EECS 16B

Spring 2022

Lecture 19

3/29/2022

LECTURE 19

- Recap of week before Spring Break
- Minimum norm (energy) solutions
- Singular Value Decomposition (SVD)

- Upper triangularization: Given $A \in \mathbb{R}^{n \times n}$ (with real evalues) we can find orthogonal $U \in \mathbb{R}^{n \times n}$ such that

$$U^{-1}AU = U^T A U \text{ is upper triangular.}$$

- Evalues of A are the diagonal entries of this upper triangular matrix:

$$U^T A U = \begin{bmatrix} \lambda_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} =: T.$$

- Induction proof $\rightarrow$ recursive procedure to find U
- $\frac{d}{dt}\vec{x}(t) = A\vec{x}(t) + B\vec{u}(t) \xrightarrow[\vec{y}=U^T\vec{x}]{} \frac{d}{dt}\vec{y}(t) = \underbrace{U^T A U}_{T}\vec{y}(t) + U^T B\,\vec{u}(t)$

Example: critically damped RLC ckt. $T = \begin{bmatrix} \lambda & * \\ 0 & \lambda \end{bmatrix}$

$e^{\lambda t}$, $te^{\lambda t}$ terms appear in the sol'n

- When A is diagonalizable, we can find V such that

$$V^{-1}AV = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

but V is not necessarily orthogonal. If we upper triangularize we get orthogonal U such that

$$U^{-1}AU = U^T A U = \begin{bmatrix} \lambda_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}.$$

Symmetric matrices get the best of both worlds.

Spectral Theorem: If $A \in \mathbb{R}^{n \times n}$ is symmetric ($A^T = A$) then

(i) Evalues of A are real;

(ii) A is diagonalizable;

(iii) Evectors of A are pairwise orthogonal; therefore, if we choose them to have length=1, then $V = [\vec{v}_1 \cdots \vec{v}_n]$ is an orthogonal matrix:

$$V^{-1}AV = V^TAV = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}.$$

Proof of (ii): upper triangularize: $U^TAU = T$

Look at T^T:

$$T^T = (U^TAU)^T = U^T \underbrace{A^T}_{=A} \underbrace{(U^T)^T}_{=U} = U^TAU = T$$

bc/ A symmetric

T upper triangular and symmetric

$$\Rightarrow T \text{ is diagonal}: \; T = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

Proof of (ii):

$$U^TAU = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

$$\underbrace{UU^T}_{=I}AU = U\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$$A[\vec{u}_1 \cdots \vec{u}_n] = [\vec{u}_1 \cdots \vec{u}_n]\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$$= [\lambda_1\vec{u}_1, \cdots, \lambda_n\vec{u}_n]$$

$$A\vec{u}_i = \lambda_i\vec{u}_i \quad i = 1, \ldots, n$$

i.e. columns of orthogonal matrix U obtained from upper triangularization are an orthonormal basis for $\mathbb{R}^n$

Once we learn Singular Value Decomposition (SVD) we will be able to:

1) Perform "Principal Component Analysis" (PCA), which is an application of SVD in statistics to find most informative directions in a data set (e.g. movie recommendations).

2) Find "minimum norm (energy)" solutions for

$$\underbrace{C}_{\text{given}}\vec{w} = \underbrace{\vec{z}}_{\text{given}} \quad -- (1)$$

where C is a wide matrix, thus, has a nontrivial null space. If a solution $\vec{w}_0$ to (1) exists, then there are INFINITELY MANY OTHERS:

$$\vec{w}_0 + \vec{n}, \quad \vec{n} \in \text{null space of } C$$

is another solution. One way to select a solution is to pick the one that has the least norm $\|\vec{w}\|$:

$$\min_{\vec{w}} \|\vec{w}\|$$

$$\text{s.t. } C\vec{w} = \vec{z}.$$

Why might we want to minimize the norm?

