Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute for Operations Research
Department of Mathematics

Robert Weismantel

Week 13: Repeated eigenvalues, symmetric matrices and the spectral theorem

Eigenvalues, eigenvectors: summary

Summary: Let $A \in \mathbb{R}^{n \times n}$

- If (λ, v) is an eigenvalue-eigenvector-pair of A, then $(\bar{\lambda}, \bar{v})$ is an eigenvalue-eigenvector-pair of A.
- The eigenvalues of A and A^T are the same, not so the corresponding eigenvectors.
- If A is invertible and (λ, v) is an eigenvalue-eigenvector-pair of A, then $(1/\lambda, v)$ is an eigenvalue-eigenvector-pair of A^{-1}.
- The eigenvalues of $A+B$ are not the sum of eigenvalues of A and B.
- The eigenvalues of AB are not the product of eigenvalues of A and B.
- Gaussian Elimination doesn't preserve eigenvalues and eigenvectors.
- Let $Q \in \mathbb{R}^{n \times n}$ be an orthogonal matrix., i.e., $Q^T Q = I$. If $\lambda \in \mathbb{C}$ is an eigenvalue of Q, then $|\lambda| = 1$.

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.

Repeated Eigenvalues

Repeated eigenvalues can pose an obstacle to building a basis.

We have shown that if $A \in \mathbb{R}^{n\times n}$ has n distinct real eigenvalues, then there is a basis of $\mathbb{R}^n$ made up of eigenvectors of A. But what if not?

Example

- $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ does not have two linearly independent eigenvectors.
  $\det(A - \lambda I) = \lambda^2$ which means that $\lambda = 0$ is the only eigenvalue and has algebraic multiplicity 2. However, $N(A - 0I) = N(A)$ only has dimension 1, so there is only one linearly independent eigenvector.
- $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ has two linearly independent eigenvectors.
  $\det(A - \lambda I) = \lambda^2$ which means that $\lambda = 0$ is the only eigenvalue and has algebraic multiplicity 2. $N(A - 0I) = N(A)$ has dim. 2.

How to proceed?

Definition

Let $A \in \mathbb{R}^{n \times n}$. If one can build a basis of $\mathbb{R}^n$ with eigenvectors of A we say that A has a complete set of real eigenvectors.

When do we have a complete set of real eigenvectors?

- A matrix with n distinct real eigenvalues always has a complete set of real eigenvectors.
- For $D \in \mathbb{R}^{n \times n}$ a diagonal matrix, the eigenvalues of D are the diagonal entries of D. The canonical basis $e_1, \ldots, e_n$ is a set of eigenvectors of D.
- When there is an eigenvalue λ with algebraic multiplicity larger than 1, it can be that $N(A - \lambda I)$ is of large enough dimension to find enough linearly independent eigenvectors.

Definition

Given a matrix $A \in \mathbb{R}^{n \times n}$ and an eigenvalue λ of A we call the dimension of $N(A - \lambda I)$ the geometric multiplicity of λ.

When do we have a complete set of real eigenvectors

Observation

If the geometric multiplicities equal the algebraic multiplicites of all eigenvalues, then such a matrix has a complete set of eigenvectors. (Note the eigenvectors corresponding to distinct eigenvalues are l.i.)

Proposition

Let P be the projection matrix on the subspace $U \subseteq \mathbb{R}^n$. Then P has two eigenvalues, 0 and 1, and a complete set of real eigenvectors.

Proof.

Let m be the dimension of U. Let $u_1, \ldots, u_m$ be an orthonormal basis of U, and $w_1, \ldots, w_{n-m}$ an orthonormal basis of $U^\perp$.

$$Pu_k = 1u_k \text{ for } 1 \leq k \leq m \text{ and } Pw_k = 0w_k \text{ for } 1 \leq k \leq n - m.$$

Hence, all n vectors are eigenvectors of P (with eigenvalues 1 or 0).

Can we use this theory for diagonalizing a matrix?

The idea:

- Let $A \in \mathbb{R}^{n \times n}$ and assume that A has a complete set of real eigenvectors. For $i \in \{1, \ldots, n\}$ let λ_i be the eigenvalue associated with eigenvector v_i.
- This fact allows us to write $x \in \mathbb{R}^n$ as

$$x = \sum_{i=1}^{n} \alpha_i v_i \ \Rightarrow Ax = \sum_{i=1}^{n} \lambda_i \alpha_i v_i, \text{ i.e.,}$$

- The linear transformation corresponding to writing an x in the basis V allows us to transform A to a diagonal matrix.

Theorem

Let $A \in \mathbb{R}^{n \times n}$ with a complete set of eigenvectors $v_1, \ldots, v_n \in \mathbb{R}^n$ associated with eigenvalues $\lambda_1, \ldots, \lambda_n$. Let V be the matrix with columns v_i. Then, $A = V \Lambda V^{-1}$, where Λ is a diagonal matrix with $\Lambda_{ii} = \lambda_i$ ($\Lambda_{ij} = 0$ for $i \neq j$).

