

Exam Location: Draft

PRINT your student ID: _____

PRINT AND SIGN your name: _____, _____
(last) (first) (sign)

PRINT your discussion sections and (u)GSIs (the ones you attend): _____

Row Number: _____ Seat Number: _____

Name and SID of the person to your left: _____

Name and SID of the person to your right: _____

Name and SID of the person in front of you: _____

Name and SID of the person behind you: _____

1. Honor Code (0 pts.)

Please copy the following statement in the space provided below and sign your name.

As a member of the UC Berkeley community, I act with honesty, integrity, and respect for others. I will follow the rules and do this exam on my own.

Note that if you do not copy the honor code and sign your name, you will get a 0 on the exam.

Solution: Any attempt to copy the honor code and sign should get full points.

2. What are you planning to do during your summer break? (2 pts.)

Solution: Any answer is sufficient.

3. What is the happiest moment of your semester so far? (2 pts.)

Solution: Any answer is sufficient.

Do not turn this page until the proctor tells you to do so.
You can work on the above problems before time starts.

4. Orthonormality, Gram-Schmidt, and SVD Potpourri (21 pts.)

- (a) (2 pts.) Suppose $Q \in \mathbb{R}^{m \times n}$ is tall (i.e. $m \geq n$) matrix and has orthonormal columns, which of the following is true:

<input type="radio"/>	$Q^T Q = I_n$
<input type="radio"/>	$Q Q^T = I_m$
<input type="radio"/>	neither
<input type="radio"/>	both

Solution: Orthonormal columns in a tall matrix only imply $Q^T Q = I_n$.

- (b) (2 pts.) Suppose $Q \in \mathbb{R}^{n \times n}$ is a square, orthonormal matrix, which of the following is true:

<input type="radio"/>	$Q^T Q = I_n$
<input type="radio"/>	$Q Q^T = I_n$
<input type="radio"/>	neither
<input type="radio"/>	both

Solution: A square orthonormal matrix implies both $Q^T Q = I_n$ and $Q Q^T = I_n$. Thus the answer to this question is **both**.

- (c) (2 pts.) Suppose $Q \in \mathbb{R}^{m \times n}$ is a wide (i.e. $m \leq n$) matrix and has orthonormal columns, which of the following is true:

<input type="radio"/>	$Q^T Q = I_n$
<input type="radio"/>	$Q Q^T = I_m$
<input type="radio"/>	neither
<input type="radio"/>	both

Solution: The original intention of this problem was to say "orthonormal rows" instead of "orthonormal columns" in which case the answer would be $Q Q^T = I_m$.

However, since a wide matrix with orthonormal columns is not possible, any answer earns full credit.

- (d) (5 pts.) Using Gram-Schmidt, find an orthonormal basis for \mathbb{R}^2 starting with the vectors $\vec{q}_1 = \begin{bmatrix} 3 \\ \sqrt{3} \end{bmatrix}$ and $\vec{q}_2 = \begin{bmatrix} -\sqrt{3} \\ 3 \end{bmatrix}$. Let \vec{v}_1 be the corresponding orthonormal basis vector for \vec{q}_1 and \vec{v}_2 be the corresponding orthonormal basis vector for \vec{q}_2 .

Solution: We run Gram-Schmidt as follows:

$$\vec{v}_1 = \frac{\vec{q}_1}{\|\vec{q}_1\|} \tag{1}$$

$$= \frac{1}{2\sqrt{3}} \begin{bmatrix} 3 \\ \sqrt{3} \end{bmatrix} \tag{2}$$

$$\tag{3}$$

$$\vec{z}_2 = \vec{q}_2 - \langle \vec{q}_2, \vec{v}_1 \rangle \vec{v}_1 \tag{4}$$

$$= \vec{q}_2 \tag{5}$$

(6)

$$\vec{v}_2 = \frac{\vec{q}_2}{\|\vec{q}_2\|} \tag{7}$$

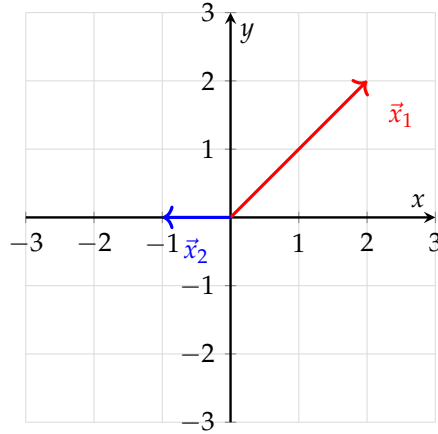
$$= \frac{1}{2\sqrt{3}} \begin{bmatrix} -\sqrt{3} \\ 3 \end{bmatrix} \tag{8}$$

(9)

Thus, our solution is given as: $\vec{v}_1 = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix}$

Suppose \vec{v}_1 and \vec{v}_2 are actually the columns of V^\top from the SVD of some arbitrary matrix A (i.e. $A = U\Sigma V^\top$). In the next several problem parts, we will evaluate how this matrix A transforms two vectors \vec{x}_1 and \vec{x}_2 .

(e) (5 pts.) The following is a graph of $\vec{x}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ and $\vec{x}_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$.



As a first step, draw $V^\top \vec{x}_1$ and $V^\top \vec{x}_2$ on the following graph.

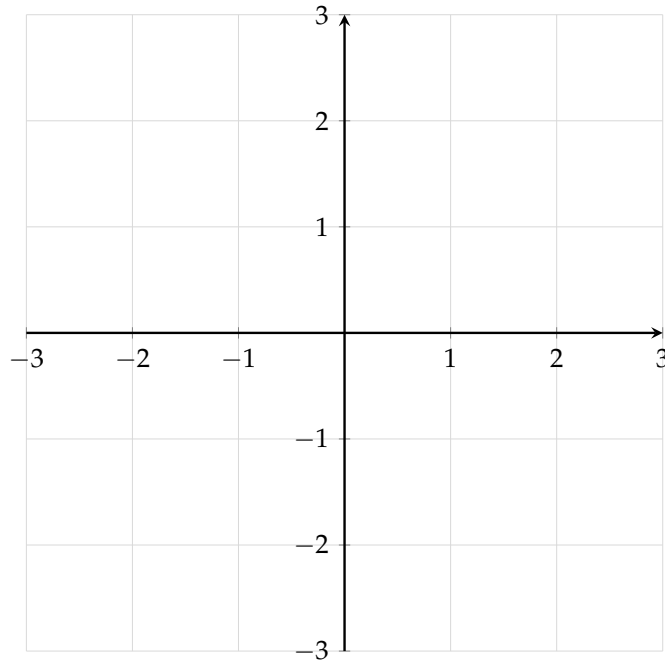
Use the fact that the rotation matrix for a counterclockwise rotation of θ (in degrees or radians)

is given by $R = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$.