Consider a controllable system: $\vec{x}[i+1] = A\vec{x}[i] + B\underbrace{u[i]}_{\text{single input}}$

and suppose we want to reach $\vec{x}_{target}$ at time step ℓ from $\vec{x}[0]$. Then $u[0], \dots u[\ell-1]$ must be selected such that (Lecture 15):

$$\underbrace{[A^{\ell-1}B, \dots, AB, B]}_{C_\ell} \overbrace{\begin{bmatrix} u[0] \\ \vdots \\ u[\ell-1] \end{bmatrix}}^{\vec{w}} = \overbrace{\vec{x}_{target} - A^{\ell}\vec{x}[0]}^{\vec{z} \in \mathbb{R}^n}.$$

By controllability, column space of C_ℓ is $\mathbb{R}^n$
$\ell \geq n$, so a solution exists. If $\ell > n$, C_ℓ is a wide matrix and we have infinitely many solutions.

Minimum norm solution is a good choice because

$$\|\vec{\omega}\| = \sqrt{u[0]^2 + u[1]^2 + \cdots + u[\ell-1]^2}$$

"Control energy:" not energy in the strict physical sense but a convenient measure of control effort

Example: longitudinal motion control of a car (Lec. 17):

$$\frac{d}{dt}p(t) = v(t) \qquad \frac{d}{dt}v(t) = \frac{1}{RM}u(t)$$

$$\xrightarrow[\text{(Lec. 17)}]{\text{discretize}} \begin{bmatrix} p_d[i+1] \\ v_d[i+1] \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \Delta \\ 0 & 1 \end{bmatrix}}_{A} \begin{bmatrix} p_d[i] \\ v_d[i] \end{bmatrix} + \underbrace{\begin{bmatrix} \frac{\Delta^2}{2RM} \\ \frac{\Delta}{RM} \end{bmatrix}}_{B} u_d[i]$$

Controllable? i.e. are B and AB linearly independent?

$$B = \frac{\Delta}{RM}\begin{bmatrix} \frac{\Delta}{2} \\ 1 \end{bmatrix} \qquad AB = \frac{\Delta}{RM}\begin{bmatrix} \frac{3\Delta}{2} \\ 1 \end{bmatrix}$$

linearly independent

Suppose $\vec{x}[0] = \begin{bmatrix} p_d[0] \\ v_d[0] \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\vec{x}_{target} = \begin{bmatrix} p_{target} \\ 0 \end{bmatrix}$

$$\begin{bmatrix} p_{target} \\ 0 \end{bmatrix} = [A^{\ell-1}B, \ldots AB, B] \begin{bmatrix} u_d[0] \\ \vdots \\ u_d[\ell-1] \end{bmatrix}$$

In theory, can find solution with $\ell=2$

$$\begin{bmatrix} p_{target} \\ 0 \end{bmatrix} = [AB, B] \begin{bmatrix} u_d[0] \\ u_d[1] \end{bmatrix}$$

$$\begin{bmatrix} u_d[0] \\ u_d[1] \end{bmatrix} = \begin{bmatrix} AB, B \end{bmatrix}^{-1} \begin{bmatrix} p_{target} \\ 0 \end{bmatrix}$$

Substitute A, B from above, take inverse, multiply:

$$\begin{bmatrix} u_d[0] \\ u_d[1] \end{bmatrix} = \left(\frac{RM}{\Delta^2} p_{target} \right) \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Assume $RM = 5000$ kg m (e.g. $R \approx 0.3$ m, $M \approx 1600$-1700 kg)

$\Delta = 0.1$ sec., $p_{target} = 1000$ m.

$$\begin{bmatrix} u_d[0] \\ u_d[1] \end{bmatrix} = 5 \cdot 10^8 \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ kg } \frac{m^2}{s^2} \text{ (Nm).}$$

five orders of magnitude larger than the torque your car can deliver

More reasonable: $\ell = 1200$ ($\ell\Delta = 120$ s. $= 2$ min)

$$\begin{bmatrix} p_{target} \\ 0 \end{bmatrix} = \underbrace{\begin{bmatrix} A^{1199}B, \; \ldots \; AB, B \end{bmatrix}}_{C_{1200}} \begin{bmatrix} u[0] \\ \vdots \\ u[1199] \end{bmatrix}$$

