Theory sheet 5

Definition of eigenvalues and eigenvectors

Definition 1. *Let $A \in M_{n,n}$ be a square matrix and $\lambda \in \mathbb{R}$ a number. We say that a non-zero vector $\boldsymbol{x} \in \mathbb{R}^n$ is an eigenvector of A corresponding to the eigenvalue λ if*

$$A\boldsymbol{x} = \lambda \boldsymbol{x}.$$

In other words, if A sends $\boldsymbol{x}$ to itself, but stretched by λ.

A few remarks:

- If $\boldsymbol{x}$ is an eigenvector of A and λ is its eigenvalue, then

  $$A^2\boldsymbol{x} = \lambda^2\boldsymbol{x}, \quad A^3\boldsymbol{x} = \lambda^3\boldsymbol{x}, \quad \ldots, \quad A^n\boldsymbol{x} = \lambda^n\boldsymbol{x}$$

  for all n.

- Compare notion of eigenvector and eigenvalue with scaling (homothetic) linear transformation from previous lecture. The transformation

  $$S_\lambda(\boldsymbol{x}) = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

  streches every vector by λ, whereas A with eigenvalue λ stretches by λ only some vectors – its eigenvectors.

- Note that λ is clearly an eigenvalue of $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$, and every vector is its eigenvector.

- If a matrix is proportional to the identity, that is, of the form $A = \lambda I$, we say that A is a scalar matrix. Scalar matrices are the only matrices which commute with every other: $AB = BA$ for all B implies that $A = \lambda I$ for some number λ. We will show later on that λ is the only eigenvalue of $A = \lambda I$.

- Note that if λ is an eigenvalue of A and we want to find $\boldsymbol{x}$ such that $A\boldsymbol{x} = \lambda\boldsymbol{x}$, then there are an infinite number of solutions. Why? Because if $\boldsymbol{x}$ is one such solution, then so is $2\boldsymbol{x}$, as well as $3\boldsymbol{x}$, $4\boldsymbol{x}$, $\pi\boldsymbol{x}$, et cetera, because $A(2\boldsymbol{x}) = 2A\boldsymbol{x} = 2\lambda\boldsymbol{x} = \lambda(2\boldsymbol{x})$.

- Eigenvalues and eigenvectors only make sense for square matrices. Indeed, if $A \in M_{m,n}$, then $A\boldsymbol{x} \in \mathbb{R}^m$, but $A\boldsymbol{x} = \lambda\boldsymbol{x} \in \mathbb{R}^n$. Hence, $m = n$.

Finding eigenvalues and eigenvectors

Remark 1. *Recall that if* $A\boldsymbol{x} = \mathbf{0}$ *and* A *is invertible, then* $A^{-1}A\boldsymbol{x} = \boldsymbol{x}$ *on one hand and* $A^{-1}A\boldsymbol{x} = A^{-1}\mathbf{0} = \mathbf{0}$ *on the other, so* $\boldsymbol{x}$ *must be zero.*

Theorem 1. λ *is an eigenvalue of* A *if and only if* $\det(A - \lambda I) = 0$ *.*

Proof. ($\Longrightarrow$) If λ is an eigenvalue, then $A\boldsymbol{x} = \lambda\boldsymbol{x}$ for some non-zero vector $\boldsymbol{x}$. Hence,

$$A\boldsymbol{x} - \lambda\boldsymbol{x} = \mathbf{0}, \quad \text{hence} \quad (A - \lambda I)\,\boldsymbol{x} = \mathbf{0}.$$

Since $A - \lambda I$ sends a non-zero vector to zero, it cannot be invertible. Hence, $\det(A - \lambda I) = 0$, as claimed.

($\Longleftarrow$) If $\det(A - \lambda I) = 0$, then $A - \lambda I$ is not invertible. Hence, there exists a non-zero vector $\boldsymbol{x}$ which it sends to zero:

$$(A - \lambda I)\,\boldsymbol{x} = \mathbf{0}.$$

Hence, $A\boldsymbol{x} = \lambda\boldsymbol{x}$, as claimed. □

The following theorem is given without proof:

Theorem 2. *If* $A \in M_{n,n}$ *, the function* $p(\lambda) = \det(A - \lambda I)$ *is a polynomial of degree* n *. It is called the characteristic polynomial of matrix* A *.*

An important corollary of this theorem is that eigenvalues always exist:

Corollary 1. *Any square matrix* $A \in M_{n,n}$ *has at least one eigenvalue.*

Proof. Eigenvalues of A are the roots of $p(\lambda)$. Since p is a polynomial, it has a root (fundamental theorem of algebra). □

Recipe for finding eigenvalues. To find eigenvalues of A ,

- denote an unknown eigenvalue by λ
- subtract λI from A
- calculate the characteristic polynomial $p(\lambda) = \det(A - \lambda I)$
- set it equal to zero: $\det(A - \lambda I) = 0$
- find solutions of this equation.

Example

$$A = \begin{pmatrix} -1 & 4 \\ -3 & -8 \end{pmatrix} \implies \det(A - \lambda I) = \det \begin{pmatrix} -1-\lambda & 4 \\ -3 & -8-\lambda \end{pmatrix}$$
$$= (-1-\lambda)(-8-\lambda) - 4 \cdot (-3)$$
$$= 8 + 8\lambda + \lambda + \lambda^2 + 12$$
$$= \lambda^2 + 9\lambda + 20.$$

Therefore, we need to solve
$$\lambda^2 + 9\lambda + 20 = 0.$$
It is clear that there are two solutions:
$$\lambda_1 = -4 \quad \text{and} \quad \lambda_2 = -5.$$

Finding eigenvectors

To find eigenvectors, we need to first find the eigenvalues. If eigenvalues are known, we can find corresponding eigenvectors by solving the linear system
$$(A - \lambda I)\,\boldsymbol{x} = \mathbf{0}.$$
To solve this system, use approach developed in the previous lecture.