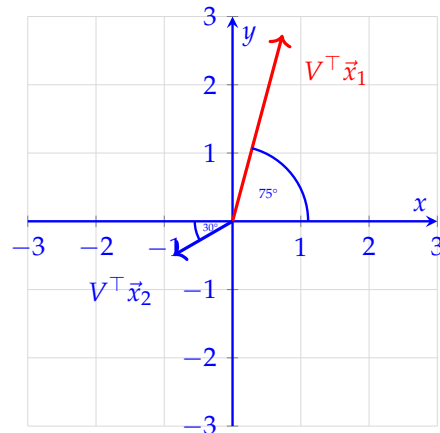
The following trigonometric values may also be of use:

	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$
$\sin(\theta)$	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$
$\cos(\theta)$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$

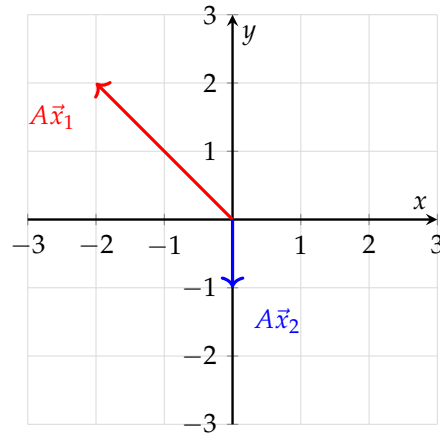
A blank plot is provided for you on the next page. *Note: In your graphs, indicate the angle of your vectors with respect to the x-axis.*



Solution: First, we have to identify how much V^T rotates a vector by matching it with the rotation matrix. We notice from inspection that $V^T = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} = R_{\frac{\pi}{6}}$. Thus, V^T imposes a $\frac{\pi}{6}$ radian rotation counterclockwise, which results in the following graph:



- (f) (5 pts.) The following graph represent how the A -matrix affects the vectors \vec{x}_1 and \vec{x}_2 where $A\vec{x}_1 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$ and $A\vec{x}_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$.



Find Σ and U .

Solution: Recall, that after the application of V^\top (which is a rotation), Σ applies a scaling to the vector, while U is another rotation.

We can first note that the scale of each of the vectors doesn't change when A is applied to each.

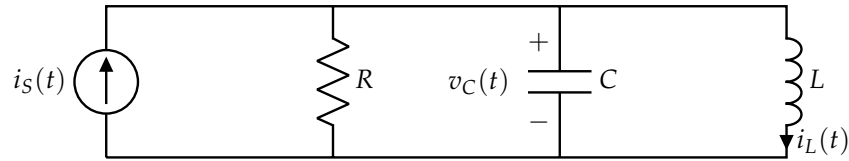
$$\text{Thus, } \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Next, we take a look at the final rotation between $V^\top \vec{x}_i$ and $A\vec{x}_i$. We notice that the rotation is a 60-degree or $\frac{\pi}{3}$ -radian rotation counterclockwise.

$$\text{Thus, } U = R_{\frac{\pi}{3}} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}.$$

5. Critically Damped Parallel RLC Circuit (15 pts.)

In this problem, we will use vector differential equations to analyze the following circuit:



This circuit can have different behavior depending on the values of the elements. In this case, we will examine the critically damped case, where the state variables ($v_C(t)$ and $i_L(t)$) decay over time without oscillation, just like the overdamped case, but with repeated eigenvalues, which could result in a system that cannot be diagonalized.

- (a) (6 pts.) **Solve for A and \vec{b}** in terms of R , L , and C such that the differential equation for $\vec{x}(t) = \begin{bmatrix} v_C(t) \\ i_L(t) \end{bmatrix}$ can be written in the following form:

$$\frac{d}{dt} \vec{x}(t) = A \vec{x}(t) + \vec{b} i_S(t) \quad (10)$$

Solution: To find the differential equation for $\vec{x}(t)$, we find the differential equations for $v_C(t)$ and $i_L(t)$ using our various circuit analysis equations.

To start off, we know that the voltage across the inductor is $v_L(t)$ and thus, using the inductor equation $v_L(t) = L \frac{d}{dt} i_L(t)$:

$$L \frac{d}{dt} i_L(t) = v_C(t) \quad (11)$$

$$\frac{d}{dt} i_L(t) = \frac{1}{L} v_C(t) \quad (12)$$

Then, we know that the current through the capacitor is $i_S(t) - i_L(t) - \frac{v_C(t)}{R}$, and thus, using the capacitor equation $i_C(t) = C \frac{d}{dt} v_C(t)$:

$$C \frac{d}{dt} v_C(t) = i_S(t) - i_L(t) - \frac{v_C(t)}{R} \quad (13)$$

$$\frac{d}{dt} v_C(t) = -\frac{1}{RC} v_C(t) - \frac{1}{C} i_L(t) + \frac{1}{C} i_S(t) \quad (14)$$

Thus, we can construct the vector differential equation:

$$\frac{d}{dt} \begin{bmatrix} v_C(t) \\ i_L(t) \end{bmatrix} = \begin{bmatrix} -\frac{1}{RC} & -\frac{1}{C} \\ \frac{1}{L} & 0 \end{bmatrix} \begin{bmatrix} v_C(t) \\ i_L(t) \end{bmatrix} + \begin{bmatrix} \frac{1}{C} \\ 0 \end{bmatrix} i_S(t) \quad (15)$$

Thus, $A = \begin{bmatrix} -\frac{1}{RC} & -\frac{1}{C} \\ \frac{1}{L} & 0 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} \frac{1}{C} \\ 0 \end{bmatrix}$.

- (b) (3 pts.) Suppose the system from the previous part has a repeated eigenvalue with only one eigenvector and is thus not diagonalizable. We will use upper triangularization to solve it. You are provided an orthonormal change of basis matrix U such that $\vec{x}(t) = U\vec{\tilde{x}}(t)$ which allows for upper triangularization of the vector differential equation with $\vec{\tilde{x}}$.

Solve for \tilde{A} and \tilde{b} in terms of A , U , and \vec{b} such that the differential equation for $\vec{\tilde{x}}(t) = \begin{bmatrix} \tilde{v}_C(t) \\ \tilde{i}_L(t) \end{bmatrix}$ can be written in the following form:

$$\frac{d}{dt}\vec{\tilde{x}}(t) = \tilde{A}\vec{\tilde{x}}(t) + \tilde{b}i_S(t) \quad (16)$$

such that \tilde{A} is an upper triangular matrix. **You are NOT allowed to use U^{-1} in your answer.**

Solution: We can see how the differential equation will change after the change of basis by replacing $\vec{x}(t)$ with $U\vec{\tilde{x}}(t)$ (remember that since U is an orthonormal matrix, $U^{-1} = U^\top$).

$$\frac{d}{dt}\vec{x}(t) = A\vec{x}(t) + \vec{b}i_S(t) \quad (17)$$

$$\frac{d}{dt}(U\vec{\tilde{x}}(t)) = AU\vec{\tilde{x}}(t) + \vec{b}i_S(t) \quad (18)$$

$$U\frac{d}{dt}\vec{\tilde{x}}(t) = AU\vec{\tilde{x}}(t) + \vec{b}i_S(t) \quad (19)$$

$$\frac{d}{dt}\vec{\tilde{x}}(t) = U^\top AU\vec{\tilde{x}}(t) + U^\top\vec{b}i_S(t) \quad (20)$$

Thus, $\tilde{A} = U^\top AU$ and $\tilde{b} = U^\top\vec{b}$.

For the rest of the problem, we will let $i_S(t) = 0$.

- (c) (3 pts.) Regardless of your answers to the previous parts, assume that we are able to reach the following vector differential equation (again, with $i_S(t) = 0$):

$$\frac{d}{dt}\vec{\tilde{x}}(t) = \begin{bmatrix} -5 & -10 \\ 0 & -5 \end{bmatrix} \vec{\tilde{x}}(t) \quad (21)$$

Solve for $\tilde{i}_L(t)$ with the initial condition $\tilde{i}_L(0) = 2$. Remember that $\vec{\tilde{x}}(t) = \begin{bmatrix} \tilde{v}_C(t) \\ \tilde{i}_L(t) \end{bmatrix}$.

