Walkthrough

Wednesday, March 31, 2021 7:43 PM

EECS 16B Designing Information Devices and Systems II
Spring 2021 Discussion Worksheet **Discussion 10B**

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:

$$U = \begin{bmatrix} \vec{v}_1 & R \end{bmatrix}. \tag{2}$$

This gave us the transformed matrix:

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

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

$$Q = \begin{bmatrix} \vec{v}_2 & Y \end{bmatrix} \begin{bmatrix} \lambda_2 & \vec{b}^\top \\ \vec{0} & P \end{bmatrix} \begin{bmatrix} \vec{v}_2 & Y \end{bmatrix}^\top \tag{4}$$

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:

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

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

(b) We have shown how to upper triangularize a 3×3 and a 2×2 matrix. How can we generalize this process to any $n \times n$ matrix M?

(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.

Show $n \times n$ $M \longrightarrow T$ upper triangular

$T = U^T M U$

3×3

$[\vec{v}_1 \ R] = U$ (GS)

$T = \begin{bmatrix} \lambda_1 & \vec{a}^T \\ \vec{0} & \boxed{Q} \end{bmatrix}$ ← 2×2, Q ← $R^T M R$

$R^T M R = Q = [\vec{v}_2 \ Y] \begin{bmatrix} \lambda_2 & \vec{b}^T \\ 0 & P \end{bmatrix} [\vec{v}_2 \ Y]^T$ (U_2, T_2, P: 1×1)

$U_2^T R^T M R U_2 = U_2^T U_2 T_2 U_2^T U_2$

$T_2 = (U_2^T R^T) M (R U_2)$

$T_2 = (R U_2)^T M (R U_2)$

$Q = R^T M R$

$[\vec{v}_1 \ RU_2] = [\vec{v}_1 \ R\vec{v}_2 \ RY] = U$ ← still orthonormal?

$T = \begin{bmatrix} \lambda_1 & \vec{a}_1 & \vec{a}_{rest} \\ 0 & \lambda_2 & \vec{b}^T \\ 0 & 0 & P \end{bmatrix}$

1a

$[\vec{v}_1 \ \boxed{R\vec{v}_2} \ RY]$ ← Show orthonormal

$\langle \vec{v}_1, R\vec{v}_2 \rangle = \vec{v}_1^T R \vec{v}_2 = [0 \ 0] \vec{v}_2 = 0$

$\langle \vec{v}_1, RY \rangle = \vec{v}_1^T R Y = [0 \ 0] Y = [0 \ 0]$

$\langle R\vec{v}_2, RY \rangle = \vec{v}_2^T \underbrace{R^T R}_{I} Y = \vec{v}_2^T Y = \vec{0}$

$\langle RY, RY \rangle = Y^T \underbrace{R^T R}_{I} Y = Y^T Y = I$

$\| R\vec{v}_2 \| = \sqrt{\langle R\vec{v}_2, R\vec{v}_2 \rangle} = \sqrt{1} = 1$

$\vec{v}_2^T \underbrace{R^T R}_{I} \vec{v}_2 = \vec{v}_2^T \vec{v}_2 = 1$

$U = [\vec{v}_1 \ R\vec{v}_2 \ RY]$ is orthonormal ✓

1b

$n \times n$ M find U (orthonormal) s.t.

$U^T M U = T$ ← upper triangular

$M_n = M$

$\underbrace{[\vec{v}_1 \ R_n]}_{U_n} \begin{bmatrix} \lambda_1 & \vec{a}_1^T \\ 0 & \boxed{M_{n-1}} \end{bmatrix} [\vec{v}_1 \ R_n]^T = M_n$ (T, T_n)

$M_{n-1} \to U_{n-1} = [v_2 \ R_{n-1}]$ (GS; v_2 ← first eigenvec (M_{n-1})) $\quad T_{n-1} = \begin{bmatrix} \lambda_2 & \vec{a}_2 \\ 0 & M_{n-2} \end{bmatrix}$

1) For $M_n \dots M_2$ find the corresponding $U_n \dots U_2$

$U_i = [v_{n-i-1} \ R_i]$ (v_{n-i-1} ← first eigenvec M_i) $\quad T_i = \begin{bmatrix} \lambda_{n-i-1} & \vec{a}_i \\ \vec{0} & \boxed{M_{i-1}} \end{bmatrix} = U_i^T M_i U_i$ (M_{i-1} ← $R_i^T M_i R_i$)

Now we have

$U_2 = [\vec{v}_{n-1} \ R_2] = U_{2f}$

$T_2 = \begin{bmatrix} \lambda_{n-1} & \vec{a}_2 \\ \vec{0} & P \end{bmatrix}$ (2×2)