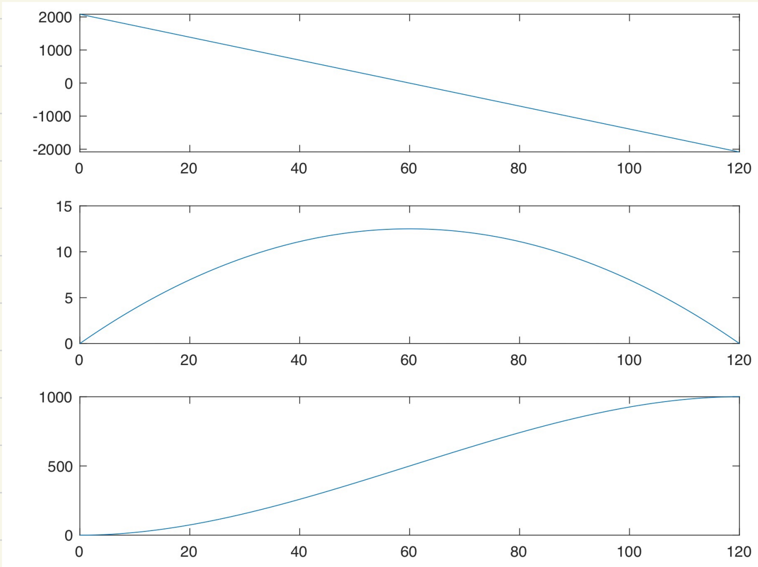
Min. norm solution gives reasonable torque magnitudes.

We haven't learned how to find min. norm sol'n yet (SVD will enable us) but here is what the sol'n looks like:

torque (N·m)

velocity $\left(\frac{m}{s}\right)$

position (m)



time(s)

"

"

SVD:

Question: What is the rank of matrix $\vec{u}\vec{v}^T$, $\vec{u}\neq 0$, $\vec{v}\neq 0$
(column times row, or "outer product")

Answer: 1, because column space spanned by $\vec{u}$:

$$\vec{u}\vec{v}^T = \vec{u}\,[v_1\ v_2 \cdots] = [v_1\vec{u},\ v_2\vec{u}, \cdots\ ]$$

SVD separates a rank-r matrix $A \in \mathbb{R}^{m\times n}$ into a sum of r rank-1 matrices, each written as an outer product. Specifically, we can find:

1) orthonormal vectors $\vec{u}_1, \ldots \vec{u}_r \in \mathbb{R}^m$

2) " " $\vec{v}_1, \ldots \vec{v}_r \in \mathbb{R}^n$

3) real, positive numbers $\sigma_1, \ldots \sigma_r$ such that

$$A = \sigma_1 \vec{u}_1\vec{v}_1^T + \cdots + \sigma_r \vec{u}_r\vec{v}_r^T.$$

(outer-product form of SVD)

$\sigma_1, \dots \sigma_r$ are called "singular values" of A and, by convention they are put in decreasing order:

$$\sigma_1 \geqslant \sigma_2 \geqslant \dots \geqslant \sigma_r > 0.$$

Alternative form of SVD (compact form):

$$A = \underbrace{\begin{bmatrix} \vec{u}_1 \cdots \vec{u}_r \end{bmatrix}}_{=: U_r} \underbrace{\begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{bmatrix}}_{=: \Sigma_r} \underbrace{\begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_r^T \end{bmatrix}}_{=: V_r^T}$$

$m \times n$ | $m \times r$ has orthonormal columns | $r \times r$ | $r \times n$

V: $n \times r$, orthonormal columns