Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute for Operations Research
Department of Mathematics

Robert Weismantel

Week 12: Eigenvalues and Eigenvectors

Eigenvalues and Eigenvectors

Informal definition:

Given a square matrix A, an eigenvalue λ and eigenvector v are a scalar and a non-zero vector satisfying

$$Av = \lambda v \iff (A - \lambda I)v = 0, \text{ i.e.,}$$

$$(A - \lambda I) \text{ is not invertible} \iff \det(A - \lambda I) = 0.$$

Strategy

We can try to find eigenvalues by inspecting the solutions of $\det(A - \lambda I) = 0$ which is a polynomial in λ but unfortunately, not all polynomials have real zeros.

Drawback

We need to take a little detour to the complex numbers to develop the tools for an analysis of this beautiful topic.

A bit more formal

Definition

Given $A \in \mathbb{R}^{n \times n}$, we say $\lambda \in \mathbb{C}$ is an eigenvalue of A and $v \in \mathbb{C}^n \setminus \{0\}$ is an eigenvector of A, associated with the eigenvalue λ if

$$Av = \lambda v.$$

We call them an eigenvalue-eigenvector pair. If $\lambda \in \mathbb{R}$ then we will call λ a real eigenvalue, and the associated eigenvalue-eigenvector pair a real pair.

Example 1

Consider $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Suppose that λ is an eigenvalue of A. Then,

$$\det(A - \lambda I) = \det(\begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix}) = \lambda^2 + 1 = 0.$$

This polynomial equation has only solutions in $\mathbb{C}$, the Complex Numbers.

Complex numbers I

What are complex numbers?

Complex numbers are of the form $z = a + ib$ for $a \in \mathbb{R}$ and $b \in \mathbb{R}$.
Our notation is $\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}$.

Keeping in mind that $i^2 = -1$ we can do operations

- $(a+ib)+(x+iy) = (a+x)+i(b+y)$,
- $(a+ib)(x+iy) = ax + i(ay+bx) + i^2 by = ax + i(ay+bx) - by = (ax-by)+i(ay+bx)$,
- $(a+ib)(a-ib) = a^2 + b^2$,
- $$\begin{aligned} \frac{a+ib}{x+iy} &= \frac{(x-iy)(a+ib)}{(x-iy)(x+iy)} = \frac{(ax+by)+i(bx-ay)}{x^2+y^2} \\ &= \left(\frac{ax+by}{x^2+y^2}\right) + i\left(\frac{bx-ay}{x^2+y^2}\right). \end{aligned}$$

$\mathbb{C}^n$ and $\mathbb{C}^{m\times n}$ denote complex valued vectors and matrices.

Complex numbers II

Given $z \in \mathbb{C}$ with $z = a + ib$ we have the following notation

$$
\begin{aligned}
\Re(a+ib) &:= a && \text{called the real part of } z = a+ib, && (1)\\
\Im(a+ib) &:= b && \text{called the imaginary part of } z = a+ib, && (2)\\
|z| &:= \sqrt{a^2+b^2} && \text{called the modulus of } z = a+ib, && (3)\\
\bar{z} &:= a - ib && \text{called the complex conjugate of } z = a+ib. && (4)
\end{aligned}
$$

Elementary calculations show

- For $z = a + ib \in \mathbb{C}$,

$$|z|^2 = a^2 + b^2 = a^2 - i^2 b^2 = (a+ib)(a-ib) = z\bar{z}.$$

- For $z = a + ib \in \mathbb{C}$, $\frac{1}{z} = \frac{\bar{z}}{|z|^2}$.
- For $z_1, z_2 \in \mathbb{C}$, $|z_1||z_2| = |z_1 z_2|$.

Complex vectors and matrices continued

Transposing complex vectors v and matrices A

$$v^* = \overline{v}^T \text{ and } A^* = \overline{A}^T. \tag{5}$$

For $v \in \mathbb{C}^n$ we have

$$\|v\|^2 = v^* v = \overline{v}^T v = \sum_{i=1}^{n} \overline{v_i} v_i = \sum_{i=1}^{n} |v_i|^2.$$

The inner-product (or dot-product) in $\mathbb{C}^n$ is given by $\langle v, w \rangle = w^* v$.

Canonical notation

- $v_1, \ldots, v_k \in \mathbb{C}^n$ are linearly independent if $\alpha_1 v_1 + \cdots + \alpha_k v_k = 0$ for $\alpha_1, \ldots, \alpha_k \in \mathbb{C}$ forces $\alpha_1 = \cdots = \alpha_k = 0$.
- $\mathrm{Span}(v_1, \ldots, v_k) = \{x \in \mathbb{C}^n \mid x = \sum_{i=1}^{k} \alpha_i v_i \text{ for } \alpha_i \in \mathbb{C}\}$.
- If $v_1, \ldots, v_k$ is a spanning set of a subspace and linearly independent we say it is a basis of that subspace.