Solution: The differential equation for $\tilde{i}_L(t)$ from the provided vector differential equation is

$$\frac{d}{dt}\tilde{i}_L(t) = -5\tilde{i}_L(t) \quad (22)$$

Notice that the differential equation does not depend on $\tilde{v}_C(t)$! We have solved this first order differential equation before, with the solution being

$$\tilde{i}_L(t) = \tilde{i}_L(0)e^{-5t} = 2e^{-5t} \quad (23)$$

- (d) (3 pts.) Using your answer from the previous part, **set up a first order differential equation for $\tilde{v}_C(t)$.** You DO NOT have to solve it.

Solution: The differential equation for $\tilde{v}_C(t)$ from the provided vector differential equation is

$$\frac{d}{dt}\tilde{v}_C(t) = -5\tilde{v}_C(t) - 10\tilde{i}_L(t) \quad (24)$$

We can use our solution for $\tilde{i}_L(t)$ to obtain the desired first order differential equation:

$$\frac{d}{dt}\tilde{v}_C(t) = -5\tilde{v}_C(t) - 10(2e^{-5t}) \quad (25)$$

$$\frac{d}{dt}\tilde{v}_C(t) = -5\tilde{v}_C(t) - 20e^{-5t} \quad (26)$$

6. Using PCA to Detect Fraudulent Transactions (15 pts.)

PCA has many different uses when applied to real-world data. One potential application is making classification of data much easier.

Suppose we are given some data, where each datapoint represents a transaction. Each one is labeled either normal or fraudulent. We will utilize PCA to develop a useful classifier.

We plot the data in two dimensions, where each dimension is some unspecified feature that will aid us in classifying the points:

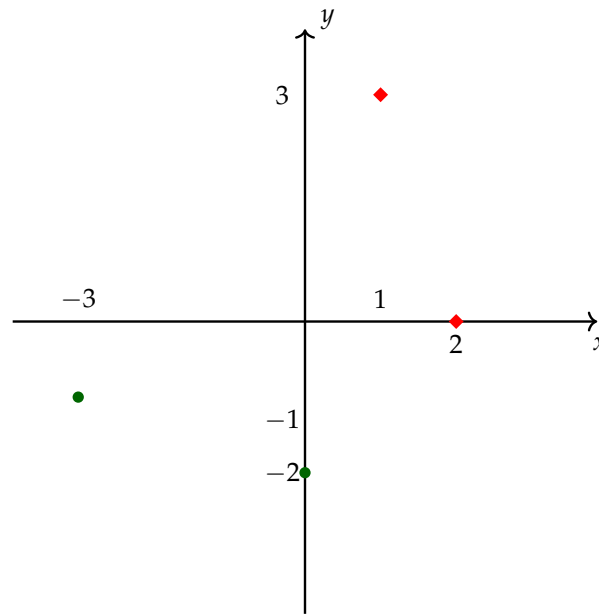


Figure 1: Plot of Transactions in 2-D

Thus, we have a total of 4 transactions, $\begin{bmatrix} -3 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ -2 \end{bmatrix}$ are normal, while $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ are fraudulent.

(a) (3 pts.) Suppose we now construct a data matrix, where the *data points are columns*.

$$X = \begin{bmatrix} -3 & 0 & 2 & 1 \\ -1 & -2 & 0 & 3 \end{bmatrix} \tag{27}$$

Using this data matrix, **calculate its first principal component \vec{u}_1** .

(HINT:

i. You may also make use of the fact that XX^T is given by:

$$XX^T = \begin{bmatrix} -3 & 0 & 2 & 1 \\ -1 & -2 & 0 & 3 \end{bmatrix} \begin{bmatrix} -3 & 0 & 2 & 1 \\ -1 & -2 & 0 & 3 \end{bmatrix}^T \tag{28}$$

$$= \begin{bmatrix} 14 & 6 \\ 6 & 14 \end{bmatrix} \tag{29}$$

ii. You may also make use of the characteristic polynomial of XX^T :

$$\lambda^2 - 28\lambda + 160 = (\lambda - 20)(\lambda - 8) = 0 \tag{30}$$

)

(HINT: Remember that your principal component should be of unit norm.)

Solution:

Our eigenvalues are $\lambda_1 = 20$ and $\lambda_2 = 8$. Thus, the eigenvector corresponding to our largest eigenvalue $\lambda_1 = 20$ is $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ from inspection.

Thus, our principal component vector is given as the normalized eigenvector: i.e. $\vec{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$.

(b) (3 pts.) It's difficult to come up with a useful classifier in two dimensions. Let's use PCA dimensionality reduction to help.

Using your answer in part (a), **project your two-dimensional data points onto one dimension. Express your answer as the vector $\vec{z} \in \mathbb{R}^{1 \times 4}$.**

Solution: To project our 2-d data into 1-d, we project our data matrix onto the first principal component as follows:

$$\vec{z} = \vec{u}_1^T X = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -3 & 0 & 2 & 1 \\ -1 & -2 & 0 & 3 \end{bmatrix} = \begin{bmatrix} -\frac{4}{\sqrt{2}} & -\frac{2}{\sqrt{2}} & \frac{2}{\sqrt{2}} & \frac{4}{\sqrt{2}} \end{bmatrix} \tag{31}$$

(c) (2 pts.) Now, **plot each of these points, on the line below. Indicate the value and label (normal as circle and fraudulent as diamond) for each point.**

Note: your plot doesn't have to be to scale.

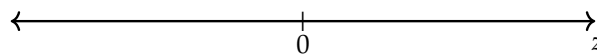


Figure 2: Plot of Transactions in 1-D

Solution:

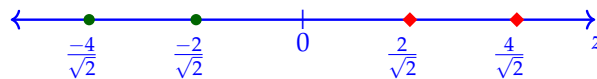


Figure 3: Plot of Transactions in 1-D

(d) (1 pts.) Suppose you are given some transaction datapoint, \vec{x}_i and you project it to be one-dimensional, i.e. z_i . **Based on your plot from part (c), come up with an inequality in terms of z_i to identify if that transaction is fraudulent.**

Note: There can be more than one answer, but please only give one.

Solution: $\vec{z}_i > 0$ is a possible answer. In fact any answer in between $\vec{z}_i \geq -\frac{2}{\sqrt{2}}$ and $\vec{z}_i \geq \frac{2}{\sqrt{2}}$ is valid.

- (e) (3 pts.) It can often be informative to compare our original data with its PCA reconstructions. **Using PCA, reconstruct \tilde{z} back into 2-D as the matrix $\tilde{X} \in \mathbb{R}^{2 \times 4}$.**

Solution: Using the reconstruction principle of PCA, we lift the 1D data back into 2 dimensions by doing the following:

$$\tilde{X} = \tilde{u}_1 \tilde{u}_1^\top X = \begin{bmatrix} -2 & -1 & 1 & 2 \\ -2 & -1 & 1 & 2 \end{bmatrix} \quad (32)$$

- (f) (3 pts.) Finally, we visualize our PCA reconstruction. **On the graph below, draw your first principal component direction (extend it as a solid line in both directions). Draw your reconstructed points from \tilde{X} , with dotted lines connecting them with the corresponding point in X .** You may draw on the plot on the next page.

