EE263: Introduction to Linear Dynamical Systems
Review Session 8

• Symmetric matrices

• Matrix inequalities

• SVD
Symmetric matrices

In the following problems you can assume that $A = A^T \in \mathbb{R}^{n \times n}$ and $B = B^T \in \mathbb{R}^{n \times n}$. We do not, however, assume that $A$ or $B$ is positive semidefinite. For $X = X^T \in \mathbb{R}^{n \times n}$, $\lambda_i(X)$ will denote its $i$th eigenvalue, sorted so $\lambda_1(X) \geq \lambda_2(X) \geq \cdots \geq \lambda_n(X)$.
Example: Similarity transformation

Is this true or false?

Suppose there is an orthogonal matrix $R$ such that $A = R^T BR$. Then the eigenvalues of $A$ and $B$ are the same, i.e., $\lambda_i(A) = \lambda_i(B)$ for $i = 1, \ldots, n$.

Solution. True.

- since $A$ is symmetric we can write $A = Q\Lambda Q^T$, where $\Lambda$ is diagonal and $Q$ is orthogonal
- but $A = R^T BR = Q\Lambda Q^T$ implies $B = R(Q\Lambda Q^T)R^T$
- thus, $B = (RQ)\Lambda(RQ)^T$ where $RQ$ is orthogonal
Example: Ellipsoid containment

Is this true or false?

If \( \{ x \mid x^T A x \leq 1 \} \subseteq \{ x \mid x^T B x \leq 1 \} \), then \( A \geq B \).

Solution. False.

• we know the statement is true when \( B > 0 \) from lecture 15-18

• consider the case where \( B \) is a negative definite matrix, the set \( \{ x \mid x^T B x \leq 1 \} \) is equal to \( \mathbb{R}^n \)

• the set \( \{ x \mid x^T A x \leq 1 \} \) is clearly a subset of \( \mathbb{R}^n \), regardless of what \( A \) is

• \( A \) can be such that \( A < B \); e.g., the scalar case \( A = -2, B = -1 \)
Example

Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times n}$ both be symmetric and positive definite. What can you say about the eigenvalues of $AB$?

Solution.

- We can choose $A^{1/2}$, and $A^{-1/2}$ so that $A^{1/2}A^{1/2} = A$, and $A^{1/2}A^{-1/2} = I$.

- The eigenvalues of $AB$ are the same as the eigenvalues of $A^{-1/2}ABA^{1/2} = A^{1/2}BA^{1/2}$.

- The matrix $A^{1/2}BA^{1/2}$ is symmetric and positive definite, which implies that the eigenvalues of $AB$ are real and positive.
Example: Matrix exponential

Is this true or false?

If $A \geq B$ then for all $t \geq 0$, $e^{At} \geq e^{Bt}$.

Solution. False.

• consider

\[
A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad B = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}.
\]

• eigenvalues of $A - B$ are 0 and 2, so $A - B \geq 0$

• eigenvalues of $e^A - e^B$ are 4.2983 and −0.2656, which means $e^A$ and $e^B$ are not comparable

• this is one of those tricky things that is true for scalars, but false for matrices
SVD Fundamentals

• Let $A \in \mathbb{R}^{m \times n}$, $A = U\Sigma V^T = \sum_{i} \sigma_i u_i v_i^T$ is the singular value decomposition of $A$.

• $S = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$ is the unit ball (ellipsoid) in $\mathbb{R}^n$,

• $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is given by the linear mapping $f(x) = Ax$

• $f(x)$ maps the unit ball $S \subseteq \mathbb{R}^n$ to an ellipsoid in $\mathbb{R}^m$
SVD Fundamentals

- Right singular vectors $v_i$ are mapped to left singular vectors $u_i$.
- Semiaxis lengths given by $\sigma_i$.

\[ x \mapsto Ax \]
SVD Properties

• $A = U\Sigma V^T$, where $U, V$ are orthogonal, $\Sigma$ diagonal.

• $r = \text{rank}(A)$ is the number of nonzero singular values of $A$, where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$.

• $\{u_1, \ldots, u_r\}$ is an orthonormal basis for $\mathcal{R}(A)$.

• $\{v_{r+1}, \ldots, v_n\}$ is an orthonormal basis for $\mathcal{N}(A)$.

• The pseudoinverse is given by

$$A^\dagger = \hat{V} \hat{\Sigma}^{-1} \hat{U}^T,$$

from the “thin” svd, where the inverses in $\hat{\Sigma}$ are taken along the diagonal.
Testing for membership in span

• Recall that \( y \in \mathcal{R}(A) \) if \( \text{rank} \begin{bmatrix} y & A \end{bmatrix} = \text{rank}(A) \). This is a numerically unsound way to check if something is in the range.