Why complex numbers, vectors and matrices?

Theorem (Fundamental Theorem of Algebra)

Any degree n non-constant ($n \geq 1$) polynomial

$$P(z) = \alpha_n z^n + \alpha_{n-1} z^{n-1} + \cdots + \alpha_1 z + \alpha_0$$

(with $\alpha_n \neq 0$) has a zero: $\lambda \in \mathbb{C}$ such that $P(\lambda) = 0$.

Once we have λ a zero of $P(z)$,

divide $P(z)$ by $(z - \lambda)$ to get $P(z) = (z - \lambda)P_1(z)$ and iterate with $P_1(z)$.

Corollary

Any degree n non-constant ($n \geq 1$) polynomial $P(z)$ (with $\alpha_n \neq 0$) has n zeros: $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, perhaps with repetitions, such that

$$P(z) = \alpha_n (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_n). \tag{6}$$

Algebraic multiplicity of $\lambda \in \mathbb{C}$ = number of times λ appears in this expansion.

Example 1 again

Suppose that $\lambda \in \mathbb{C}$ is an eigenvalue of A

- $\det(A - \lambda I) = \lambda^2 + 1 = 0$. We obtain two solutions $\lambda_1 = i$, $\lambda_2 = -i$.
- Let us now try to find an eigenvector $v \in \mathbb{C}^2$ for $\lambda_1 = i$.

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} a + ib \\ x + iy \end{pmatrix} \text{ and } Av = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} v = \begin{pmatrix} -v_2 \\ v_1 \end{pmatrix}.$$

- This leads to the following system of equations

$$Av = iv \iff -x - iy = -b + ia \text{ and } a + ib = -y + xi.$$

- This leads to the solution $x = b$ and $y = -a$. Hence,

$$v = \begin{pmatrix} a + ib \\ b - ia \end{pmatrix} = (a + ib) \begin{pmatrix} 1 \\ -i \end{pmatrix}.$$

- We conclude that an eigenvector v for λ_1 is the vector $\begin{pmatrix} 1 \\ -i \end{pmatrix}$.

The theory in general I

Observation

Let $A \in \mathbb{R}^{n\times n}$. $\lambda \in \mathbb{R}$ is a (real) eigenvalue of $A \iff \det(A - \lambda I) = 0$.
A vector $v \in \mathbb{R}^n$ is an eigenvector associated with $\lambda \iff v \in N(A - \lambda I)$.

Lemma

$\det(A - \lambda I)$ *is a polynomial in* λ *of degree* n *with coefficient* $(-1)^n$ *for* λ^n.

From the Fundamental Theorem of Algebra:

Every matrix $A \in \mathbb{R}^{n\times n}$ has an eigenvalue (perhaps complex-valued).

Lemma

Let $A \in \mathbb{R}^{n\times n}$. If $\lambda \in \mathbb{C}$ and $v \in \mathbb{C}^n$ are a complex eigenvalue-eigenvector pair, then also $\bar{\lambda} \in \mathbb{C}$ and $\bar{v} \in \mathbb{C}^n$ are a complex eigenvalue-eigenvector pair.

The theory in general II

Lemma

If λ and v are an eigenvalue-eigenvector pair of a matrix A, then, for $k \geq 1$, λ^k and v are an eigenvalue-eigenvector pair of the matrix A^k.

Proof by induction. $k = 1$ is trivial.

For the induction, if λ^{k-1} and v are an eigenvalue-eigenvector pair for A^{k-1},

$$A^k v = A\left(A^{k-1} v\right) = A\left(\lambda^{k-1} v\right) = \lambda^{k-1} A v = \lambda^k v.$$

Proposition

Let A be an invertible matrix. If λ and v are an eigenvalue-eigenvector pair of A, then, $\frac{1}{\lambda}$ and v are an eigenvalue-eigenvector pair of the matrix A^{-1}.

Proof.

A is invertible and hence, in the statement $\lambda \neq 0$. Since $Av = \lambda v$ we have

$$A^{-1}(\lambda v) = v \Rightarrow \lambda A^{-1} v = v \iff A^{-1} v = \frac{1}{\lambda} v.$$

The theory in general III

Proposition

Let $A \in \mathbb{R}^{n \times n}$ and let $v_1, \ldots, v_k \in \mathbb{R}^n$ be eigenvectors corresponding to eigenvalues $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$. If $\lambda_1, \ldots, \lambda_k$ are all distinct, the eigenvectors $v_1, \ldots, v_k$ are linearly independent.

Proof.

