EE263: Introduction to Linear Dynamical Systems
Review Session 8

- Symmetric matrices
- Matrix inequalities
- SVD

Symmetric matrices

In the following problems you can assume that $A = A^T \in \mathbf{R}^{n\times n}$ and $B = B^T \in \mathbf{R}^{n\times n}$. We *do not*, however, assume that A or B is positive semidefinite. For $X = X^T \in \mathbf{R}^{n\times n}$, $\lambda_i(X)$ will denote its ith 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 $A = R^T B R$. 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 Λ is diagonal and Q is orthogonal
- but $A = R^T B R = Q \Lambda Q^T$ 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?

$$\text{If } \{x|x^T Ax \le 1\} \subseteq \{x|x^T Bx \le 1\}, \text{ then } A \ge B.$$

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 $\{x|x^T Bx \le 1\}$ is equal to $\mathbf{R}^n$
- the set $\{x|x^T Ax \le 1\}$ is clearly a subset of $\mathbf{R}^n$, regardless of what A is
- A can be such that $A < B$; $e.g.$, the scalar case $A = -2$, $B = -1$

Example

Let $A \in \mathbf{R}^{n \times n}$ and $B \in \mathbf{R}^{n \times n}$ 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?

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

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

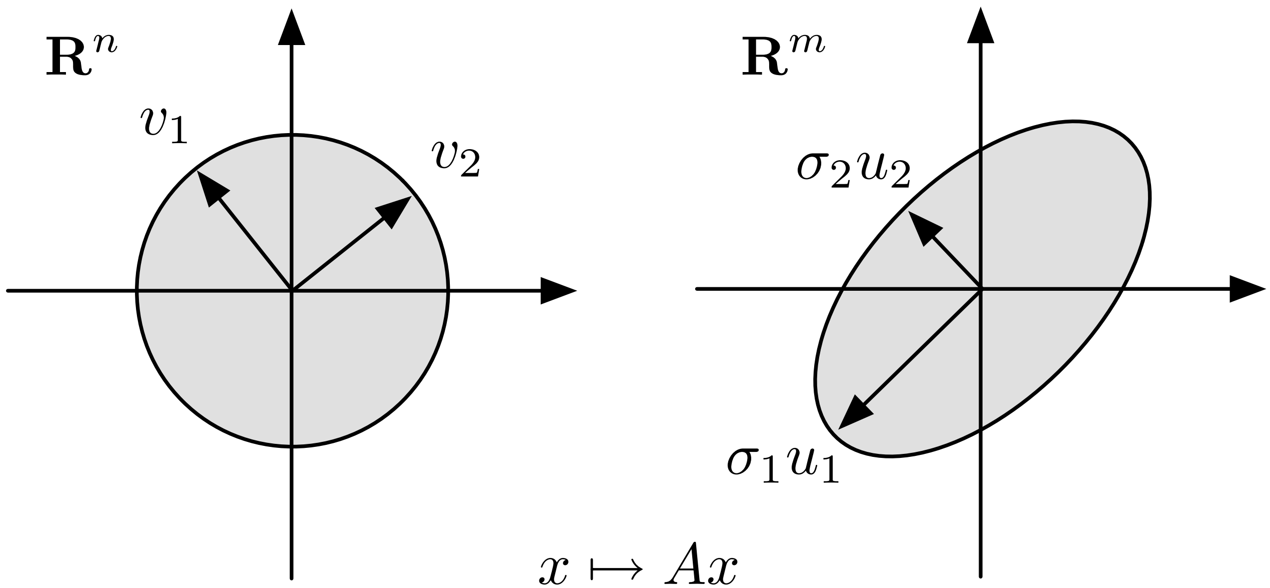
- this is one of those tricky things that is true for scalars, but false for matrices

SVD Fundamentals

- Let $A \in \mathbf{R}^{m \times n}$, $A = U \Sigma V^T = \sum_i \sigma_i u_i v_i^T$ is the singular value decomposition of A.
- $S = \{x \in \mathbf{R}^n : \|x\| \leq 1\}$ is the unit ball (ellipsoid) in $\mathbf{R}^n$,
- $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ is given by the linear mapping $f(x) = Ax$
- $f(x)$ maps the unit ball $S \subseteq \mathbf{R}^n$ to an ellipsoid in $\mathbf{R}^m$

SVD Fundamentals

- Right singular vectors v_i are mapped to left singular vectors u_i.
- Semiaxis lengths given by σ_i.



SVD Properties

- $A = U\Sigma V^T$, where U, V are orthogonal, Σ diagonal.
- $r = \mathbf{rank}(A)$ 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 $\mathbf{rank}\,[y\ A] = \mathbf{rank}(A)$. 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 $\mathrm{span}\{u_1, \ldots, u_r\}$, *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 $A^TA = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^2V^T$ is symmetric
- Similarly, $AA^T = U\Sigma^2U^T$
- Thus, the singular values are the square roots of the eigenvalues of A^TA or AA^T:

$$\sigma_i = \sqrt{\lambda_i(A^TA)} = \sqrt{\lambda_i(AA^T)}$$

- Right singular vectors v_i are the eigenvectors of A^TA, 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 $c \in \mathbf{R}$
- Triangle inequality: $\|A + B\| \leq \|A\| + \|B\|$
- Definiteness: $\|A\| = 0$ if and only if $A = 0$
- Submultiplicative Identity: $\|Ax\| \leq \|A\|\|x\|$ for $x \in \mathbf{R}^n$

Example

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

Solution.

- We know that

$$y^T A x \leq \|y\|\|Ax\| \leq \|y\|\|A\|\|x\| = \|A\|.$$

- This upper bound is achieved by $y = u_1$, and $x = v_1$, which means that

$$\max_{\|y\|=1, \|x\|=1} y^T A x = \|A\| = \sigma_1.$$

Example

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

$$d = \frac{w^TB^TBw}{w^TA^TAw}.$$

Solution.

- We know how to solve the problem in the case $A^TA = I$
- Define $z = (A^TA)^{1/2}w$, so we have $w = (A^TA)^{-1/2}z$. Then we can write

$$\max_{w\neq 0} \frac{w^TB^TBw}{w^TA^TAw} = \max_{z\neq 0} \frac{z^T(A^TA)^{-1/2}B^TB(A^TA)^{-1/2}z}{z^Tz}.$$

- Thus, $d_{\max} = \lambda_{\max}\left((A^TA)^{-1/2}B^TB(A^TA)^{-1/2}\right)$

- 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 $r = \mathbf{rank}(A)$.
- We want to find a matrix $\hat{A}$, with $\mathbf{rank}(\hat{A}) \leq p < r$, 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 $\|A - \hat{A}\| = \left\| \sum_{i=p+1}^{r} \sigma_i u_i v_i^T \right\| = \sigma_{p+1}$