Example

We have found above that $\lambda_1 = -4$ and $\lambda_2 = -5$ are eigenvalues of $\begin{pmatrix} -1 & 4 \\ -3 & -8 \end{pmatrix}$. Let us find the eigenvectors corresponding to λ_1:
$$A - \lambda_1 I = A + 4I = \begin{pmatrix} 3 & 4 \\ -3 & -4 \end{pmatrix}.$$
Hence, we need to find $\boldsymbol{x}$ such that
$$\begin{pmatrix} 3 & 4 \\ -3 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$
Adding first line to the second, we obtain
$$\begin{pmatrix} 3 & 4 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$
We can drop the second line because there is zero against it, hence any $\boldsymbol{x}$ satisfying
$$3x_1 + 4x_2 = 0$$

is an eigenvector. We can parametrize these eigenvectors by x_2:

$$\left\{ \begin{pmatrix} -4x_2/3 \\ x_2 \end{pmatrix}, \quad x_2 \in \mathbb{R} \right\}.$$

We also need to find the eigenvectors corresponding to $\lambda = -5$:

$$(A - \lambda I)\,\boldsymbol{x} = \mathbf{0} \iff \begin{pmatrix} 4 & 4 \\ -3 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and the set of solutions is given by

$$\left\{ \begin{pmatrix} -x_2 \\ x_2 \end{pmatrix}, \quad x_2 \in \mathbb{R} \right\}.$$

Properties of eigenvalues and eigenvectors

Definition 2. *The set of eigenvectors $\boldsymbol{x}$ corresponding to an eigenvalue λ of A is called the eigenspace.*

Proof of this theorem is left as an exercise:

Theorem 3. *The eigenspace of A corresponding to λ is a vector space. In other words, if $\boldsymbol{x}, \boldsymbol{y}$ satisfy $A\boldsymbol{x} = \lambda\boldsymbol{x}$ and $A\boldsymbol{y} = \lambda\boldsymbol{y}$, then (a) $\boldsymbol{x}+\boldsymbol{y}$ also satisfies this: $A(\boldsymbol{x}+\boldsymbol{y}) = \lambda(\boldsymbol{x}+\boldsymbol{y})$ and (b) for any number μ, $A(\mu\boldsymbol{x}) = \lambda(\mu\boldsymbol{x})$.*

The following theorem is a simple yet deep result:

Theorem 4. *Determinant of A is equal to the product of all eigenvalues of A:*

$$\det A = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n = \prod_{i=1}^{n} \lambda_i.$$

Proof. Note that $\det A = \det(A - \lambda I)\,|_{\lambda=0} = p(0)$, where p is the characteristic polynomial of A. By Vieta's theorem, $p(0)$ is equal to the product of all roots of p for all polynomials (not just characteristic polynomials of matrices). □

Corollary 2. *A is invertible if and only if it does not have zero eigenvalues.*

Proof. If $\lambda_i \neq 0$ for all i, then $\det A = \prod_{i=1}^{n} \lambda_i \neq 0$, so the matrix is invertible. On the other hand, if A is invertible, we have that $\det A \neq 0$ and by this product representation there cannot be $\lambda_i = 0$ among eigenvalues of A. □

Remark 2. *Hence, if $A \in M_{2,2}$ and we somehow know one eigenvalue, we can find the other using previous theorem. For example, if we are given that $\lambda_1 = 1$ is an eigenvalue of $\begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$, we can find the second eigenvalue λ_2 by*

$$\lambda_2 = \lambda_1 \cdot \lambda_2 = \det \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} = (1-a)(1-b) - ab = 1 - a - b.$$

Theorem 5. *If A is symmetric ($A^\top = A$), then eigenvectors corresponding to different eigenvalues are orthogonal. In other words, $\boldsymbol{x}^\top \boldsymbol{y} = 0$ for every $\boldsymbol{x}$ such that $A\boldsymbol{x} = \lambda \boldsymbol{x}$, $\boldsymbol{y}$ such that $A\boldsymbol{y} = \mu \boldsymbol{y}$ and $\lambda \neq \mu$.*

Proof. Since $\lambda \neq \mu$, we have $\lambda - \mu \neq 0$, therefore

$$\begin{aligned}
(\lambda - \mu)\, \boldsymbol{x}^\top \boldsymbol{y} &= \lambda\, \boldsymbol{x}^\top \boldsymbol{y} - \mu\, \boldsymbol{x}^\top \boldsymbol{y} \\
&= \lambda\, \boldsymbol{y}^\top \boldsymbol{x} - \mu\, \boldsymbol{x}^\top \boldsymbol{y} \\
&= \boldsymbol{y}^\top (\lambda \boldsymbol{x}) - \boldsymbol{x}^\top (\mu \boldsymbol{y}) \\
&= \boldsymbol{y}^\top A \boldsymbol{x} - \boldsymbol{x}^\top A \boldsymbol{y} \\
&= \boldsymbol{x}^\top A^\top \boldsymbol{y} - \boldsymbol{x}^\top A \boldsymbol{y} \\
&= 0,
\end{aligned}$$

where the last line follows from $A^\top = A$. Hence, $\boldsymbol{x}^\top \boldsymbol{y} = 0$. □

Remark 3. *Eigenvectors may be orthogonal without $A = A^\top$.*

Remark 4. *Two different matrices may have the same characteristic polynomial. For example,*

$$A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 5 \\ 0 & 4 \end{pmatrix}.$$

Hence, they have the same eigenvalues, but their eigenvectors are different.