• The component \( \hat{y} \) of \( y \) in \( \mathcal{R}(A) \) is computed by projecting \( y \) onto \( \text{span}\{u_1, \ldots, u_r\} \), i.e.,

\[
\hat{y} = \sum_{i=1}^{r} u_i u_i^T y
\]

• Thus, \( y \in \mathcal{R}(A) \) if and only if the component \( z \) of \( y \) in \( \mathcal{R}(A)^\perp \) is \( z = y - \hat{y} = 0 \).

• Note that \( z \) can be written as \( z = (I - \hat{U}\hat{U}^T)y \)

• So, \( y \in \mathcal{R}(A) \) if and only if \( (I - \hat{U}\hat{U}^T)y = 0 \)
Computing SVD “by hand”

- $A = U\Sigma V^T$ means $A^T A = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^2 V^T$ is symmetric

- Similarly, $AA^T = U\Sigma^2 U^T$

- Thus, the singular values are the square roots of the eigenvalues of $A^T A$ or $AA^T$:

  \[ \sigma_i = \sqrt{\lambda_i(A^T A)} = \sqrt{\lambda_i(AA^T)} \]

- Right singular vectors $v_i$ are the eigenvectors of $A^T A$, the left singular vectors $u_i$ are the eigenvectors of $AA^T$

- Much better algorithms exist!
Matrix norm

- The (induced) matrix norm of $A$ is $\|A\| = \max_{\|x\|=1} \|Ax\|$

- Also $\|A\| = \max_{x \neq 0} \|Ax\|/\|x\| = \sigma_1(A)$

Obeys “norm” properties, like

- Scaling: $\|cA\| = |c|\|A\|$ for $c \in \mathbb{R}$

- Triangle inequality: $\|A + B\| \leq \|A\| + \|B\|$

- Definiteness: $\|A\| = 0$ if and only if $A = 0$

- Submultiplicative Identity: $\|Ax\| \leq \|A\|\|x\|$ for $x \in \mathbb{R}^n$
Example

We are given \( A \in \mathbb{R}^{m \times n} \), with \( \text{svd} \ A = U \Sigma V^T \). How can we find vectors \( x \) and \( y \) that maximize \( y^T Ax \), subject to \( \|y\| = 1, \|x\| = 1 \)?

**Solution.**

• We know that

\[
y^T Ax \leq \|y\| \|Ax\| \leq \|y\| \|A\| \|x\| = \|A\|.
\]

• This upper bound is achieved by \( y = u_1 \), and \( x = v_1 \), which means that

\[
\max_{\|y\|=1,\|x\|=1} y^T Ax = \|A\| = \sigma_1.
\]
Example

We are given $A \in \mathbb{R}^{m \times n}$, and $B \in \mathbb{R}^{k \times n}$. Assume that $A^T A$ is invertible. Find a nonzero vector $w \in \mathbb{R}^n$ that maximizes

$$d = \frac{w^T B^T Bw}{w^T A^T A w}.$$ 

Solution.

- We know how to solve the problem in the case $A^T A = I$

- Define $z = (A^T A)^{1/2} w$, so we have $w = (A^T A)^{-1/2} z$. Then we can write

$$\max_{w \neq 0} \frac{w^T B^T Bw}{w^T A^T A w} = \max_{z \neq 0} \frac{z^T (A^T A)^{-1/2} B^T B (A^T A)^{-1/2} z}{z^T z}.$$ 

- Thus, $d_{\text{max}} = \lambda_{\text{max}} \left( (A^T A)^{-1/2} B^T B (A^T A)^{-1/2} \right)$
The value of \( z \) that maximizes the ratio is the eigenvector associated with the maximum eigenvalue above. To find the \( w \) that maximizes \( d \), we simply multiply this eigenvector by \((A^T A)^{-1/2}\).
Low rank approximations

- Let $A = U\Sigma V^T$ be the SVD of $A$ with $r = \text{rank}(A)$.

- We want to find a matrix $\hat{A}$, with $\text{rank}(\hat{A}) \leq p < r$, so that $\|A - \hat{A}\|$ is minimized. (where $\| \cdot \|$ can refer to either the matrix norm, or the Frobenius norm — the solution is the same in both cases)

- The optimal rank $p$ approximator of $A$ is

$$\hat{A} = \sum_{i=1}^{p} \sigma_i u_i v_i^T,$$

- The optimal approximation error is $\|A - \hat{A}\| = \left\| \sum_{i=p+1}^{r} \sigma_i u_i v_i^T \right\| = \sigma_{p+1}$