EECS 16B

Spring 2022

Lecture 17

3/15/2022

LECTURE 17 : – upper triangularization

Recall: diagonalization of an $n \times n$ matrix requires n linearly independent evectors: $A\vec{v}_i = \lambda_i \vec{v}_i \quad i=1,\dots,n$

$$A \underbrace{[\vec{v}_1 \cdots \vec{v}_n]}_{=:V} = [\lambda_1\vec{v}_1 \;\cdots\; \lambda_n\vec{v}_n] = \underbrace{[\vec{v}_1 \cdots \vec{v}_n]}_{=V}\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$$\Rightarrow \quad V^{-1} A V = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

Today: we can upper-triangularize even if we can't diagonalize. Upper-triangular form has some of the benefits of diagonal matrices:

1) Evalues are the diagonal entries:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & \\ 0 & a_{22} & & \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix} \qquad \lambda I - A = \begin{bmatrix} \lambda - a_{11} & -a_{12} & \cdots & \vdots \\ 0 & \lambda - a_{22} & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & \lambda - a_{nn} \end{bmatrix}$$

$\lambda = a_{nn}$: bottom row is zero; row space has dim $< n \Rightarrow$ rank drops $\Rightarrow a_{nn}$ is evalue

$\lambda = a_{11}$ zeroes out this column $\Rightarrow$ rank drops therefore a_{11} is an evalue

If $\lambda = a_{ii}$

$1 < i < n$

$$\lambda I - A = \begin{bmatrix} a_{ii}-a_{11} & & * & * & * \\ 0 & \ddots & a_{ii}-a_{i-1,i-1} & * & \\ \vdots & & & 0 & a_{ii}-a_{i+1,i+1} \ \ddots \\ 0 & - & - & - & \end{bmatrix}$$

i columns;

only top $i-1$ entries can be nonzero

$\Rightarrow$ can span at most the $i-1$ dimensional space spanned by first $i-1$ unit vectors

$\Rightarrow$ these i columns are linearly dependent

$\Rightarrow$ Columns of $\lambda I - A$ are linearly dependent

$\Rightarrow$ not full rank $\Rightarrow$ $\lambda = a_{ii}$ is an evalue

2) Sol'n of vector diff. eq. or diffence eq. broken down into scalar equations:

$$\frac{d}{dt}\vec{x}(t) = \begin{bmatrix} a_{11} & * & \cdots & * \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \\ 0 & & 0 & a_{nn} \end{bmatrix} \vec{x}(t) + B\vec{u}(t)$$

$$\frac{d}{dt}x_n(t) = a_{nn}x_n(t) + (B\vec{u}(t))_n$$

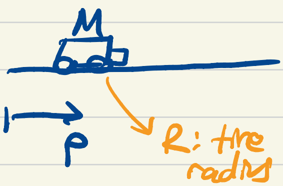
$$x_n(t) = \underbrace{e^{a_{nn}(t-t_0)}x_n(t_0) + \text{term due to } (B\vec{u}(t))_n}_{\text{substitute}}$$

$$\frac{d}{dt}x_{n-1}(t) = a_{n-1,n-1}x_{n-1}(t) + \underbrace{a_{n-1,n}x_n(t) + (B\vec{u}(t))_{n-1}}_{\text{known function of time: treat as input to scalar diff. eq. for } x_{n-1}}$$

$$x_{n-1}(t) = e^{a_{n-1,n-1}(t-t_0)} x_{n-1}(t_0) + \int_{t_0}^{t} \underbrace{- - - -}_{\text{Combined input terms from above go here}}$$

$$\frac{d}{dt} x_{n-2}(t) = a_{n-2,n-2} x_{n-2}(t) + \underbrace{a_{n-2,n-1} x_{n-1}(t) + a_{n-2,n} x_n(t) + (B\vec{u}(t))_{n-2}}_{\text{combined input terms}}$$

