Session 19
April 20, 2020

We continue our study of eigenvalues and eigenvectors.

↳ consider an $n\times n$ matrix A

↳ eigenvalues λ and eigenvectors $\vec{x}$ satisfy $A\vec{x}=\lambda\vec{x}$

↳ so the matrix A acts on the eigenvectors in a very special, very simple way

↳ we know how to compute them: find solutions λ to $\det(A-\lambda I)=0$ and then compute $\vec{x}$'s by solving the system of linear equations $A\vec{x}=\lambda\vec{x}$ (given λ)

Today, let us ask the question of when and how many eigenvalues/eigenvectors exist more systematically, and establish more properties.

Recall that eigenvalues are exactly the solution to the characteristic equation $\det(A-\lambda I)=0$, where $\det(A-\lambda I)=(-\lambda)^n+\dots$ is a polynomial of degree n in λ.

So how many solutions are there?

Example: $n=2$

- $\lambda^2=1$ has two solutions: -1 and $+1$
- $(\lambda-2)^2=0$ has one solution: 2
- $\lambda^2=-1$ has no real solution, but two complex solutions: $+i,-i$ (recall: $i^2=-1$, where i is the imaginary unit)

Here are a few facts about solutions to polynomial equations $\sum_{i=0}^{n} c_i\lambda^i=0$ of degree n ($c_n\neq 0$):

a) There are always at most n distinct solutions $\lambda_1,\dots,\lambda_n$, which can possibly be complex.

b) Thus, every polynomial equation of degree n can be written as

$$\boxed{(\lambda_1-\lambda)^{m_1}(\lambda_2-\lambda)^{m_2}\cdot\ldots\cdot(\lambda_k-\lambda)^{m_k}=0}$$ where $k\leq n$, each $m_i\in\mathbb{N}$ and $\underbrace{\sum_{i=1}^{k} m_k=n}_{\text{such that the degree is indeed } n.}$

(Note: One could also write $(\lambda_1-\lambda)(\lambda_2-\lambda)\cdots(\lambda_n-\lambda)=0$, without the m_i, if one allows some λ_i to be the same. But here it is nicer to include the information about how often a factor $(\lambda_i-\lambda)$ appears.)

The m_i are called algebraic multiplicity.

c) If λ is a solution that is not real, then also the complex conjugate $\bar{\lambda}$ is a solution. As a consequence, for n odd, there must be at least one real solution.
(Since complex solutions always come in pairs: λ and $\bar{\lambda}$.)

So if we also allow for complex eigenvalues, we have:

- There are at most n distinct possibly complex eigenvalues.
- Their multiplicities always add up to n
- Complex eigenvalues always come in a pair with its complex conjugate and for n odd there is always at least one real eigenvalue.

In the following, things become much easier if we assume that we can always write

$$\det(A-\lambda I) = (\lambda_1-\lambda)^{m_1}(\lambda_2-\lambda)^{m_2}\cdot\ldots\cdot(\lambda_k-\lambda)^{m_k} \text{ for some } m_1,\ldots,m_k\in\mathbb{N} \text{ with } \sum_{i=1}^k m_i = n,$$

i.e., if we also allow for complex eigenvalues $\lambda_1,\ldots,\lambda_k$. (Later we will look at a class of matrices that only has real eigenvalues.)

Let us collect some more properties of eigenvalues:

One could also write down expressions for the other terms, but these are lengthy and not very useful.

- We have $\det(A-\lambda I) = (-\lambda)^n + \underbrace{(A_{11}+A_{22}+\ldots+A_{nn})}_{\text{think about why this is so using the Laplace expansion or Leibniz formula}}(-\lambda)^{n-1} + \ldots + \underbrace{\det(A)}_{\text{this follows from setting } \lambda=0 \text{ on the left and right-hand side}}$

See the end of the notes for a more detailed explanation of this formula.

But we also have $\det(A-\lambda I) = (\lambda_1-\lambda)^{m_1}\cdots(\lambda_k-\lambda)^{m_k}$ (with $\sum_{i=1}^k m_i = n$)

$$= (-\lambda)^n + (m_1\lambda_1 + m_2\lambda_2 + \ldots + m_k\lambda_k)(-\lambda)^{n-1} + \ldots + \lambda_1^{m_1}\lambda_2^{m_2}\cdots\lambda_k^{m_k}$$

Recall that $\operatorname{tr} A := \sum_{i=1}^{n} A_{ii} = A_{11}+A_{22}+\dots+A_{nn}$ is called trace of A

So we have found: • $\boxed{\operatorname{tr} A = \sum_{i=1}^{k} m_i \lambda_i}$ = sum of all eigenvalues including their multiplicities

• $\boxed{\det A = \prod_{i=1}^{k} \lambda_i^{m_i}}$ = product of all eigenvalues including their multiplicities

• Zero is an eigenvalue if and only if $0 = \det(A - 0\cdot I) = \det A$, i.e., A is singular.