The proof

$v_1, \ldots, v_n$ is a basis, hence V is invertible and it remains to prove

$$V^{-1}AV = \Lambda.$$

For $1 \leq j \leq n$, the j-th column of the matrix $V^{-1}AV$ is given by

$$\left(V^{-1}AV\right)_{\cdot j} := \left(V^{-1}AV\right)e_j = V^{-1}Av_j = V^{-1}\lambda_j v_j = \lambda_j V^{-1}v_j = \lambda_j e_j,$$

since $V^{-1}v_j = V^{-1}Ve_j = e_j$. Note that $\lambda_j e_j$ is the j-th column of Λ. Hence, we have that

$$V^{-1}AV = \Lambda.$$

Definition (Diagonalizable Matrix)

A matrix $A \in \mathbb{R}^{n\times n}$ is called diagonalizable if there exists an invertible matrix V and a diagonal matrix Λ such that

$$V^{-1}AV = \Lambda.$$

Why do we want diagonizable matrices?

It allows us to perform a change of basis

using eigenvectors so that the matrix A becomes diagonalizable.

The idea more general

Let $u_1, \ldots, u_n$ be a basis for $\mathbb{R}^n$ and $v_1, \ldots, v_m$ a basis of $\mathbb{R}^m$. Consider the transformation L that maps $x = \sum_{j=1}^n \alpha_j u_j$ to $L(x) = Ax = \sum_{j=1}^n \beta_i v_i$.
We want to compute the matrix B that takes

$$\alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \text{ to } \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}, \text{i.e.}, B\alpha = \beta.$$

Let $U \in \mathbb{R}^{n \times n}$ and $V \in \mathbb{R}^{m \times m}$ have columns $u_1, \ldots, u_n$ and $v_1, \ldots, v_m$.
Then, $x = U\alpha$ and $L(x) = V\beta$ and so $\beta = V^{-1}AU\alpha$ and hence,

$$B = V^{-1}AU.$$

In summary,

$L : \mathbb{R}^n \to \mathbb{R}^m, \qquad L(x) = Ax.$

$$L\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{i=1}^{n} (Ax)_i e_i \qquad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{1}$$

$$L\left(\sum_{j=1}^{n} \alpha_j u_j\right) = \sum_{i=1}^{n} (B\alpha)_i v_i \qquad \alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix},$$

$$B = V^{-1}AU \in \mathbb{R}^{m \times n}$$

$U = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \in \mathbb{R}^{n \times n}, \quad V = \begin{bmatrix} v_1 & \cdots & v_m \end{bmatrix} \in \mathbb{R}^{m \times m}$

Specifically if A is square and diagonalizable

$U = V$ can be chosen and B becomes a diagonal matrix!

Similiar matrices

Definition (Similar Matrices)

We say that $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times n}$ are similar matrices if there exists an invertible matrix S such that $B = S^{-1}AS$.

Proposition

Similar matrices have the same eigenvalues.

Proof.

A and B are similiar, i.e., $B = S^{-1}AS$. Let λ be an eigenvalue of A with associated eigenvector v. Then $Av = \lambda v$. Define $w = S^{-1}v$. We obtain

$$Bw = S^{-1}ASw = S^{-1}ASS^{-1}v = S^{-1}Av = \lambda S^{-1}v = \lambda w.$$

Conversely, let λ be an eigenvalue of B with associated eigenvector w. Then $Bw = \lambda w$. Define $v = Sw$. We obtain

$$Av = SBS^{-1}v = SBS^{-1}Sw = SBw = \lambda Sw = \lambda v.$$

Our next target: Symmetric matrices

Target

Here we consider symmetric matrices $A \in \mathbb{R}^{n \times n}$, $A^T = A$. Our target is to show that such a matrix always has a complete set of eigenvectors.

Proposition

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix and λ an eigenvalue of A, then $\lambda \in \mathbb{R}$.

Proof

Let $v \in \mathbb{C}^n$ be an eigenvector associated with the eigenvalue $\lambda \in \mathbb{C}$. We have $Av = \lambda v$. Recall that, for a matrix (or vector) M, its Hermitian conjugate is given by $M^* = \overline{M}^\top$. Since A is real symmetric we have $A^* = A$. Thus

$$\overline{\lambda} \|v\|^2 = \overline{\lambda} v^* v = (\lambda v)^* v = (Av)^* v = v^* A^* v = v^* A v = v^* \lambda v = \lambda \|v\|^2.$$

Since $v \neq 0$, then $\|v\| \neq 0$ and so $\lambda = \overline{\lambda}$. This implies that $\lambda \in \mathbb{R}$.

A second preparatory step

Proposition

Let $A \in \mathbb{R}^{n\times n}$ be symmetric and $\lambda_1 \neq \lambda_2$ two distinct eigenvalues of A with corresponding eigenvectors v_1, v_2. Then v_1 and v_2 are orthogonal.

Proof.