U_{3f} should upper triangulate M_3 (3×3)

$U_3 = [\vec{v}_{n-2} \ R_3]$

$U_{3f} = [\vec{v}_{n-2} \ R_3\vec{v}_{n-1} \ R_3 R_2] = [\vec{v}_{n-2} \ R_3 U_{2f}]$

$U_{4f} = [\vec{v}_{n-3} \ R_4\vec{v}_{n-2} \ R_4 R_3 U_{2f}]$

$= [\vec{v}_{n-3} \ R_4 U_{3f}]$ (4×4)

$U_{if} = [\vec{v}_{n-i-1} \ R_i U_{(i-1)f}]$

$\boxed{U_{nf}} = [\vec{v}_1 \ R_1 U_{(n-1)f}] = U$ $\quad M_n$ ← n×n

$T = U^T M U$ ← upper triangular form

1)c) Show $\text{poly}(M) = \text{poly}(UMU^{-1})$ where U is invertible

$\det(M - \lambda I) = \det(UMU^{-1} - \lambda I)$ $\quad \det(AB) = \det(A)\det(B)$

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

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

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

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

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

→ $x[k+1] = a x[k] + b u[k]$

desired destination $x[K] = \gamma$

what is the smallest set of inputs $u[0]$

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:

$$x[k+1] = ax[k] + bu[k] \tag{6}$$

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.]*

(b) Consider the scalar system:

$$x[k+1] = 1.0x[k] + 0.7u[k] \tag{7}$$

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

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

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.

(c) Consider the scalar system:

$$x[t+1] = 0.5x[t] + 0.7u[t] \tag{8}$$

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.

Contributors:

- Yuxun Zhou.
- Edward Wang.
- Anant Sahai.
- Sanjit Batra.
- Pavan Bhargava.
- Kareem Ahmad.
- Nathan Lambert.

$\rightarrow$ $x[k+1] = ax[k] + bu[k]$

desired destination $x[K] = \gamma$

what is the smallest set of inputs $u[0] \ldots u[K-1]$

$x[K] = \gamma$

$\underset{u[k]}{\text{argmin}} \sum_{k=0}^{K} (u[k])^2 = \|\vec{u}\|^2$

a) $x[K]$ = in terms of inputs

$x[1] = ax[0] + bu[0] = bu[0]$

$x[2] = ax[1] + bu[1]$

$= a(bu[0]) + bu[1]$

$x[3] = a^2 bu[0] + abu[1] + bu[2]$

$x[K] = a^{K-1} bu[0] + a^{K-2} bu[1] + \ldots$

$abu[K-2] + bu[K-1]$

$x[k+1] = \sum_{i=0}^{K-1} a^i b u[k-i]$

$= \langle [b \;\; ab \ldots a^{K-1} b]^T, \vec{u} \rangle$

B

$x[10] = 14 = \langle [0.7 \;\; 0.7 \ldots 0.7]^T, \vec{u} \rangle$

$\vec{u} = \alpha \cdot [1 \ldots 1]^T$

$14 = \langle \vec{v}, \alpha \vec{1} \rangle = \alpha \vec{v}^T \vec{1} = 7\alpha$

$\alpha = 2$

$\vec{u} = \begin{bmatrix} 2 \\ 2 \\ \vdots \\ 2 \end{bmatrix}$

$u[k] = 2$ for $k \in \{0 \ldots 9\}$

Cauchy Schwarz $\langle \vec{a}, \vec{b} \rangle \leq \|\vec{a}\| \|\vec{b}\|$

equality only holds when $\vec{a}$ and $\vec{b}$ are linearly dependent

C $x[k+1] = 0.5x[k] + 0.7u[k]$ $x[10] = 14$

$x[K] = \langle [0.7 \;\; 0.7/2 \;\; 0.7/4 \ldots 0.7/2^{K-1}]^T, \vec{u} \rangle$

$x[10] = 14 = \langle \vec{v}, \alpha\vec{v} \rangle = \alpha \cdot \vec{v}^T\vec{v} = \alpha \cdot \|\vec{v}\|^2$

$= \alpha (.808)^2$

$\alpha = \frac{14}{(.808)^2} \approx 21.44$ $\vec{u} = \alpha \cdot \vec{v}$

$\vec{u} \approx 21.44 [0.7 \;\; 0.7/2 \;\; 0.7/4 \ldots 0.7/2^9]^T$

$\vec{u} = 15 [(\frac{1}{2})^0 \; (\frac{1}{2})^1 \; (\frac{1}{2})^2 \ldots (\frac{1}{2})^9]^T$

$u[0] = 15(\frac{1}{2})^9$

$\vec{u}[k] = 15 \cdot 0.5^{(9-k)}$