• Let us consider the class of real symmetric matrices, i.e., $A = A^T$, or, in components: $A_{ij} = A_{ji}\ \forall i,j$. Then all eigenvalues are real.

e.g., $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$ is a symmetric 3x3 matrix

(same matrix if we interchange rows and columns)

(Proof for those who are interested: If λ is an eigenvalue, $\vec{x}$ an eigenvector, then $A\vec{x} = \lambda\vec{x}$, where λ and $\vec{x}$ might be complex. But then

$$\overline{\lambda}\underbrace{|\vec{x}|^2}_{\text{real and } >0} = \overline{\lambda}\,\overline{\vec{x}}\cdot\vec{x} = \overline{(\lambda\vec{x})}\cdot\vec{x} = \overline{(A\vec{x})}\cdot\vec{x} = \sum_{i,j}\overline{(A_{ij}x_j)}\,x_i = \sum_{i,j}\overbrace{A_{ij}}^{=A_{ji}\text{ by assumption}}\overline{x_j}\,x_i = \sum_{i,j}\overline{x_j}\,A_{ji}x_i$$

$$= \overline{\vec{x}}\cdot A\vec{x} = \overline{\vec{x}}\cdot\lambda\vec{x} = \lambda|\vec{x}|^2,\ \text{so } \overline{\lambda} = \lambda,\ \text{i.e., } \lambda \text{ was real.})$$

More generally for complex self-adjoint matrices, i.e., $A = \overline{A}^T$, all eigenvalues are real (really important in quantum mechanics!)).

• Suppose λ is an eigenvalue of A with eigenvector $\vec{x}$.

Then $A^2\vec{x} = A(A\vec{x}) = A\lambda\vec{x} = \lambda A\vec{x} = \lambda^2\vec{x}$, so λ^2 is an eigenvalue of A^2 (with the same eigenvector).

In the same way we find that λ^k is an eigenvalue of A^k.

What about A^{-1}, if it exists?

$A\vec{x} = \lambda\vec{x}$, so $\underbrace{A^{-1}(A\vec{x})}_{=\vec{x}} = A^{-1}\lambda\vec{x}$, i.e., $\vec{x} = \lambda A^{-1}\vec{x}$, i.e., $A^{-1}\vec{x} = \frac{1}{\lambda}\vec{x}$.

So $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} (with the same eigenvector).
Note that we assumed A^{-1} exists, i.e., A is non-singular, i.e., 0 is <u>not</u> an eigenvalue (so $\frac{1}{\lambda}$ always makes sense).

- Let me also mention, without proof, the Cayley-Hamilton theorem:
  Any $n \times n$ matrix satisfies its own characteristic equation, meaning if $P(\lambda)$ is the characteristic polynomial, then $\underbrace{P(A)}_{\text{a polynomial of matrices}} = \overbrace{0}^{\text{= the matrix with only zeroes}}$.

This can, e.g., be used to express A^n in terms of A^{n-1}, then A^{n-1} in terms of A^{n-2} and so on. We could also use this to find a formula for A^{-1} in terms of A, but let us skip this here.

Next, let us think about "how many" eigenvectors there are. This brings us to the next chapter.

Extra notes explaining the formula

$$\det(A-\lambda I) = (-\lambda)^n + (A_{11} + A_{22} + \dots + A_{nn})(-\lambda)^{n-1} + \dots \quad .$$

Let us compute $\det(A-\lambda I)$ by repeated Laplace expansion, but only keep track of the terms where powers λ^n or λ^{n-1} appear (we are not interested in the other ones).

$$\det(A-\lambda I) = \det\begin{pmatrix} A_{11}-\lambda & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22}-\lambda & & \vdots \\ \vdots & & \ddots & \\ A_{n1} & \cdots & & A_{nn}-\lambda \end{pmatrix}$$

