5374 Fall '22

Numerical Linear Algebra — Lecture 20

Today: * Symmetric matrices and eigenvectors

Theorem: If A is a symmetric $n\times n$ matrix then there exists n vectors

$$v_1, \ldots, v_n$$

which are orthonormal and n numbers $\lambda_1, \ldots, \lambda_n$ such that for each $k = 1, \ldots, n$ we have

$$A v_k = \lambda_k v_k$$

Remarks: In particular, since the $\{v_k\}$ form an orthonormal matrix, given $x \in \mathbb{R}^n$, we have

$$x = (x, v_1) v_1 + \cdots + (x, v_n) v_n$$

$$A x = \lambda_1 (x, v_1) v_1 + \cdots + \lambda_n (x, v_n) v_n$$

This last equation is the same as saying

$$A x = V D V^t x \quad \forall\ x \in \mathbb{R}^n$$

where

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}, \quad V = \left(v_1 \,|\, v_2 \,|\, \cdots \,|\, v_n \right)$$

In other words, A can be factorized as (note $V^{-1} = V^t$)

$$A = VDV^{-1} \qquad (*)$$

In other words, the theorem says any symmetric matrix A is conjugate to a diagonal matrix and the conjugation is done by an orthogonal matrix.

The factorization $(*)$ is sometimes called the eigenvector decomposition of A (even if A is not symmetric and V is not orthogonal).

Exercise: ① Show that if $v_1, \dots, v_m$ are eigenvectors of A ($n \times n$, $n \geq m$) with respective eigenvalues $\lambda_1, \dots, \lambda_m$ all different, then the $v_1, \dots, v_m$ are linearly independent

② Show that if the characteristic polynomial of A ($n \times n$) has n different real roots then A has an eigenvector decomposition

Definition: If A has an eigenvector decomposition, it is said that A is diagonalizable

Remark: Here is another way of thinking about this theorem, and one that naturally extends to infinite dimensions and is important in functional analysis.

Given a matrix A and eigenvalue λ, define

$$E_\lambda = \{ v \in \mathbb{R}^n \mid Av = \lambda v \}$$

and define

$$P_\lambda = \text{orthogonal projection into } E_\lambda$$

$$\left(\text{ie. } \quad P_\lambda^2 = P_\lambda, \quad P_\lambda^t = P_\lambda, \quad \operatorname{Im}(P_\lambda) = E_\lambda \right)$$

Then the theorem stated earlier can also be stated as follows:

If A is symmetric and A's set of eigenvalues is $\lambda_1, \dots, \lambda_m$ $(m \leq n)$ then, we have

$$A = \lambda_1 P_{\lambda_1} + \lambda_2 P_{\lambda_2} + \cdots + \lambda_m P_{\lambda_m}$$

where $\quad I = P_{\lambda_1} + P_{\lambda_2} + \cdots + P_{\lambda_m}$

and $\quad P_{\lambda_k} P_{\lambda_j} = 0 \quad$ if $k \neq j$.

Note: In functional analysis, certain linear operators L can be decomposed as

$$L = \int_{\mathbb{R}} \lambda P_\lambda \, d\mu(\lambda)$$

where μ is something called the spectral measure

Remark: All of the above statements have straightforward analogues for Hermitian matrices A, i.e. matrices A with complex entries such that

$$A^* = A$$

There, the orthogonality is in terms of the usual Hermitian product in $\mathbb{C}^n$, and "orthogonal matrix" is replaced with "unitary matrix."

Moreover, for such Hermitian matrices, the eigenvalues are all real.

Proof of the eigenvalue decomposition for symmetric matrices

Lemma: Let A be symmetric with two eigenvectors v_1, v_2 with different eigenvalues λ_1, λ_2 THEN $v_1 \perp v_2$. In other words, the spaces E_{λ_1} and E_{λ_2} are orthogonal to each other.

Proof:

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

Since A is symmetric, $(Av_1, v_2) = (v_1, Av_2)$

$$\Rightarrow \quad \lambda_1 (v_1, v_2) = \lambda_2 (v_1, v_2)$$

$$(\lambda_1 - \lambda_2)(v_1, v_2) = 0$$

$$\Rightarrow \quad (v_1, v_2) = 0 \quad (\text{since } \lambda_1 - \lambda_2 \neq 0)$$

∎

This shows how the symmetry of A guarantees the orthonormality of the eventual eigenvector basis.

