Upper-triangularization

\[ M: \text{any square matrix } n \times n \]

Want: \[ M = UTU^\dagger \]

\( U: \text{orthogonal matrix} \)

\( T: \text{upper triangular matrix} \)

2x2 case: \[ M = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \]

let \( \vec{v}: \text{be an eigenvector of } M \text{ with } \|v\| = 1 \)

i.e. \( M\vec{v} = \lambda \vec{v} \)

let \( U = \begin{bmatrix} \vec{v} & R \end{bmatrix} \]

2x2 from Gram-Schmidt

\[ T = U^\dagger MU \]

\[ = \begin{bmatrix} \lambda & \vec{v}^TMR \\ 0 & R^TMR \end{bmatrix} \]

upper-triangular!
\[ 3 \times 3 \text{ case: } M \in \mathbb{R}^{3 \times 3} \]

Let \( \tilde{v}_i \) be an eigenvector of \( M \) with norm = 1.

Let \( U = [\tilde{v}_i \ R] \) from Crank-Schmidt.

\[
U = \begin{bmatrix} \tilde{v}_i \ R \end{bmatrix} \quad 3 \times 3 \quad 3 \times 1 \quad 3 \times 2
\]

Let \( T = U^{-1} M U \)

\[
T = \begin{bmatrix} \lambda_i & \tilde{v}_i^T MR \\ 0 & R^T MR \end{bmatrix} = Q
\]

\((2 \times 3) (3 \times 3) (3 \times 2) = 2 \times 2\)

Based on the 2x2 case, we know

\[
\exists U_2 \text{ s.t. } (U_2)^{-1} Q U_2 = 2 \times 2 \quad = (U_2)^{-1} R^T MR U_2 = U_2^T R^T MR U_2
\]

is upper-triangular.

\[
\Rightarrow \text{ Edit current } U = [\tilde{v}_i \ R] \text{ to be } U = [\tilde{v}_i \ R U_2] \text{ instead}
\]
\[ T = U^T M U \]
\[ = \begin{bmatrix} \lambda_1 & \bar{U} M R U z \\ 0 & \text{red box} \end{bmatrix} \]
\[ = \begin{bmatrix} \lambda_1 & \bar{U} M R U z \\ 0 & U z R M R U z \end{bmatrix} \]
\[ = \begin{bmatrix} \lambda_1 & \bar{U} M R U z \\ 0 & \text{red box} \end{bmatrix} \]
\[ = U z^T Q U z \]
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:

$$T = U^{-1}MU.$$  \hspace{1cm} (1)

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

$$U = \begin{bmatrix} \vec{v}_1 & R \end{bmatrix}.$$  \hspace{1cm} (2)

This gave us the transformed matrix:

$$T = \begin{bmatrix} \lambda_1 & \vec{a}^T \\ 0 & \vec{Q} \end{bmatrix}.$$  \hspace{1cm} (3)

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

$$Q = \begin{bmatrix} \vec{v}_2 & Y \end{bmatrix} \begin{bmatrix} \lambda_2 & \vec{b}^T \\ 0 & \vec{P} \end{bmatrix} \begin{bmatrix} \vec{v}_2 & Y^T \end{bmatrix}.$$  \hspace{1cm} (4)

where $\vec{v}_2$ is the first eigenvector of matrix $Q$ and corresponds to $\lambda_2$. At this point, since $Q$ was a $2 \times 2$ 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:

$$M = \begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix} \begin{bmatrix} \lambda_1 & \vec{a}_1 & \vec{a}_{rest} \\ 0 & \lambda_2 & \vec{b} \\ 0 & 0 & \vec{P} \end{bmatrix} \begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix}^T.$$  \hspace{1cm} (5)

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

From Gram-Schmidt:

$$R\vec{v}_2 \text{ and } RY : (R\vec{v}_2)^T RY = \vec{v}_2^T Y = 0$$

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

$$\vec{v}_1 \text{ and } \vec{v}_2 : \vec{v}_1^T \vec{v}_2 = 1$$

Normality: $\vec{v}_1^T \vec{v}_2 = 1$.
\[(R\tilde{v}_2)^T R\tilde{v}_2 = \tilde{v}_2^T R^T R \tilde{v}_2 = \tilde{v}_2^T \tilde{v}_2 = 1\]

\[(RY)^T RY = Y^T R^T RY = Y^T Y = 1\]