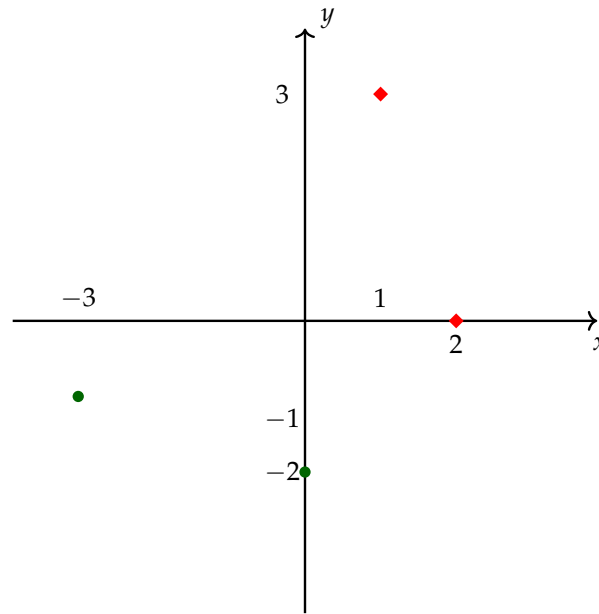


Figure 4: Plot of Transactions in 2-D

Solution:

Using the answer from part (d), students should end up with the following graph:

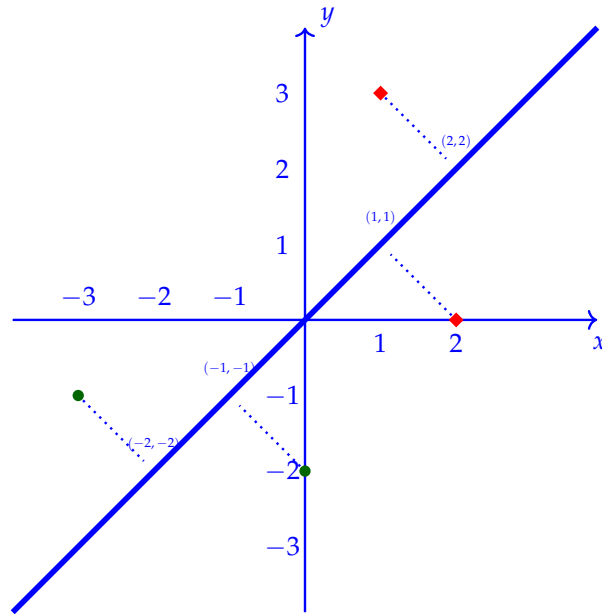
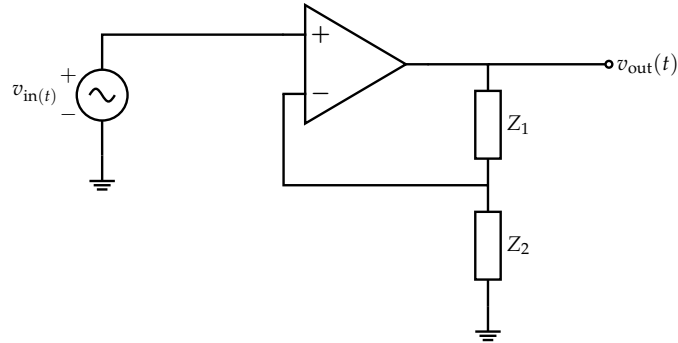


Figure 5: Plot of Transactions in 2-D

7. Op-Amp Transfer Function (25 pts.)

In this problem, we will analyze the transfer function of a non-inverting amplifier implemented with an op-amp. For the entire problem, you can assume that all of the op-amp Golden Rules apply.

- (a) (7 pts.) Consider the following op-amp non-inverting amplifier, with generic impedances Z_1 and Z_2 present on the feedback loop:



Show that the transfer function for the circuit above is $H(j\omega) = \frac{\tilde{V}_{out}}{\tilde{V}_{in}} = 1 + \frac{Z_1}{Z_2}$.

Solution: Since no current flows through the op-amp terminals, the voltage at the negative terminal of the op-amp is the output of a voltage divider of \tilde{V}_{out} between Z_1 and Z_2 :

$$\tilde{V}^- = \frac{Z_2}{Z_1 + Z_2} \tilde{V}_{out} \tag{33}$$

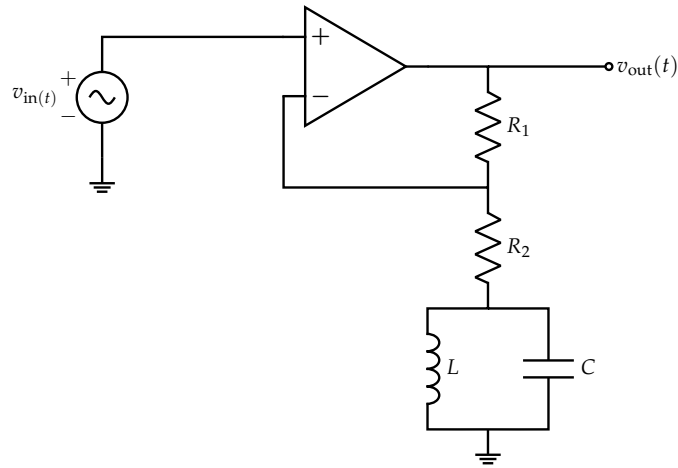
The op-amp is assumed to be in negative feedback so $\tilde{V}^+ = \tilde{V}^-$.

$$\tilde{V}^- = \tilde{V}^+ \tag{34}$$

$$\frac{Z_2}{Z_1 + Z_2} \tilde{V}_{out} = \tilde{V}_{in} \tag{35}$$

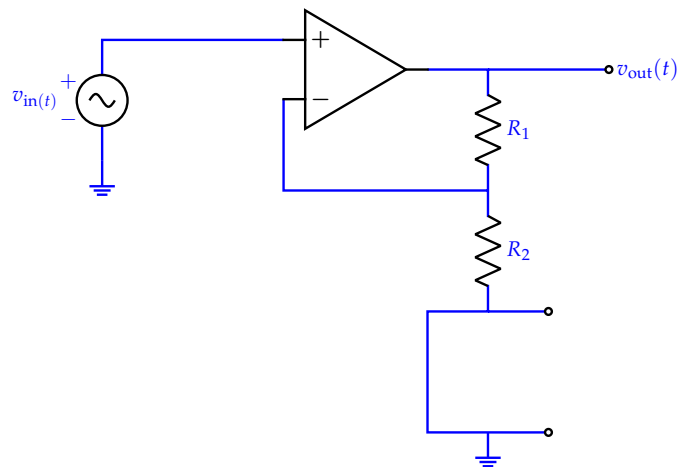
$$H(j\omega) = \frac{\tilde{V}_{out}}{\tilde{V}_{in}} = \frac{Z_2 + Z_1}{Z_2} = 1 + \frac{Z_1}{Z_2} \tag{36}$$

For the remaining parts, we will consider the following non-inverting amplifier, which is a specific instance of the amplifier from part (a).

