EECS 16B DIS 10B Gaoyue

Upper-triangularization

M: any square matrix $n \times n$

want: $M = U T U^{-1}$

U: orthogonal matrix

T: upper triangular matrix

<u>2x2 case:</u> $M = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}$

let $\vec{v}_1$ be an eigenvector of M with norm $=1$

i.e. $M\vec{v}_1 = \lambda_1 \vec{v}_1$

let $\underset{2\times 2}{U} = [\vec{v}_1 \quad \underset{\text{from Gram-Schmidt}}{R}]$

$\Rightarrow T = U^{-1} M U$

$= \begin{bmatrix} \lambda_1 & \vec{v}_1^T M R \\ 0 & R^T M R \end{bmatrix}$

upper-triangular !

3×3 case: $\quad M \in \mathbb{R}^{3\times3}$

let $\vec{v}_1$ be an eigenvector of M with norm $=1$

let $\underbrace{U}_{3\times3} = [\underbrace{\vec{v}_1}_{3\times1} \quad \underbrace{R}_{3\times2}]$ from Gram-Schmidt

let $T = U^{-1} M U$

$= \ldots$

$$= \begin{bmatrix} \lambda_1 & \vec{v}_1^{\,T} M R \\ 0 & \\ 0 & \boxed{R^T M R} \end{bmatrix} \quad \boxed{R^T M R} \neq Q$$

$(2\times3)(3\times3)(3\times2) = 2\times2$

Based on the 2×2 case, we know

$\exists \underbrace{U_2}_{2\times2}$ s.t. $(U_2)^{-1} \, Q \, U_2$

$= (U_2)^{-1} R^T M R \, U_2$

$= \boxed{U_2^T R^T M R U_2}$

is upper-triangular

$\Rightarrow$ Edit current ~~$U = [\vec{v}_1 \; R]$~~

to be $U = [\vec{v}_1 \;\; R U_2]$ instead

$$T = \underline{U^{-1} M U}$$

$$= \cdots$$

$$= \left[\begin{array}{c|c} \lambda_1 & \vec{v}_1 M R U_2 \\ \hline 0 & \\ 0 & \boxed{U_2^T \underbrace{R^T M R}_{Q} U_2} \end{array}\right] \quad 2\times 2.$$

$$U_2^{-1} Q U_2$$

1. **Towards upper-triangulation by an orthonormal basis**

This problem is a continuation of problem 1 from Discussion 10A.

Recall that in the last discussion we set out to show that any matrix M can be upper triangularized. In particular we want to find the coordinate transformation U such that M becomes upper triangular when represented in this coordinate system:

$$T = U^{-1}MU. \tag{1}$$

In the previous discussion we began with the example of a 3x3 matrix M. We first constructed U by extending the first eigenvector of M, $\vec{v}_1$, into an orthonormal basis using Gram-Schmidt:

$$U = \begin{bmatrix} \underbrace{\vec{v}_1}_{3\times1} & \underbrace{R}_{3\times2} \end{bmatrix}. \tag{2}$$

This gave us the transformed matrix:

$$\underbrace{T}_{3\times3} = \begin{bmatrix} \lambda_1 & \vec{a}^\top \\ \vec{0} & \underbrace{Q}_{2\times2} \end{bmatrix}, \tag{3}$$

which is upper triangular if Q is upper triangular. Then we realized that we could do a similar transformation on Q to get

$$\underbrace{Q}_{2\times2} = \overbrace{\begin{bmatrix} \underbrace{\vec{v}_2}_{2\times1} & \underbrace{Y}_{2\times1} \end{bmatrix}}^{U_2} \begin{bmatrix} \lambda_2 & \vec{b}^\top \\ \vec{0} & P \end{bmatrix} \overbrace{\begin{bmatrix} \vec{v}_2 & Y \end{bmatrix}^\top}^{U_2^T} \tag{4}$$

where $\vec{v}_2$ is the first eigenvector of matrix Q and corresponds to λ_2. At this point, since Q was a 2x2 matrix, we can see that the transformed Q is upper triangular. We then plugged this Q in for M, and simplified to get the final result:

$$M = \begin{bmatrix} \vec{v}_1 & \underbrace{R\vec{v}_2 \quad RY}_{RU_2} \end{bmatrix} \begin{bmatrix} \lambda_1 & \breve{a}_1 & \breve{a}_{\text{rest}} \\ 0 & \lambda_2 & \vec{b} \\ 0 & 0 & P \end{bmatrix} \begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix}^\top. \tag{5}$$

(a) Show that the matrix $\begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix}$ is still orthonormal.