Example: longitudinal motion of a car described by:



$$\frac{d}{dt} p(t) = v(t)$$

$$M \frac{d}{dt} v(t) = \frac{1}{R} u(t)$$

p: position
v: velocity
u: torque

$$\frac{d}{dt} \underbrace{\begin{bmatrix} p(t) \\ v(t) \end{bmatrix}}_{\vec{x}(t)} = \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}_{A_c} \begin{bmatrix} p(t) \\ v(t) \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \\ \frac{1}{RM} \end{bmatrix}}_{B_c} u(t)$$

Evalues of A_c : $\{0, 0\}$

Evector : $A_c \vec{v} = 0 \qquad \vec{v} = \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \alpha \neq 0$

no two lin. ind. evectors exist

$\Rightarrow$ not diagonalizable

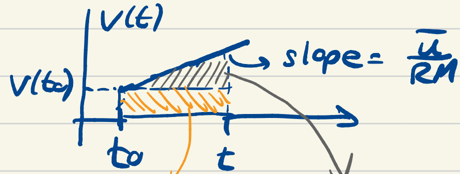
but in upper triangular form

Suppose $u(t) = \bar{u} = \text{constant}$. Solution of diff. eq.?

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

$$v(t) = v(t_0) + \frac{\bar{u}}{RM}(t-t_0)$$

$$\frac{d}{dt}p(t) = v(t) = v(t_0) + \frac{\bar{u}}{RM}(t-t_0)$$



$$p(t) = p(t_0) + \overbrace{v(t_0)(t-t_0)} + \overbrace{\frac{1}{2}\frac{\bar{u}}{RM}(t-t_0)^2}$$

Can find discrete time model if we set

$$t_0 = i\Delta, \quad t = (i+1)\Delta, \quad \bar{u} = ud[i]:$$

$$\underbrace{p((i+1)\Delta)}_{pd[i+1]} = \underbrace{p(i\Delta)}_{pd[i]} + \underbrace{v(i\Delta)}_{vd[i]}\Delta + \frac{\Delta^2}{2RM}ud[i]$$

$$\underbrace{v((i+1)\Delta)}_{vd[i+1]} = \underbrace{v(i\Delta)}_{vd[i]} + \frac{\Delta}{RM}ud[i]$$

$$\begin{bmatrix} pd[i+1] \\ vd[i+1] \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \Delta \\ 0 & 1 \end{bmatrix}}_{Ad} \begin{bmatrix} pd[i] \\ vd[i] \end{bmatrix} + \underbrace{\begin{bmatrix} \frac{\Delta^2}{2RM} \\ \frac{\Delta}{RM} \end{bmatrix}}_{Bd} ud[i]$$

Next: prove any square matrix can be brought to an upper triangular form. Will prove for real matrices with real evalues, but a complex version can be proven along the same lines.

Theorem: For any matrix $A \in \mathbb{R}^{n \times n}$ with real eigenvalues there exists orthogonal matrix $U \in \mathbb{R}^{n \times n}$ such that

$$U^{-1} A U = U^T A U$$

is upper triangular.

Thus, if $\frac{d}{dt}\vec{x}(t) = A\vec{x}(t) + B\vec{u}(t)$, then $\underbrace{\vec{y} = U^T \vec{x}}_{}$ $\quad \vec{x} = U\vec{y}$

satisfies $\frac{d}{dt}\vec{y}(t) = U^T A \underbrace{\vec{x}(t)}_{U\vec{y}} + U^T B \vec{u}(t)$

$$= \underbrace{U^T A U}_{\text{upper triangular}} \vec{y}(t) + U^T B \vec{u}(t)$$

We will prove the theorem by induction.

Proof by induction: Suppose we have a statement that depends on an integer $n = 1, 2, 3, \ldots$ S_n.
To prove it by induction, show:

- S_1 is true
- for any $k \geq 1$, if we assume S_k is true then S_{k+1} is true.

Back to the Theorem: take the statement as S_n.

- Show S_1 is true: true b/c scalars are "upper triangular"
- Show if S_k is true, S_{k+1} is also true.

i.e.,

Assume: any real $k \times k$ matrix with real evalues can be upper triangularized with orthogonal U.

Show: same is true for $(k+1) \times (k+1)$.