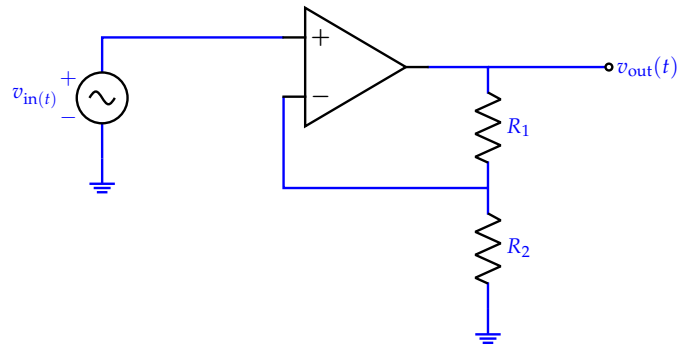


- (b) (4 pts.) **Draw an equivalent circuit if the input frequency is $\omega = 0$.** Then, using the equation for $H(j\omega)$ from part (a), **find a simplified expression for $H(j0)$** in terms of R_1 , R_2 , L , and C .
 (HINT: To draw the equivalent circuit, think about the impedance of inductors and capacitors at $\omega = 0$ (DC).)

Solution: At $\omega = 0$, the impedances of the inductor and capacitor are $Z_L = 0$ and $Z_C = \infty$, which means the inductor behaves as a short circuit and the capacitor behaves as an open circuit (you can also reach this conclusion by thinking about DC steady state since this is what $\omega = 0$ corresponds to). With this, the equivalent circuit is:



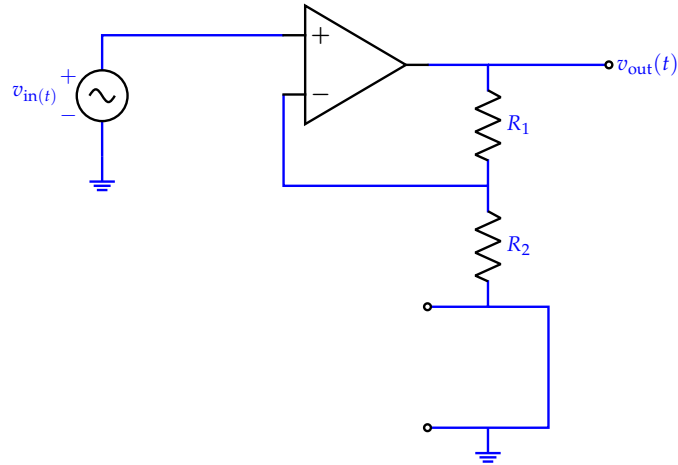
If we remove unnecessary information (not needed to have a valid equivalent circuit, though it helps to see what is going on), the simplified circuit diagram would be:



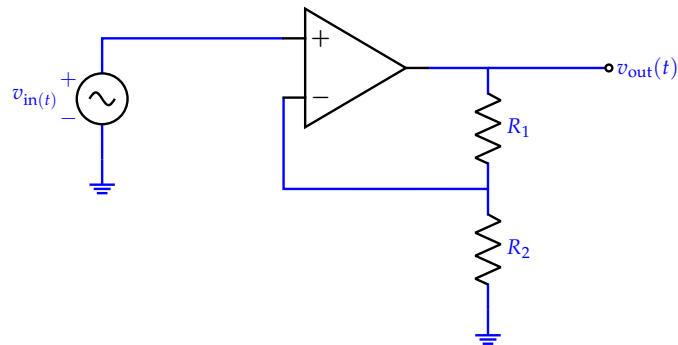
Thus, comparing with the circuit from part (a), $Z_1 = R_1$ and $Z_2 = R_2$, so $H(j0) = 1 + \frac{R_1}{R_2}$. This is the standard non-inverting op-amp amplifier!

- (c) (4 pts.) **Draw an equivalent circuit if the input frequency is $\omega = \infty$.** Then, the equation for $H(j\omega)$ from part (a), **find a simplified expression for $H(j\infty)$** in terms of R_1 , R_2 , L , and C .

Solution: At $\omega = \infty$, the impedances of the inductor and capacitor are $Z_L = \infty$ and $Z_C = 0$, which means the inductor behaves as an open circuit and the capacitor behaves as a short circuit. With this, the equivalent circuit is:



If we remove unnecessary information again, the simplified circuit diagram would be (notice that this is the exact same as the previous part):



Just as with the previous part, $Z_1 = R_1$ and $Z_2 = R_2$, so $H(j\infty) = 1 + \frac{R_1}{R_2}$.

- (d) (2 pts.) **Find an expression for ω_r , the resonant frequency**, in terms of R_1 , R_2 , L , and C . For this circuit, the resonant frequency is defined as the frequency at which the equivalent impedance of the inductor and capacitor is $Z_{eq} = \infty$.

Solution: The equivalent impedance of the inductor and capacitor in parallel is:

$$Z_{eq} = Z_L || Z_C \tag{37}$$

$$= j\omega L || \frac{1}{j\omega C} \tag{38}$$

$$= \frac{(j\omega L)(\frac{1}{j\omega C})}{j\omega L + \frac{1}{j\omega C}} \tag{39}$$

$$= \frac{j\omega L}{1 + j^2\omega^2 LC} \tag{40}$$

$$= \frac{j\omega L}{1 - \omega^2 LC} \quad (41)$$

From our obtained expression, we can see that $Z_{\text{eq}} = \infty$ when $1 - \omega^2 LC = 0$, so the equation to solve for the resonant frequency in this case is:

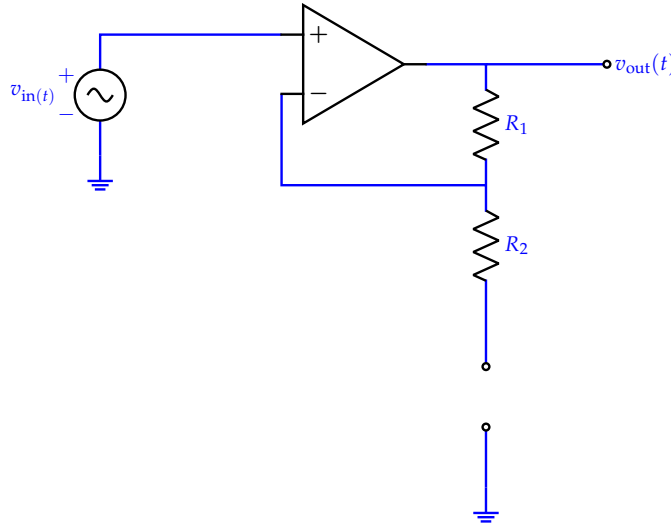
$$1 - \omega_r^2 LC = 0 \quad (42)$$

$$\omega_r^2 = \frac{1}{LC} \quad (43)$$

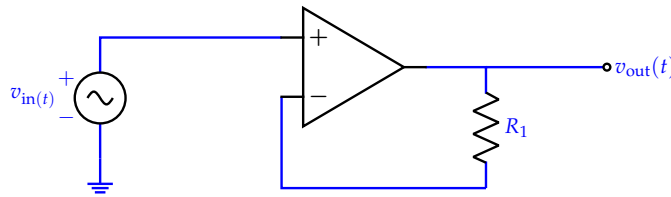
$$\omega_r = \frac{1}{\sqrt{LC}} \quad (44)$$

(e) (2 pts.) Now that we have found the resonant frequency, let's analyze the circuit for $\omega = \omega_r$. **Draw an equivalent circuit if the input frequency is $\omega = \omega_r$.** Then, using the equation for $H(j\omega)$ from part (a), **find a simplified expression for $H(j\omega_r)$** in terms of R_1 , R_2 , L , and C .