Laplace expansion along first row

$$= (A_{11}-\lambda)\det\begin{pmatrix} A_{22}-\lambda & A_{23} & \cdots & A_{2n} \\ A_{32} & A_{33}-\lambda & & \vdots \\ \vdots & & \ddots & \\ A_{n2} & \cdots & & A_{nn}-\lambda \end{pmatrix}$$

$$\left.\begin{aligned} &+ (-1)^{1+2} A_{12} \det\begin{pmatrix} A_{21} & A_{23} & \cdots & A_{2n} \\ A_{31} & A_{33}-\lambda & & \vdots \\ \vdots & & \ddots & \\ A_{n1} & \cdots & & A_{nn}-\lambda \end{pmatrix} \\ &+ \cdots \\ &+ (-1)^{n+1} A_{1n} \det(\cdots) \end{aligned}\right\}$$

all of these terms have at most λ^{n-2} in them, but there are no λ^n or λ^{n-1} terms

$$= (A_{11}-\lambda)\det\begin{pmatrix} A_{22}-\lambda & A_{23} & \cdots & A_{2n} \\ A_{32} & A_{33}-\lambda & & \vdots \\ \vdots & & \ddots & \\ A_{n2} & \cdots & & A_{nn}-\lambda \end{pmatrix} + \ldots\lambda^{n-2} + \ldots + \lambda + \text{const}$$

now do a Laplace expansion in the first row again, and again neglect all terms with powers λ^{n-2} or less

$$= (A_{11}-\lambda)(A_{22}-\lambda)\det\begin{pmatrix} A_{33}-\lambda & \cdots & \vdots \\ \vdots & \ddots & \\ \cdots & & A_{nn}-\lambda \end{pmatrix} + \ldots\lambda^{n-2} + \ldots + \ldots\lambda + \text{const}$$

repeat this until no determinants are left
$$= \underbrace{(A_{11}-\lambda)(A_{22}-\lambda)\cdot\ldots\cdot(A_{nn}-\lambda)} + \ldots\lambda^{n-2} + \ldots + \ldots\lambda + \text{const}$$

now we multiply this out and again neglect terms with powers λ^{n-2} or less; we get:

$$(A_{11}-\lambda)\cdots(A_{nn}-\lambda) = (-\lambda)^n + (A_{11}+A_{22}+\ldots+A_{nn})(-\lambda)^{n-1} + \ldots(-\lambda)^{n-2} + \ldots + \ldots\lambda + \text{const}$$

So in total, we get

$$\det(A-\lambda I) = (-\lambda)^n + \left(\sum_{i=1}^{n} A_{ii}\right)(-\lambda)^{n-1} + \ldots\lambda^{n-2} + \ldots + \ldots\lambda + \text{const},$$

which is what we wanted to show.

4.4 Eigenspaces

In this chapter we basically just introduce some terminology, and discuss one interesting theorem.

In our motivational example, we saw that the matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ had two eigenvalues $\lambda_- = -1$ and $\lambda_+ = 1$, and that the corresponding eigenvectors are $\vec{x}_- = t\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ for any $t \neq 0$ and $\vec{x}_+ = s\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ for any $s \neq 0$. So the eigenvectors to an eigenvalue span a whole subspace.

Let us make the corresponding general definition:

Definition: For an $n\times n$ matrix A and eigenvalue λ, we define the eigenspace

$$E_\lambda(A) := \{\vec{x}\in\mathbb{R}^n : A\vec{x} = \lambda\vec{x}\} = \{\text{all eigenvectors corresponding to } \lambda\} \cup \underbrace{\{\vec{0}\}}_{\text{we want the 0 vector to be included here, even though it is never an eigenvector}}.$$

The dimension of $E_\lambda(A)$ is called geometric multiplicity of λ.

In other words, $E_\lambda(A) =$ nullspace of $A-\lambda I$ (by def. of nullspace)

Note that $E_\lambda(A)$ is indeed a subspace, i.e., taking scalar multiples and sums does not lead out of $E_\lambda(A)$. Why?

↳ If $\vec{x}\in E_\lambda(A)$, $t\in\mathbb{R}$, then $A(t\vec{x}) = t\,A\vec{x} = t\lambda\vec{x} = \lambda(t\vec{x})$, i.e., $t\vec{x}\in E_\lambda(A)$

↳ If $\vec{x}$ and $\vec{y}\in E_\lambda(A)$, then $A(\vec{x}+\vec{y}) = A\vec{x}+A\vec{y} = \lambda\vec{x}+\lambda\vec{y} = \lambda(\vec{x}+\vec{y})$, i.e., also $\vec{x}+\vec{y}\in E_\lambda(A)$

E.g., in the example of the matrix $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 6 \\ 0 & 1 & 2 \end{pmatrix}$ from last time, we found that there are three eigenvalues $\lambda_0 = 0$, $\lambda_+ = 4$, $\lambda_- = -1$ and that the corresponding eigenspaces are

$$E_{\lambda_0}(A) = \left\{ t\begin{pmatrix}1\\0\\0\end{pmatrix}, \forall t\in\mathbb{R}\right\},\quad E_{\lambda_+} = \left\{ t\begin{pmatrix}\frac{1}{2}\\2\\1\end{pmatrix}, \forall t\in\mathbb{R}\right\},\quad E_{\lambda_-} = \left\{ t\begin{pmatrix}3\\-3\\1\end{pmatrix}, \forall t\in\mathbb{R}\right\}.$$

But eigenspaces might be two-dimensional, or have any higher dimension up to n.

