vec{v}_j = 0$ because we chose an orthonormal basis for the λ -eigenspace. Also, the \vec{v}_i all have length 1 because we chose an orthonormal basis in this place.

If \vec{v}_i and \vec{v}_j come from different eigenvalues, then the Key Fact tells us that \vec{v}_i and \vec{v}_j are perpendicular. So we also have perpendicularity between these vectors.

Since the \vec{v}_i are orthonormal, they are linearly independent. From Lucky Fact 2, the number of \vec{v}_i coming from each λ is the same as the number of times $(\lambda - k)$ divides $f(k)$, and the polynomial f has degree n . So the total number of vectors \vec{v}_i is n . We have n linearly independent vectors in \mathbb{R}^n , so we have a basis for \mathbb{R}^n .

We have now explained all of our Lucky Facts, and shown that every symmetric matrix has a basis of eigenvectors.