Solution: At $\omega = \omega_r$, by definition, the parallel impedance of the inductor and capacitor is $Z_{eq} = \infty$, which means the inductor and capacitor in parallel behaves as an open circuit. With this, the equivalent circuit is:

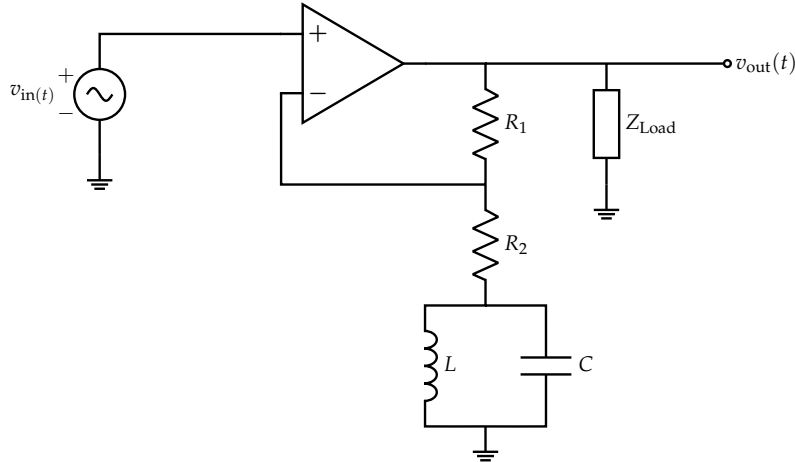


If we remove unnecessary information again, the simplified circuit diagram would be:



Thus, in this case, $Z_1 = R_1$ and $Z_2 = \infty$, so $H(j\omega_r) = 1$.

(f) (6 pts.) Suppose we have a load impedance $Z_{\text{Load}} = |Z_{\text{Load}}|e^{j\frac{\pi}{4}}$ at the output of this circuit:



Suppose the input phasor is \tilde{V}_{in} , and the input is at a cutoff frequency ω_c . **Find the average power P_{avg} delivered to the load impedance Z_{Load}** in terms of $|\tilde{V}_{\text{in}}|$, $|H(j\omega)|_{\text{max}}$ (the maximum magnitude of the transfer function), and the provided element variables.

(HINT: You DO NOT need to solve for $H(j\omega)$. Recall that $|H(j\omega_c)| = \frac{|H(j\omega)|_{\text{max}}}{\sqrt{2}}$.)

Solution: There are many different equivalent equations we can use to solve for the average power P_{avg} . Here, we will use

$$P_{\text{avg}} = \frac{1}{2} |\tilde{V}| |\tilde{I}| \cos(\phi) \quad (45)$$

where \tilde{V} and \tilde{I} are the phasors for the voltage and current through the load impedance Z_{Load} , respectively, and ϕ is the phase difference between \tilde{V} and \tilde{I} ($\phi = \angle \tilde{V} - \angle \tilde{I}$).

We know that the voltage phasor across Z_{Load} will be $\tilde{V} = \tilde{V}_{\text{out}}$, which can be found using the transfer function at $\omega = \omega_c$, the input frequency:

$$\tilde{V}_{\text{out}} = H(j\omega_c) \tilde{V}_{\text{in}} \quad (46)$$

$$|\tilde{V}_{\text{out}}| = |H(j\omega_c)| |\tilde{V}_{\text{in}}| \quad (47)$$

The current phasor \tilde{I} is found using Ohm's law:

$$\tilde{I} = \frac{\tilde{V}_{\text{out}}}{Z_{\text{Load}}} \quad (48)$$

$$|\tilde{I}| = \frac{|\tilde{V}_{\text{out}}|}{|Z_{\text{Load}}|} = |H(j\omega_c)| \frac{|\tilde{V}_{\text{in}}|}{|Z_{\text{Load}}|} \quad (49)$$

Since $\tilde{I} = \frac{\tilde{V}_{\text{out}}}{Z_{\text{Load}}}$, $\angle \tilde{I} = \angle \tilde{V}_{\text{out}} - \angle Z_{\text{Load}}$.

Thus, $\phi = \angle \tilde{V}_{\text{out}} - \angle \tilde{I} = \angle Z_{\text{Load}} = \frac{\pi}{4}$.

Finally, using the hint $|H(j\omega_c)| = \frac{|H(j\omega)|_{\text{max}}}{\sqrt{2}}$, we can solve for P_{avg} :

$$P_{\text{avg}} = \frac{1}{2} |\tilde{V}| |\tilde{I}| \cos(\phi) \quad (50)$$

$$= \frac{1}{2} (|H(j\omega_c)| |\tilde{V}_{\text{in}}|) \left(|H(j\omega_c)| \frac{|\tilde{V}_{\text{in}}|}{|Z_{\text{Load}}|} \right) \cos\left(\frac{\pi}{4}\right) \quad (51)$$

$$= \frac{1}{2} |H(j\omega_c)|^2 \frac{|\tilde{V}_{in}|^2}{|Z_{Load}|} \left(\frac{1}{\sqrt{2}} \right) \quad (52)$$

$$= \frac{1}{2\sqrt{2}} \left(\frac{1}{2} |H(j\omega)|^2_{max} \right) \frac{|\tilde{V}_{in}|^2}{|Z_{Load}|} \quad (53)$$

$$= \frac{1}{4\sqrt{2}} |H(j\omega)|^2_{max} \frac{|\tilde{V}_{in}|^2}{|Z_{Load}|} \quad (54)$$

For reference, $|H(j\omega)|_{max} = 1 + \frac{R_1}{R_2}$ for this circuit. This is because

$$H(j\omega) = 1 + \frac{Z_1}{Z_2} = 1 + \frac{R_1}{R_2 + Z_{eq}} = 1 + \frac{R_1}{R_2 + \frac{j\omega L}{1 - \omega^2 LC}} \quad (55)$$

which is maximized when $Z_{eq} = \frac{j\omega L}{1 - \omega^2 LC} = 0$ (which corresponds to the $\omega = 0$ and $\omega = \infty$ cases discussed in earlier parts of the problem).

8. Controllability (10 pts.)

(a) (3 pts.) **Is the following, discrete-time system controllable? Explain your answer.**

$$\vec{x}[i + 1] = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \vec{x}[i] + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u[i] \tag{56}$$

Solution: We compute the following quantities:

$$\vec{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{57}$$

$$A\vec{b} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \tag{58}$$

which means we construct the controllability matrix as follows:

$$C = [A\vec{b} \quad \vec{b}] = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \tag{59}$$

To check whether this matrix is full rank, we can compute the determinant to obtain $\det\{C\} = 2 - 1 = 1 \neq 0$ which shows that it is full rank. Hence, the system is controllable.

(b) (3 pts.) **Is the following, discrete-time system controllable? Explain your answer.**

$$\vec{x}[i + 1] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \vec{x}[i] + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \vec{u}[i] \tag{60}$$

Solution: The system shown is in controllable canonical form. From lecture and Homework 9 Question 4, we know that the system is controllable. One may also compute the controllability matrix mechanically, which would yield

$$C = [A^3\vec{b} \quad A^2\vec{b} \quad A\vec{b} \quad \vec{b}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 \\ -4 & 2 & -1 & 1 \end{bmatrix} \tag{61}$$

Based on the way the matrix is structured (each column sets a new element to 0, going from left to right), we can see that it is full rank. We can also notice that the controllability matrix will be full rank without computing all of the elements. Based on the structure of A , multiplying \vec{b} by increasing powers of A will repeatedly add a nonzero element in a bottom-top fashion, which indicates that the resulting controllability matrix will be lower triangular. In this case, since the last element of \vec{b} is 1, the elements along the diagonal of this lower triangular matrix will be 1, which implies that the controllability matrix will be full rank.