Consider two examples:

a) $A = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ (just the identity)

Clearly there is only one eigenvalue $\lambda=+1$ with algebraic multiplicity $m=2$, since

$$\det(A-\lambda I)=\det\begin{pmatrix}1-\lambda & 0\\ 0 & 1-\lambda\end{pmatrix}=(1-\lambda)^2=(\lambda_1-\lambda)^{m_1} \text{ with } \lambda_1=1,\ m_1=2=\text{alg. mult.}$$

And clearly any vector $\vec{x}\in\mathbb{R}^2$ is an eigenvector ($A\vec{x}=I\vec{x}=1\cdot\vec{x}$, and $\vec{x}=\vec{x}$ is true for any $\vec{x}\in\mathbb{R}^2$), so $E_{\lambda=+1}(A=I)=\mathbb{R}^2$ is 2-dimensional, i.e., the geometric multiplicity is 2.

b) $A=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$.

Here, we have $\det(A-\lambda I)=\det\begin{pmatrix}1-\lambda & 1\\ 0 & 1-\lambda\end{pmatrix}=(1-\lambda)^2$, so again there is only one eigenvalue $\lambda=+1$ with algebraic multiplicity $m_1=2$.

Then $(A-\lambda I)\vec{x}=\left[\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}-\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right]\vec{x}=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\vec{0}$ implies $y=0$ (the second line does not tell us anything). So $E_{\lambda=1}(A)=\{t\begin{pmatrix}1\\0\end{pmatrix},\forall t\in\mathbb{R}\}=\mathrm{span}\{\begin{pmatrix}1\\0\end{pmatrix}\}$, which is one-dimensional, so the geometric multiplicity is 1.

x is arbitrary

$=\{\begin{pmatrix}x\\0\end{pmatrix} \text{ s.t. } x\in\mathbb{R}\}$

Conclusion: Sometimes algebraic multiplicity = geometric multiplicity, but sometimes not.

However it looks like when all eigenvalues are different, then the algebraic multiplicity = the geometric multiplicity = 1. In fact, the following is true:

Theorem: Let A be an $n\times n$ matrix with n distinct eigenvalues. Then the corresponding eigenvectors are linearly independent.

so the algebraic multiplicities are all 1

As a consequence, all eigenspaces of A are one-dimensional.

so the geometric multiplicities are all 1

So this means that if we choose n eigenvectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ corresponding to the n distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, then $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ is a basis.

This theorem is something we should prove:

Proof: Let us call the n distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, and let us choose n corresponding eigenvectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$, i.e., $A\vec{v}_i = \lambda_i \vec{v}_i$.

We now want to prove that $\vec{v}_1, \dots, \vec{v}_n$ are linearly independent, i.e., whenever a linear combination $\sum_{i=1}^{n} c_i \vec{v}_i = \vec{0}$, then this implies $c_i = 0 \ \forall\, i = 1, \dots, n$. (In other words none of the $\vec{v}_i$'s can be written as a linear combination of the others.)

So let's see: We take some linear combination $\sum_{i=1}^{n} c_i \vec{v}_i$ and apply A to it:

$$A\left(\sum_{i=1}^{n} c_i \vec{v}_i\right) = \sum_{i=1}^{n} c_i \underbrace{A\vec{v}_i}_{=\lambda_i \vec{v}_i} = \sum_{i=1}^{n} c_i \lambda_i \vec{v}_i .$$

So we need to satisfy both $\sum_{i=1}^{n} c_i \vec{v}_i = \vec{0}$ and $\sum_{i=1}^{n} c_i \lambda_i \vec{v}_i = \vec{0}$.

So, e.g., for $n=2$, we have $c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$ and $c_1\lambda_1\vec{v}_1 + c_2\lambda_2\vec{v}_2 = \vec{0}$. If we multiply the first equation by $-\lambda_1$ and add it to the second, this gives

$$\vec{0} = -\lambda_1 c_1 \vec{v}_1 - \lambda_1 c_2 \vec{v}_2 + c_1\lambda_1\vec{v}_1 + c_2\lambda_2\vec{v}_2 = c_2 \underbrace{(\lambda_2 - \lambda_1)}_{\neq 0,\ \text{since } \lambda_1 \neq \lambda_2} \vec{v}_2 .$$

But the right-hand side can only vanish if $c_2 = 0$ ($\vec{v}_2$ was an eigenvector, i.e., $\vec{v}_2 \neq \vec{0}$). But then also $c_1 = 0$, which proves linear independence for $n=2$. Higher dimensions work in the same way by just repeating the argument above.

$\square$