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

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

\[M_i = \begin{bmatrix} \tilde{v}_i & R_i \end{bmatrix} \begin{bmatrix} \lambda_i & \tilde{a}_i^T \\ 0 & M_2 \end{bmatrix} \begin{bmatrix} \tilde{v}_i^T \\ R_i^T \end{bmatrix} \]

\[M_2 = \begin{bmatrix} \tilde{v}_2 & R_2 \end{bmatrix} \begin{bmatrix} \lambda_2 & \tilde{a}_2^T \\ 0 & M_3 \end{bmatrix} \begin{bmatrix} \tilde{v}_2^T \\ R_2^T \end{bmatrix} \]

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

\[U_i = \begin{bmatrix} \tilde{v}_i & R_i U_{i+1} \end{bmatrix} = \begin{bmatrix} \tilde{v}_i & R_i \tilde{v}_{i+1} & R_i R_{i+1} \end{bmatrix} \]

Until we have $U_i$

\[M_i = U_i^T \Lambda_i U_i^T \]

\[T = U_i^T M_i U_i \]

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

\[\boxed{UMU^{-1}}\]
The characteristic polynomial of \( M \) is:

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

\[
= \det(U M U^{-1} - \lambda U U^{-1})
\]

\[
= \det(U(M - \lambda I)U^T)
\]

\[
= \det(U) \det(M - \lambda I) \det(U^T)
\]

\[
\det(U) \cdot \det(U^T) = 1
\]

\[
= \det(M - \lambda I)
\]

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:

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

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 \( [u[0] \ u[1] \ \cdots \ u[K - 2] \ u[K - 1]]^T \) and figure out the combination of \( a \) and \( b \) that gives you the state at time \( K \).*
(b) Consider the scalar system:

\[ x[k+1] = 1.0x[k] + 0.7u[k] \]

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

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

**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.

\[
\begin{align*}
X[k] &= \begin{bmatrix} 0.7 & 0.7 & \cdots & 0.7 \end{bmatrix} \begin{bmatrix} u[k-1] \\ u[k-2] \\ \vdots \\ u[0] \end{bmatrix} \\
14 &= \begin{bmatrix} 0.7 & 0.7 & \cdots & 0.7 \end{bmatrix} \begin{bmatrix} u[9] \\ u[8] \\ \vdots \\ u[0] \end{bmatrix} \\
14 &= \vec{v}^\top \tilde{u} = \langle \tilde{u}, \vec{v} \rangle
\end{align*}
\]

Cauchy-Schwarz: \( \langle \tilde{u}, \vec{v} \rangle \leq ||\tilde{u}|| ||\vec{v}|| \)

\( 14 \) is fixed want to minimize

\( \langle \tilde{u}, \vec{v} \rangle = ||\tilde{u}|| ||\vec{v}|| \) when \( \tilde{u}, \vec{v} \) linearly dependent
\[ \begin{bmatrix} x[t + 1] \end{bmatrix} = 0.5 x[t] + 0.7 u[t] \]  

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.

\[ x[K] = \begin{bmatrix} 0.7 \cdot 0.5^{10} & \cdots & 0.7 \cdot 0.5^{K-1} \end{bmatrix} \begin{bmatrix} u[K-1] \\ \vdots \\ u[0] \end{bmatrix} \]

\[ = \langle \vec{u}, \vec{u} \rangle \]

Cauchy-Schwarz

\[ \Rightarrow \vec{u}, \vec{u} \text{ linear dependent} \]

i.e. \( \vec{u} = \alpha \vec{u} \)

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

\[ = \alpha \langle \vec{u}, \vec{u} \rangle \]
\[ \alpha \approx 0.9 \cdot 0.8^2 \]

\[ \alpha = 21.42 \]

\[ \vec{w} = 21.42 \vec{v} \]