Lecture 20

SYMMETRIC MATRICES

* PSET 3 is out! $\boxed{\text{Due } 10/30 \text{ 9PM.}}$
(has Julia component, so start early!)

* Mid-course Survey (please fill it in!)

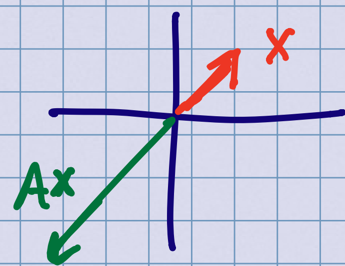
* Checkups (revamped!)

* Midterm Re-dos (signup sheet)

Eigenvalues / Eigenvectors

$A \in \mathbb{R}^{n \times n}$ (square)

$$\underset{\text{matrix}}{A}\, x = \underset{\text{scalar (eigenvalue)}}{\lambda}\, x$$



DIAGONALIZATION: (sometimes)

$$A = \underset{\text{eigenvectors}}{T}\, \overset{\text{eigenvalues}}{D}\, T^{-1}$$

Singular Value Decomposition (SVD)

$A \in \mathbb{R}^{m \times n}$ (rectangular)

(always exists)

$$\underset{m \times n}{A} = \underset{\substack{m \times m \\ \text{orthogonal}}}{U}\, \underset{\substack{m \times n \\ \text{diagonal}}}{\overset{\text{singular values}}{\Sigma}}\, \underset{\substack{n \times n \\ \text{orthogonal}}}{V^T}$$

(low-rank approximation, PCA, ..)

$$A = \sum_{i=1}^{\overset{\text{rank}}{r}} \sigma_i \overbrace{u_i v_i^T}^{\text{rank 1}}$$

(sum of rank 1 matrices)

$$\overbrace{\sigma_1 \geq \cdots \geq \sigma_r}^{\text{nonzero}} > \sigma_{r+1} = \cdots = 0$$

EVEN FOR SQUARE MATRICES,

EIGENVALUES AND SINGULAR VALUES

CAN BE QUITE DIFFERENT,
SINCE THEY REVEAL DIFFERENT
ASPECTS OF A MATRIX.

EXAMPLE:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 10^6 \\ 0 & 0 & 0 \end{bmatrix}$$

Eigenvalues ?

Singular Values ?

Q: How do they relate? (if they do)

Q: How to compute the SVD??

Eigenvalues
Eigenvectors

Singular
Value
Decomposition

SYMMETRIC
MATRICES

OPTIMIZATION

(as we have seen)

If a matrix can be diagonalized,
all kind of good things happen.

$$\overset{\text{eigenvectors}}{\mathbf{A} = \mathbf{T}\underbrace{\mathbf{D}}_{\text{eigenvalues}}\mathbf{T}^{-1}}$$

But sometimes, we want even more!

E.G., some additional conditions
on T (eigenvectors)

A (very important!) special case:

SYMMETRIC MATRICES.

Def: $A \in \mathbb{R}^{n \times n}$ is symmetric if

$$\boxed{A = A^T} \qquad (a_{ij} = a_{ji})$$

e.g.

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix} \text{ is symmetric}$$

(identity matrix) I_n is symmetric

$\underset{n\times m}{W^T} \underset{m \times m}{B} \underset{m \times n}{W}$ is symmetric if B is symmetric

Why study symmetric matrices?

They appear naturally in many areas
(including , in optimization).

Some examples:

- Second derivatives: (Hessians)
$$\frac{\partial f}{\partial x_i \partial x_j} = \frac{\partial f}{\partial x_j \partial x_i}$$

- Quadratic forms: (multivariate polynomials of degree 2)
$$q(x) = x^T Q x = \sum Q_{ij} x_i x_j$$

- Covariance matrices: (correlations betweeen random variables)
$$\mathbb{E}[x x^T]$$

- Physics: classical mechanics (inertia tensor)
quantum mechanics (density matrix)
relativity ... (metric tensor of spacetime)

Quadratic forms / polynomials

E.g. $p(x,y,z) = x^2 + 2xz + 3y^2 - 10zy - z^2$

$x,y,z \in \mathbb{R}$

all monomials have degree 2.
"homogeneous quadratic".