1. Let $A \in \mathbb{R}^{(k+1)\times(k+1)}$ be a matrix with real evalues and let $\lambda_1, \vec{q}_1$ be an evalue/evector pair for A. We will assume $\|\vec{q}_1\| = 1$. (No loss of generality because we can normalize the evector if its length is not 1 and it remains an evector.)

2. Choose an orthonormal basis for $\mathbb{R}^{k+1}$ that includes $\vec{q}_1$: $\{\vec{q}_1, \vec{q}_2, \dots \vec{q}_{k+1}\}$.

   How to find them? Pick any k vectors in $\mathbb{R}^{k+1}$ which, when combined with $\vec{q}_1$, form a basis for $\mathbb{R}^{k+1}$. Then apply Gram-Schmidt with $\vec{q}_1$ as the first vector.

3. Then $Q = [\vec{q}_1 \; \dots \; \vec{q}_{k+1}]$ is an orthogonal matrix and

$$AQ = [A\vec{q}_1 \ A\vec{q}_2 \cdots A\vec{q}_{k+1}]$$

$$= [\lambda_1 \vec{q}_1, A\vec{q}_2, \ldots, A\vec{q}_{k+1}]$$

$$Q^T A Q = \begin{bmatrix} \vec{q}_1^{\,T} \\ \vec{q}_2^{\,T} \\ \vdots \\ \vec{q}_{k+1}^{\,T} \end{bmatrix} [\lambda_1 \vec{q}_1 \ A\vec{q}_2 \cdots A\vec{q}_{k+1}]$$

$$= \begin{bmatrix} \lambda_1 \vec{q}_1^{\,T}\vec{q}_1 & * & \cdots & * \\ \lambda_1 \vec{q}_2^{\,T}\vec{q}_1 & \vdots & & \\ \vdots & \vdots & & \\ \lambda_1 \vec{q}_{k+1}^{\,T}\vec{q}_1 & * & \cdots & * \end{bmatrix}$$

$\Big\downarrow$ by orthonormality

$$= \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \\ 0 & * & \cdots & * \end{bmatrix}$$

call this $\vec{p}^{\,T}$

call this A_0 $(k \times k)$

$$Q^T A Q = \begin{bmatrix} \lambda_1 & \vec{p}^{\,T} \\ \vec{0} & A_0 \end{bmatrix} \qquad \text{---}(1)$$

A_0 is not necessarily upper triangular but it is $k \times k$ so, by assumption at the top of last page, $U_0^T A_0 U_0$ upper triangular for some orthogonal U_0.

(Note: we also need to know A_0 has real evalues. That follows because $Q^T A Q$ has the same evalues as A and because evalues of (1) are those of A_0 combined with λ_1.)

Define $U = Q \begin{bmatrix} 1 & 0 \\ 0 & U_0 \end{bmatrix}$ which is orthogonal:

$$U^T = \begin{bmatrix} 1 & 0 \\ 0 & U_0^T \end{bmatrix} Q^T \Rightarrow U^T U = \begin{bmatrix} 1 & 0 \\ 0 & U_0^T \end{bmatrix} \underbrace{Q^T Q}_{=1} \begin{bmatrix} 1 & 0 \\ 0 & U_0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & \underbrace{U_0^T U_0}_{I_{k \times k}} \end{bmatrix} = I_{(k+1)\times(k+1)}.$$

Note $$U^T A U = \begin{bmatrix} 1 & 0 \\ 0 & U_0^T \end{bmatrix} \underbrace{Q^T A Q}_{\begin{bmatrix} \lambda_1 & \vec{p}^T \\ 0 & A_0 \end{bmatrix}} \begin{bmatrix} 1 & 0 \\ 0 & U_0 \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & \vec{p}^T \\ 0 & U_0^T A_0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & U_0 \end{bmatrix} = \underbrace{\begin{bmatrix} \lambda_1 & \vec{p}^T U_0 \\ 0 & \underbrace{U_0^T A_0 U_0}_{\text{upper triangular}} \end{bmatrix}}_{\text{also upper triangular}}$$

<u>Summary</u>: assuming that a $k \times k$ matrix can be upper triangularized, we have shown the same is true for a $(k+1)\times(k+1)$ matrix.