The problem we may face when doing an eigenvector decomposition is: not enough eigenvectors exist to span $\mathbb{R}^n$.

Example: Let

$$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

The matrix has exactly one eigenvalue, λ, which is a double root of the characteristic polynomial of A. However,

$$\begin{aligned} E_\lambda &= \{ v \in \mathbb{R}^2 \mid Av = \lambda v \} \\ &= \operatorname{span}\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right] \neq \mathbb{R}^2 \end{aligned}$$

Definition: A subspace $V \subseteq \mathbb{R}^n$ is said to be **stable** with respect to a matrix A if (AKA, invariant)
$v \in V \Rightarrow Av \in V$.

Lemma: Let A be symmetric and $V \subseteq \mathbb{R}^n$ a subspace stable under A. $V \neq \{0\}$ Then there exists an eigenvector of A in V.

Proof: It's a variational proof. You consider

$$f(x) := \frac{1}{2}\langle Ax, x\rangle \quad \forall\ x \in \mathbb{R}^n$$

It's gradient is $\nabla f(x) = Ax$ (This uses the symmetry of A, in general, if A is not symmetric

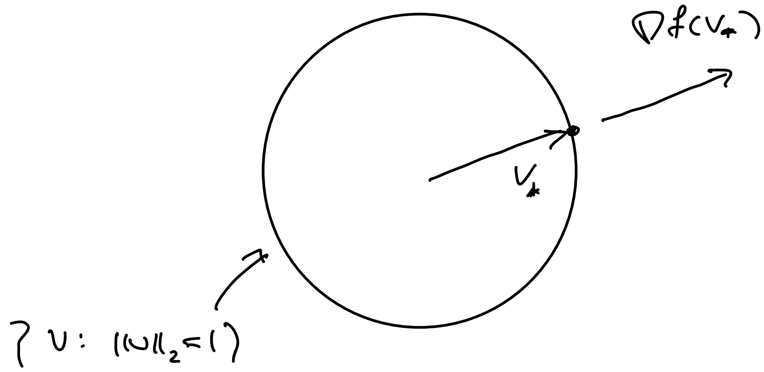
$$\nabla f(x) = \left(\tfrac{1}{2}A + \tfrac{1}{2}A^t\right)x \quad)$$

Consider f restricted to the following set:

$$\{x \in \mathbb{R}^n \mid \|x\|_2 = 1 \text{ and } x \in V\}$$

This set is compact, and so f restricted to it achieves a maximum and a minimum.

Let V_* be an element of the set where the minimum of $f(x)$ is achieved.



So the gradient $\nabla f(V_*)$ must be parallel to V_*, or rather there must be $\alpha \in \mathbb{R}$ s.t.

$$\nabla f(V_*) = \alpha V_*$$

i.e.
$$A V_* = \alpha V_*$$

Since $\|V_*\|_2 = 1$, $V_* \neq 0$ so V_* is an eigvect of A in V.

$\boxtimes$

Now we combine the lemmas and finish the proof of the theorem:

We will construct an orthonormal basis of eigenvectors for A.

Step 1: $\mathbb{R}^n$ is not $\{0\}$ and is stable under A, so by the last lemma there exists a unit vector, which we call v_1, and $\lambda_1 \in \mathbb{R}$ s.t.

$$Av_1 = \lambda_1 v_1$$

Step 2: Suppose you have constructed k orthonormal eigenvectors $v_1, \dots, v_k$ $(k \geq 1)$, if $k = n$, you are done. If not, then $k < n$, and then we consider

$$\mathcal{V} = \{ x \in \mathbb{R}^n \mid (x, v_j) = 0 \text{ for } j = 1, \dots, k \}$$

Since $k < n$, $\mathcal{V}$ is not the null subspace.
I claim $\mathcal{V}$ is stable under A:

let $x \in \mathcal{V}$, then $(x, v_j) = 0$ $\forall j = 1, \dots, k$, let's show $Ax \in \mathcal{V}$. Indeed,

$$(Ax, v_j) = (x, Av_j) = (x, \lambda_j v_j) = \lambda_j (x, v_j) = 0$$

so $(Ax, v_j) = 0$ for $j = 1, \dots, k$, and Ax also lies in V.

So V is stable, and not null, so by the last lemma there is a unit eigenvector of A in V, we call it v_{k+1}.

We repeat step 2 until it hits $k = n$, and we are done.