Can represent in matrix form:

$$p(x,y,z) = \begin{bmatrix} x \\ y \\ z \end{bmatrix}^T \underbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & -5 \\ 1 & -5 & -1 \end{bmatrix}} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

More generally, if $x \in \mathbb{R}^n$ and

$$p(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij} x_i x_j + \sum_{i}^{n} \beta_i x_i$$

Can represent as.

$$p(x) = x^T \underbrace{A}_{\text{symmetric matrix}} x + b^T x$$

quadratic forms $\Longleftrightarrow$ symmetric matrices

Q: What does the eigenvalue decomposition of the matrix A tell me?

What's particularly nice about real symmetric matrices?

Let $A \in \mathbb{R}^{n \times n}$, $A = A^T$.

- **All** eigenvalues λ_i **are real**.

$$Ax = \lambda x \overset{\text{(conjugate)}}{\Longrightarrow} A\bar{x} = \bar{\lambda}\bar{x} \overset{\text{(transpose)}}{\Longrightarrow} \bar{x}^T A = \bar{\lambda}\bar{x}^T$$

right-multiply by x

$$\Longrightarrow \quad \bar{x}^T A x = \bar{\lambda}\,\bar{x}^T x$$

$$\lambda(\bar{x}^T x) = \bar{\lambda}(\bar{x}^T x)$$

$$\Longrightarrow \quad \boxed{\lambda = \bar{\lambda}} \iff \lambda \in \mathbb{R}$$

- Eigenvectors can always be chosen to be orthogonal. $(v_i \cdot v_j = 0)$

v_i, v_j eigenvectors of different eigenvalues.

$$A v_i = \lambda_i v_i \xrightarrow{\text{transpose}} v_i^T A^T = \lambda_i v_i^T$$

$$= \quad v_i^T A = \lambda_i v_i^T$$

$$\xrightarrow{v_j} \quad v_i^T A v_j = \lambda_i v_i^T v_j$$

$$\lambda_j (v_i^T v_j) = \lambda_i (v_i^T v_j)$$

$$(\lambda_i - \lambda_j)(v_i^T v_j) = 0$$

$$\Rightarrow \underline{v_i^T v_j = 0.}$$

(also for same λ_i)

$\Rightarrow$ $T = \begin{bmatrix} v_1 \dots & v_n \end{bmatrix}$ is an orthogonal matrix!

if $\|v_i\| = 1$

This is great! Then A is diagonalizable, and furthermore: $\boxed{T^{-1} = T^{T}}$ $(A = TDT^{-1})$

A is symmetric $\Rightarrow$ $\boxed{A = TDT^{T}}$ (always!)

or equivalently, $\boxed{A = \sum_{i=1}^{n} \lambda_i v_i v_i^{T}}$

(with $\lambda_i \in \mathbb{R}$, $v_i \in \mathbb{R}^n$)

Structure of symmetric matrices is very simple / natural:

$TT^{T} = I$ (orthogonal)

$$\boxed{A = \underbrace{T}_{\text{"rotation"}} \underbrace{D}_{\text{diagonal}} \underbrace{T^{T}}_{\text{inverse "rotation"}}}$$

"After a suitable rotation, every symmetric matrix is diagonal"

Hmmm. This reminds me of something, doesn't it?

Indeed! Start with $A = U \Sigma V^T$

What are the eigenvalues of AA^T and A^TA?

(both symmetric).

$$AA^T = U \Sigma V^T V \Sigma U^T = U \Sigma^2 U^T$$

eigenvalues!

$$A^T A = V \Sigma U^T U \Sigma V^T = V \Sigma^2 V^T$$

eigenvalues!

"Reversing", we can use this to compute the SVD.

⇒ The (nonzero) eigenvalues of AA^T and A^TA are the same, and they are equal to σ_i^2.

Alternatively,

$$\boxed{\sigma_i = \lambda_i^{1/2}(AA^T) = \lambda_i^{1/2}(A^TA)}$$

⇒ The singular vectors u_i, v_j are the eigenvectors of AA^T and A^TA, respectively.

(In practice, algorithms are a bit different, to preserve numerical accuracy.)

A bit about computation...

Basic tasks in linear algebra:

- Solve linear systems
- Compute matrix inverses
- Eigenvalues / Eigenvectors
- Singular Value Decomposition

The basic algorithms we learned (e.g. Gaussian elimination) are ok for small/medium problems. Typical complexity is $O(n^3)$ or $O(n^2 m)$. (dense matrices)

A lot of research about systems/algorithms with particular structure (e.g., sparse, PDEs, ML, graphics, randomized, ...)