(c) (4 pts.) **Prove or disprove: if the matrix A in the discrete-time system $\vec{x}[i + 1] = A\vec{x}[i] + \vec{b}u[i]$ is full rank, then the system is controllable.** Assume $\vec{b} \neq \vec{0}$. If you believe the statement is

TRUE, then prove it by showing that the controllability matrix is always full rank for any full-rank matrix $A \in \mathbb{R}^{n \times n}$. If you believe the statement is FALSE, give a counterexample with a specific matrix A and \vec{b} that produces a rank deficient controllability matrix.

Solution: This statement is false. We can take the following discrete-time system as an example:

$$\vec{x}[i+1] = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_A \vec{x}[i] + \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{\vec{b}} u[i] \quad (62)$$

where $A = I$ is clearly full rank. However, the controllability matrix is

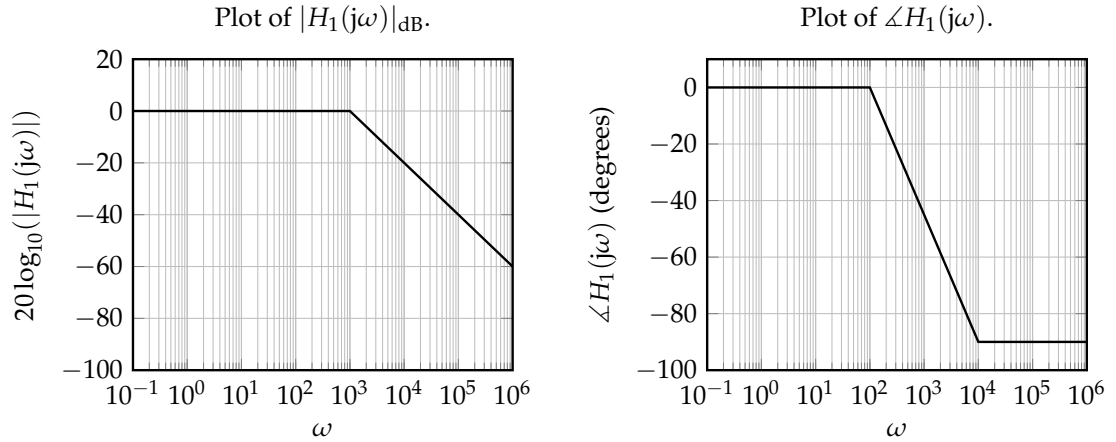
$$\mathcal{C} = \begin{bmatrix} A^2 \vec{b} & A \vec{b} & \vec{b} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad (63)$$

which has rank 1 (the first column is equal to the second and third columns). Hence, the system is not controllable.

9. Bode Plot Analysis (20 pts.)

In this problem, we will analyze various Bode plots and implement them.

- (a) (2 pts.) Here are the Bode magnitude and phase plots (straight line approximations) for transfer function $H_1(j\omega)$:



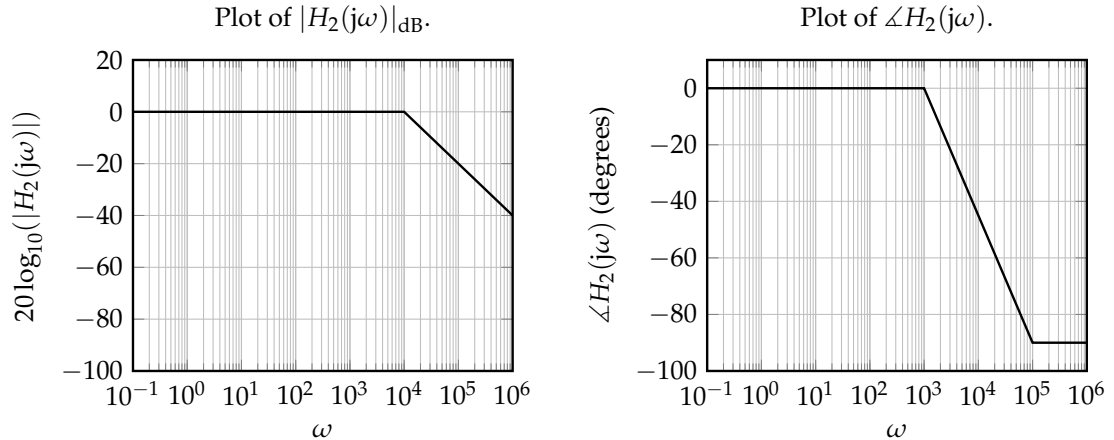
Find an expression for $H_1(j\omega)$. What type of filter is H_1 ?

Solution: We can notice that the Bode plots have the shape of a standard low pass filter, with cutoff frequency $\omega_c = 10^3$.

Thus, the transfer function should be

$$H_1(j\omega) = \frac{1}{1 + j\frac{\omega}{10^3}} \tag{64}$$

- (b) (2 pts.) Here are the Bode magnitude and phase plots (straight line approximations) for transfer function $H_2(j\omega)$:



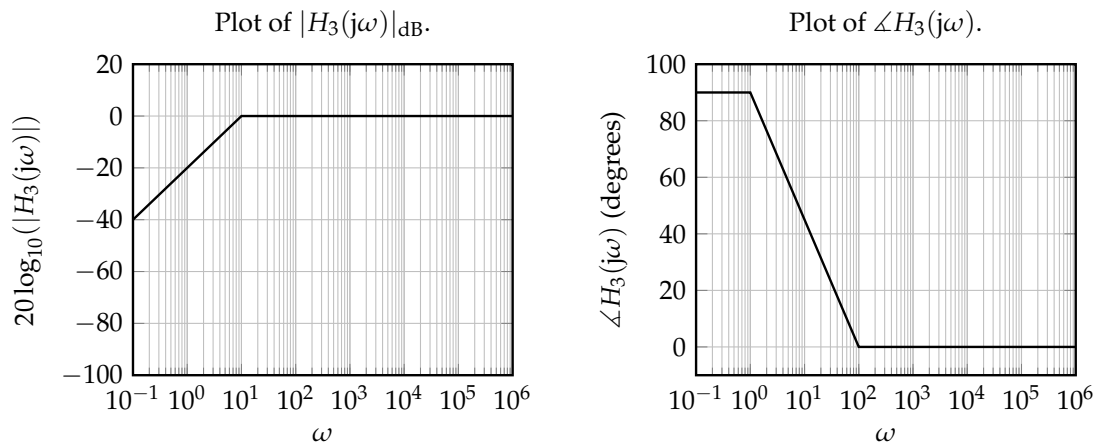
Find an expression for $H_2(j\omega)$. What type of filter is H_2 ?

Solution: We can notice that the Bode plots have the shape of a standard low pass filter, with cutoff frequency $\omega_c = 10^4$.

Thus, the transfer function should be

$$H_2(j\omega) = \frac{1}{1 + j\frac{\omega}{10^4}} \quad (65)$$

- (c) (2 pts.) Here are the Bode magnitude and phase plots (straight line approximations) for transfer function $H_3(j\omega)$:



Find an expression for $H_3(j\omega)$. What type of filter is H_3 ?

Solution: We can notice that the Bode plots have the shape of a standard high pass filter, with cutoff frequency $\omega_c = 10^1$.