$v_1, v_2 \neq 0$ and hence,

$$\lambda_1 v_1^\top v_2 = (Av_1)^\top v_2 = v_1^\top A^\top v_2 = v_1^\top A v_2 = v_1^\top (Av_2) = \lambda_2 v_1^\top v_2,$$

since $\lambda_1 \neq \lambda_2$ we must have that $v_1^\top v_2 = 0$.

Theorem (Spectral Theorem)

Every symmetric matrix $A \in \mathbb{R}^{n\times n}$ has n real eigenvalues and an orthonormal basis made of eigenvectors of A.

First a few consequences of the spectral theorem

Corollary

For any symmetric matrix $A \in \mathbb{R}^{n\times n}$ there exists an orthogonal matrix $V \in \mathbb{R}^{n\times n}$ (whose columns are eigenvectors of A) such that

$$A = V\Lambda V^\top,$$

where $\Lambda \in \mathbb{R}^{n\times n}$ is a diagonal matrix with the eigenvalues of A in its diagonal (and $V^\top V = I$).

Let A be a real $n \times n$ symmetric matrix

Let $v_1, \ldots, v_n$ be an orthonormal basis of eigenvectors of A and $\lambda_1, \ldots, \lambda_n$ the associated eigenvalues. Then $A = \sum_{k=1}^n \lambda_i v_i v_i^\top$

Corollary

The rank of a real symmetric matrix A is the number of non-zero eigenvalues (counting repetitions).

Proof of the spectral theorem I

The point of departure

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix. We will prove the following by induction, which for $k = n$ implies the theorem we want to show:

- For any $k \in \{1, \ldots, k\}$ there are k orthogonal eigenvectors of A corresponding to k real eigenvalues of A.
- If $k = 1$, this statement is true.

The inductive step

Assume the statement is true for k, i.e., A has k (with $1 \leq k < n$) orthonormal eigenvectors. Then we can build an extra one, orthogonal to the others.

- $v_1, \ldots, v_k$ denote k orthonormal eigenvectors of A and $\lambda_1, \ldots, \lambda_k$ the respective eigenvalues.
- Let $u_{k+1}, \ldots, u_n$ be an orthonormal basis of the orthogonal complement of the span of $v_1, \ldots, v_k$.
- Let V_k be the $n \times n$ matrix whose i-th column is v_i if $i \leq k$ and u_i if $i > k$. V_k is an orthogonal matrix.

Proof of the spectral theorem II: Define

$$
B = V^\top A V = \begin{bmatrix} — & v_1^\top & — \\ & \vdots & \\ — & v_k^\top & — \\ — & u_{k+1}^\top & — \\ & \vdots & \\ — & u_n^\top & — \end{bmatrix} \begin{bmatrix} | & & | & | & & | \\ Av_1 & \cdots & Av_k & Au_{k+1} & \cdots & Au_n \\ | & & | & | & & | \end{bmatrix}
$$

$$
= \begin{bmatrix} — & v_1^\top & — \\ & \vdots & \\ — & v_k^\top & — \\ — & u_{k+1}^\top & — \\ & \vdots & \\ — & u_n^\top & — \end{bmatrix} \begin{bmatrix} | & & | & | & & | \\ \lambda v_1 & \cdots & \lambda v_k & Au_{k+1} & \cdots & Au_n \\ | & & | & | & & | \end{bmatrix}
$$

$$
= \begin{bmatrix} \Lambda_k & 0_{k\times(n-k)} \\ 0_{(n-k)\times k} & C \end{bmatrix},
$$

Λ_k is diagonal with entries $\lambda_1, \dots, \lambda_k$, C is a $(n-k)\times(n-k)$ symmetric matrix.

Proof of the spectral theorem III

- Since C is a $(n-k) \times (n-k)$ symmetric matrix, it has a real eigenvalue λ_{k+1} and a real eigenvector $y \in \mathbb{R}^{n-k}$.
- Let $w \in \mathbb{R}^n$,

$$w_i = \begin{cases} 0 & \text{if} \quad i \leq k \\ y_{i-k} & \text{if} \quad i > k. \end{cases}$$

-

$$Bw = \begin{bmatrix} \Lambda_k & 0_{k\times(n-k)} \\ 0_{(n-k)\times k} & C \end{bmatrix} \begin{bmatrix} 0_{k\times 1} \\ y \end{bmatrix} = \begin{bmatrix} 0_{k\times 1} \\ Cy \end{bmatrix} = \begin{bmatrix} 0_{k\times 1} \\ \lambda_{k+1} y \end{bmatrix} = \lambda_{k+1} w.$$

- Let $v_{k+1} := Vw$. V is orthogonal and $A = VBV^\top$. Thus,

$$Av_{k+1} = VBV^\top v_{k+1} = VBw = V\lambda_{k+1} w = \lambda_{k+1} v_{k+1},$$

  so v_{k+1} is an eigenvector of A.
- Show that v_{k+1} is orthogonal to $v_1, \ldots, v_k$!
- The inner products $v_i^\top v_{k+1}$ for $i \leq k$ appear in the first k entries of $V^\top v_{k+1} = w$ and w has its first k coordinates equal to 0.
- By normalizing the vector we can have it attain unit norm.