$U = [\vec{v}_1 \quad R]$

$\vec{v}_1$ and $R\vec{v}_2$: $\vec{v}_1^T R\vec{v}_2 = 0$

from Gram-Schmidt,

$\therefore \vec{v}_1^T R = [0 \quad 0]$

$\vec{v}_1$ and RY : $\vec{v}_1^T RY = 0$

$[2\times3][3\times2]$

$R\vec{v}_2$ and RY : $(R\vec{v}_2)^T RY = \vec{v}_2^T R^T RY$

$= \vec{v}_2^T Y = 0$

Normality: $\vec{v}_1^T \vec{v}_1 = 1$

$$(R\vec{v}_2)^T R\vec{v}_2 = \vec{v}_2^{\,T} R^T R\vec{v}_2$$
$$= \vec{v}_2^{\,T}\vec{v}_2 = 1$$

$$(RY)^T RY = Y^T R^T RY$$
$$= Y^T Y = 1$$

(b) We have shown how to upper triangularize a 3×3 and a 2×2 matrix. How can we generalize this process to an nxn matrix?

Any $M_1 \in \mathbb{R}^{n\times n}$

$$M_1 = [\vec{v}_1 \;\; R_1] \begin{bmatrix} \lambda_1 & \vec{a}_1^{\,T} \\ \vec{0} & \underbrace{M_2}_{(n-1)\times(n-1)} \end{bmatrix} \begin{bmatrix} \vec{v}_1^{\,T} \\ R_1^T \end{bmatrix}$$

$$M_2 = [\vec{v}_2 \;\; R_2] \begin{bmatrix} \lambda_2 & \vec{a}_2 \\ \vec{0} & \underbrace{M_3}_{(n-2)\times(n-2)} \end{bmatrix} \begin{bmatrix} \vec{v}_2^{\,T} \\ R_2^T \end{bmatrix}$$

$\ldots$

Continue until $M_i \in \mathbb{R}^{2\times 2}$

$$U_i = [\vec{v}_i \;\; R_i U_{i+1}]$$
$$= [\vec{v}_i \;\; R_i\vec{v}_{i+1} \;\; R_i R_{i+1}]$$

until we have U_1

$$M_1 = U_1 T U_1^T$$
$$T = U_1^T M_1 U_1$$

(c) Show that the characteristic polynomial of square matrix M is the same as that of the square matrix UMU^{-1} for any invertible U.

Δ

Char Poly of M: $\det(M - \lambda I)$

— — of UMU^{-1}: $\det(UMU^{-1} - \lambda I)$

$$\det(UMU^{-1} - \lambda I)$$
$$= \det(UMU^{-1} - \lambda UU^{-1})$$
$$= \det(U(M - \lambda I)U^{-1})$$
$$= \det(U)\det(M - \lambda I)\det(U^{-1})$$

$\det(U) \cdot \det(U^{-1}) = 1$

$$= \det(M - \lambda I)$$

2. **Minimum Energy Control**

In this question, we build up an understanding for how to get the minimum energy control signal to go from one state to another

(a) Consider the scalar system:

$$x[k+1] = ax[k] + bu[k] \tag{21}$$

where $x[0] = 0$ is the initial condition and $u[k]$ is the control input we get to apply based on the current state. Consider if we want to reach a certain state, at a certain time, namely $x[K]$. Write a matrix equation for how a choice of values of $u[k]$ for $k \in \{0, 1, \ldots K-1\}$ will determine the output at time K.

