ECE 602: LUMPED LINEAR SYSTEMS

Professor Jianghai Hu

Matrix Norm and Singular Value Decomposition

(Induced) Matrix Norm

Given a matrix $A \in \mathbb{R}^{m\times n}$ (not necessarily square)

- A defines a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$
- Denote by $\|\cdot\|$ the Euclidean (or L^2) norms on both $\mathbb{R}^n$ and $\mathbb{R}^m$
- $\|Ax\|/\|x\|$ is the amplification factor or gain of A in the direction of x

The induced matrix norm of $A \in \mathbb{R}^{m\times n}$ is the maximum possible gain:

$$\|A\| := \sup_{x\in\mathbb{R}^n,\, x\neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{x\in\mathbb{R}^n,\, \|x\|=1} \frac{\|Ax\|}{\|x\|}$$

- $\|\cdot\|$ is called the spectrum norm (or L^2-induced norm)

Example: $A = \begin{bmatrix} 0.5 & 100 \\ 0 & 0.5 \end{bmatrix}$

Properties of Induced Matrix Norm

Induced matrix norm is a norm on the vector space $\mathbb{R}^{m \times n}$:

- **Homogeneity**: $\|\alpha A\| = |\alpha| \, \|A\|$, $\forall \alpha \in \mathbb{R}$
- **Triangle Inequality**: $\|A + B\| \leq \|A\| + \|B\|$
- **Positive Definiteness**: $\|A\| = 0$ if and only if $A = 0$

Moreover, it has the additional properties

- $\|Ax\| \leq \|A\| \, \|x\|$, $\forall x \in \mathbb{R}^n$
- $\|AB\| \leq \|A\| \cdot \|B\|$ (assume the product AB is well defined)

Characterizing Spectrum Norm

For $A \in \mathbb{R}^{m\times n}$, its spectrum norm is

$$\|A\| = \sqrt{\lambda_{\max}(A^T A)}$$

Proof: $\|A\|^2 = \sup_{x\neq 0} \frac{\|Ax\|^2}{\|x\|^2} = \sup_{x\neq 0} \frac{x^T A^T A x}{x^T x} = \lambda_{\max}(A^T A)$ ∎

- Maximum gain input direction is the eigenvector of $A^T A$ for $\lambda_{\max}(A^T A)$
- Alternatively, can switch the positions of A and A^T:

$$\|A\| = \|A^T\| = \sqrt{\lambda_{\max}(AA^T)}$$

Singular Value Decomposition (SVD)

Any matrix $A \in \mathbb{R}^{m \times n}$ can be decomposed as

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

- $\Sigma = \begin{bmatrix} \Sigma_+ & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{m \times n}$ where $\Sigma_+ = \mathrm{diag}\,(\sigma_1, \ldots, \sigma_r)$, $r = \mathrm{rank}\,(A)$
  - $\sigma_1 \geq \cdots \geq \sigma_r > 0$ are the singular values of A
- $U = \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix} \in \mathbb{R}^{m \times m}$ is orthogonal: $U^T U = UU^T = I_m$
  - $u_1, \ldots, u_m$ are the left or output singular vectors of A
- $V = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \in \mathbb{R}^{n \times n}$ is orthogonal: $V^T V = VV^T = I_n$
  - $v_1, \ldots, v_n$ are the right or input singular vectors of A
- Matlab command: `svd`

Transformation Interpretation of SVD

$$A = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_+ & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 & V_2 \end{bmatrix}^T = U_1 \Sigma_+ V_1^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

Consider A as a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$

- maps v_1 to $\sigma_1 u_1$ (most sensitive input/output direction)
- $\cdots$
- maps v_r to $\sigma_r u_r$
- maps $v_{r+1}, \ldots, v_n$ to 0

- Range space of A is $\mathcal{R}(A) = \mathcal{R}(U_1) = \text{span}\,\{u_1, \ldots, u_r\}$
- Null space of A is $\mathcal{N}(A) = \mathcal{R}(V_2) = \text{span}\,\{v_{r+1}, \ldots, v_n\}$
- From $A^T = V_1 \Sigma_+ U_1^T$, we have $\mathcal{R}(A^T) = \mathcal{R}(V_1)$ and $\mathcal{N}(A^T) = \mathcal{R}(U_2)$

Finding U and V in SVD

Write $A^TA = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^T\Sigma V^T$

- v_i is a unit eigenvector of A^TA for eigenvalues $\lambda_i(A^TA)$
- $\sigma_i = \sqrt{\lambda_i(A^TA)}$, $i = 1, \ldots, r$, and $\lambda_i(A^TA) = 0$ for $i = r+1, \ldots, n$

Write $AA^T = (U\Sigma V^T)(U\Sigma V^T)^T = U\Sigma\Sigma^T U^T$

- u_i is a unit eigenvector of AA^T for eigenvalues $\lambda_i(AA^T)$
- $\sigma_i = \sqrt{\lambda_i(AA^T)}$, $i = 1, \ldots, r$, and $\lambda_i(AA^T) = 0$ for $i = r+1, \ldots, m$

Remark:

- Symmetric matrices AA^T and A^TA have the same set of nonzero eigenvalues
- **Spectrum norm of A is the largest singular value of A**:

$$\|A\| = \sigma_1 = \sqrt{\lambda_{\max}(AA^T)} = \sqrt{\lambda_{\max}(A^TA)}$$

Constructive Proof of SVD

1. Find an orthogonal $V \in \mathbb{R}^{n\times n}$ to diagonalize $A^T A$:

$$A^T A = V \Lambda V^T$$

where $\Lambda = \operatorname{diag}\left(\sigma_1^2, \ldots, \sigma_r^2, 0, \ldots, 0\right)$ with $\sigma_1 \geq \cdots \geq \sigma_r > 0$

2. Rewrite

$$A^T A = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \begin{bmatrix} \Sigma_+^2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 & V_2 \end{bmatrix}^T$$

where $V_1 \in \mathbb{R}^{n\times r}$, $V_2 \in \mathbb{R}^{n\times(n-r)}$, and $\Sigma_+ = \operatorname{diag}\left(\sigma_1, \ldots, \sigma_r\right)$

3. Define $U_1 = A V_1 \Sigma_+^{-1} \in \mathbb{R}^{m\times r}$, whose columns are orthogonal: $U_1^T U_1 = I_r$

4. Choose any $U_2 \in \mathbb{R}^{m\times(m-r)}$ so that $U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}$ is orthogonal

5. Then $A = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_+ & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 & V_2 \end{bmatrix}^T$.