- We prove by induction that

$$j = \dim(\mathrm{Span}(\{v_1, \ldots, v_j\})) = \dim(\{x \in \mathbb{R}^n \mid x = \sum_{l=1}^{j} \mu_l v_l \text{ for } \mu \in \mathbb{R}^l\}).$$

- For $j = 1$ the statement is correct since $v_1 \neq 0$ by definition.
- Suppose the statement is correct for index $j - 1$.
- For the purpose of deriving a contradiction, assume that

$$v_j = \alpha_1 v_1 + \cdots \alpha_{j-1} v_{j-1}, \ \alpha \in \mathbb{R}^{j-1}. \quad (7)$$

Proof continued

Assume $v_j = \alpha_1 v_1 + \cdots \alpha_{j-1} v_{j-1}$.

- If we multiply by A both sides we get

$$\lambda_j v_j = A v_j = A\left(\alpha_1 v_1 + \cdots \alpha_{j-1} v_{j-1}\right) = \alpha_1 \lambda_1 v_1 + \cdots \lambda_{j-1} \alpha_{j-1} v_{j-1}.$$

- Replacing v_j with $\alpha_1 v_1 + \cdots \alpha_{j-1} v_{j-1}$ we get

$$\lambda_j\left(\alpha_1 v_1 + \cdots \alpha_{j-1} v_{j-1}\right) = \alpha_1 \lambda_1 v_1 + \cdots \lambda_{j-1} \alpha_{j-1} v_{j-1} \iff$$

$$\alpha_1\left(\lambda_j - \lambda_1\right) v_1 + \alpha_2\left(\lambda_j - \lambda_2\right) v_2 + \cdots + \alpha_{j-1}\left(\lambda_j - \lambda_{j-1}\right) v_{j-1} = 0. \quad (8)$$

- Since $\lambda_j - \lambda_i \neq 0$ for all $i \leq j-1$ and not all α_i's are zero, this is a non-zero linear combination of $v_1, \ldots, v_{j-1}$ adding to zero contradicting the hypothesis.

Theorem

Let $A \in \mathbb{R}^{n \times n}$ with n distinct real eigenvalues. There is a basis of $\mathbb{R}^n$, $v_1, \ldots, v_n$, made up of eigenvectors of A.

The characteristic polynomial

Definition

The polynomial (9) in variable $z \in \mathbb{C}$ is the Characteristic Polynomial of the matrix A:

$$(-1)^n \det(A - zI) = \det(zI - A) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_n). \tag{9}$$

The right hand side is its factorization. $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A.

Remark

Note that in Eq. (9) not all λ_i are distinct. The number of times an eigenvalue λ appears in this expression is called the algebraic multiplicity of λ.

Lemma

For $A \in \mathbb{R}^{n \times n}$ the eigenvalues of A and $A^\top$ are the same.

follows from (9) and $\det(A^\top - zI^\top) = \det((A - zI)^\top) = \det(A - zI)$.

The trace of a matrix

Definition

Given a matrix $A \in \mathbb{R}^{n\times n}$, the trace of A is defined as

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} A_{ii}.$$

Lemma

For matrices $A, B, C \in \mathbb{R}^{n\times n}$ the following relations hold.

(i) $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$

(ii) $\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA) = \mathrm{Tr}(CAB)$.

Proof.

$$\mathrm{Tr}(AB) = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}B_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{n} B_{ji}A_{ij} = \mathrm{Tr}(BA).$$

$$\mathrm{Tr}(ABC) = \mathrm{Tr}(A(BC)) = \mathrm{Tr}((BC)A) = \mathrm{Tr}(B(CA)) = \mathrm{Tr}(CAB).$$

The trace of a matrix II

Theorem

Let $A \in \mathbb{R}^{n \times n}$ and $\lambda_1, \ldots, \lambda_n$ its n eigenvalues as they show up in (9)*.*

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} \lambda_i \quad \textit{and} \quad \det(A) = \prod_{i=1}^{n} \lambda_i.$$

Proof.

$$\begin{aligned}(-1)^n \det(A - zI) &= (z-\lambda_1)(z-\lambda_2)\cdots(z-\lambda_n) \\ &= z^n + (-\textstyle\sum_{i=1}^n \lambda_i) z^{n-1} + \sum_{k=1}^{n-2} b_k z^k + (-1)^n \prod_{i=1}^n \lambda_i,\end{aligned}$$

where $b_k \in \mathbb{C}$.

- Set $z = 0$. It gives $(-1)^n \det(A) = (-1)^n \prod_{i=1}^n \lambda_i$.
- For the second relation the coefficient of z^{n-1} in the characteristic polynomial is given in the right hand side by $(-\sum_{i=1}^n \lambda_i)$.
- On the left hand side the coefficient of z^{n-1} can only come from the permutation that takes all diagonal elements in the matrix $zI - A$.
- Hence it is the coefficient of z^{n-1} of $\prod_{i=1}^n (z - A_{ii})$ which is $-\sum_{i=1}^n A_{ii} = -\mathrm{Tr}(A)$.