[Hint: write out all the inputs as a vector $\begin{bmatrix} u[0] & u[1] & \cdots & u[K-2] & u[K-1] \end{bmatrix}^T$ *and figure out the combination of* a *and* b *that gives you the state at time* K*.]*

$$X[K] = \begin{bmatrix} b & ab & \cdots & a^{K-2}b & a^{K-1}b \end{bmatrix} \begin{bmatrix} u[K-1] \\ u[K-2] \\ \vdots \\ u[1] \\ u[0] \end{bmatrix}$$

(b) Consider the scalar system:

$$x[k+1] = \overset{a}{1.0}x[k] + \overset{b}{0.7}u[k] \tag{22}$$

where $x[0] = 0$ is the initial condition and $u[k]$ is the control input we get to apply based on the current state. Suppose if we want to reach a certain state, at a certain time, namely $x[K] = 14$. With our dynamics $a = 1$, solve for the best way to get to a specific state $x[K] = 14$, when $K = 10$. When we say **best way** to control a system, we want the sum squared of the inputs to be minimized

$$\operatorname*{argmin}_{u[k]} \sum_{k=0}^{K} u[k]^2.$$

Hint: recall the Cauchy-Schwarz inequality $\langle \vec{a}, \vec{b} \rangle \leq \|\vec{a}\| \|\vec{b}\|$ *where equality holds if* $\vec{a}$ *and* $\vec{b}$ *are linearly dependent.*

$$X[K] = 0.7\,u[K-1] + 0.7\,u[K-2] + \cdots + 0.7\,u[0]$$

$$\parallel$$

$$14 = \underbrace{\begin{bmatrix} 0.7 & 0.7 & \cdots & 0.7 \end{bmatrix}}_{\vec{v}^T} \begin{bmatrix} u[9] \\ \vdots \\ u[1] \\ u[0] \end{bmatrix}$$

$$14 = \vec{v}^T \vec{u} = \langle \vec{u}, \vec{v} \rangle$$

Cauchy-Schwarz: $\underbrace{\langle \vec{u}, \vec{v} \rangle}_{14 \text{ fixed}} \leq \|\vec{u}\| \|\vec{v}\|$ want to minimize

$\langle \vec{u}, \vec{v} \rangle = \|\vec{u}\| \|\vec{v}\|$ when $\vec{u}, \vec{v}$ linearly dependent

$$14 = \begin{bmatrix} 0.7 & \cdots & 0.7 \end{bmatrix} \begin{bmatrix} u[9] \\ \vdots \\ u[0] \end{bmatrix}$$

$$14 = \begin{bmatrix} 0.7 & \cdots & 0.7 \end{bmatrix} \begin{bmatrix} \bar{u} \\ \vdots \\ \bar{u} \end{bmatrix} \qquad \bar{u}$$

$$\bar{u} = 2$$

(c) Consider the scalar system:

$$x[t+1] = 0.5x[t] + 0.7u[t] \tag{23}$$

where $x[0] = 0$ is the initial condition and $u[t]$ is the control input we get to apply based on the current state. Consider if we want to reach a certain state, at a certain time, namely $x[K] = 14$, when $K = 10$. Explain in words the trend of the control input that will be used to solve this problem.

$$x[K] = \underbrace{\begin{bmatrix} 0.7 & 0.7 \cdot 0.5 & \cdots & 0.7 \cdot 0.5^{K-1} \end{bmatrix}}_{\vec{v}} \begin{bmatrix} u[K-1] \\ \vdots \\ u[0] \end{bmatrix}$$

$$= \langle \vec{v}, \vec{u} \rangle$$

Cauchy-Schwarz

$\Rightarrow \vec{u}, \vec{v}$ Linearly dependent

i.e $\vec{u} = \alpha \vec{v}$

$$x[10] = 14 = \langle \vec{v}, \vec{u} \rangle$$

$$= \alpha \langle \vec{v}, \vec{v} \rangle$$

$\approx \alpha \cdot 0.808^2$

$\alpha = 21.42$

$\vec{u} = 21.42\,\vec{v}$