Thus, the transfer function should be

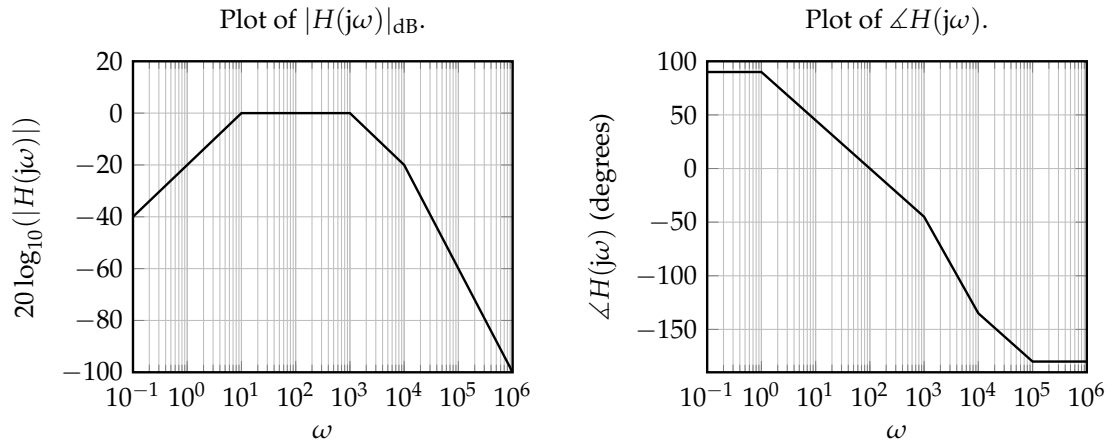
$$H_3(j\omega) = \frac{j\frac{\omega}{10^1}}{1 + j\frac{\omega}{10^1}} \quad (66)$$

PRINT your name and student ID: _____

EECS 16B Final

2023-11-21 14:00:56-08:00

- (d) (4 pts.) Here are the Bode magnitude and phase plots (straight line approximations) for transfer function $H(j\omega)$:



Find an expression for $H(j\omega)$. What type of filter is H ?

(HINT: Consider the plots from the previous parts.)

Solution: As the hint suggests, the Bode plot in this part is actually the combination of the Bode plots from the previous parts. To be specific, we can observe the following details:

$$20 \log(H(j\omega)) = 20 \log(H_1(j\omega)) + 20 \log(H_2(j\omega)) + 20 \log(H_3(j\omega)) \quad (67)$$

$$\angle H(j\omega) = \angle H_1(j\omega) + \angle H_2(j\omega) + \angle H_3(j\omega) \quad (68)$$

From this, we can determine that

$$H(j\omega) = H_1(j\omega)H_2(j\omega)H_3(j\omega) = \frac{j\frac{\omega}{10^1}}{(1 + j\frac{\omega}{10^1})(1 + j\frac{\omega}{10^3})(1 + j\frac{\omega}{10^4})} \quad (69)$$

As the Bode plot indicates, this is a band pass filter.

- (e) (5 pts.) Suppose we have the following input into the filter with transfer function $H(j\omega)$:

$$v_{\text{in}}(t) = 1 + \cos(10^2 t) + \sin(10^4 t) + \cos(10^{10} t) \quad (70)$$

Find an expression for $v_{\text{out}}(t)$, the output of the filter for this input. You can approximate any values you use as part of this calculation.

(HINT: You can use the plot of $H(j\omega)$ from the previous part.)

Solution: To deal with an input that is the sum of different sinusoidal inputs, we can use superposition and apply the filter to each sinusoidal input independently. To do this, we need to evaluate the transfer function $H(j\omega)$ for each of the relevant frequencies, which we can do using the provided plot of $H(j\omega)$ from the previous part.

$$|H(j0)| = 0 \quad (71)$$

$$|H(j10^2)| \approx 0\text{dB} = 1 \quad \angle H(j10^2) \approx 0 \quad (72)$$

$$|H(j10^4)| \approx -20\text{dB} = \frac{1}{10} \quad \angle H(j10^4) \approx -135^\circ \quad (73)$$

$$|H(j10^{10})| \approx 0 \tag{74}$$

Thus, applying the transfer function to the given input will produce approximately the following output:

$$v_{\text{out}}(t) = |H(j0)|(1) + |H(j10^2)| \cos(10^2 t + \angle H(j10^2)) + |H(j10^4)| \sin(10^4 t + \angle H(j10^4)) + |H(j10^{10})| \cos(10^{10} t + \angle H(j10^{10})) \tag{75}$$

$$v_{\text{out}}(t) \approx \cos(10^2 t) + \frac{1}{10} \sin(10^4 t - 135^\circ) \tag{76}$$

- (f) (5 pts.) Now, suppose we have the following transfer function, independent of the previous parts of this problem:

$$F(j\omega) = \frac{(j\frac{\omega}{10^3})(j\frac{\omega}{10^5})}{(1 + j\frac{\omega}{10^3})(1 + j\frac{\omega}{10^5})(1 + j\frac{\omega}{10^6})} \tag{77}$$

Now, suppose we want to build a filter that implements $F(j\omega)$. You are provided with 3 1 kΩ resistors, 2 ideal op-amps, and 3 capacitors whose values you can choose as necessary. **Draw a filter that implements $F(j\omega)$. Make sure to label the values of any capacitors you use.**

(HINT: Try to break up the filter $F(j\omega)$ into multiple first order low pass / high pass filters.)

Solution: We can construct the overall filter $F(j\omega)$ by cascading the three standard filters it is a product of ($F(j\omega) = F_1(j\omega)F_2(j\omega)F_3(j\omega)$).

$$F_1(j\omega) = \frac{1}{1 + j\frac{\omega}{10^6}} \tag{78}$$

$$F_2(j\omega) = \frac{j\frac{\omega}{10^5}}{1 + j\frac{\omega}{10^5}} \tag{79}$$

$$F_3(j\omega) = \frac{j\frac{\omega}{10^3}}{1 + j\frac{\omega}{10^3}} \tag{80}$$

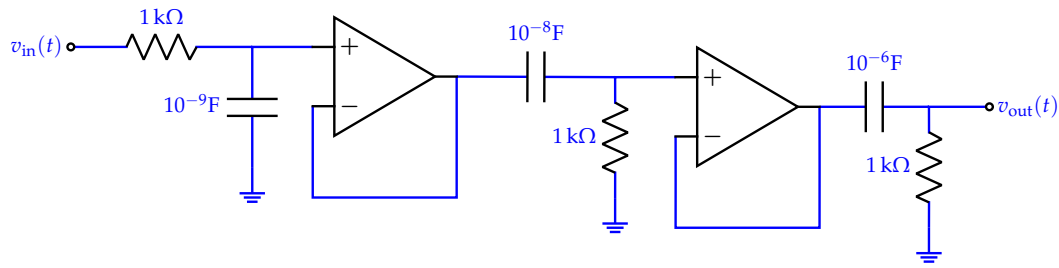
The low pass filter (F_1) can be constructed using an RC filter, and the high pass filters (F_2 and F_3) can be constructed using CR filters. Since the cutoff frequency for both RC and CR filters is $\omega_c = \frac{1}{RC}$, we can find the required capacitance value using $C = \frac{1}{R\omega_c}$:

$$C_1 = \frac{1}{(10^3)(10^6)} = 10^{-9}\text{F} \tag{81}$$

$$C_2 = \frac{1}{(10^3)(10^5)} = 10^{-8}\text{F} \tag{82}$$

$$C_3 = \frac{1}{(10^3)(10^3)} = 10^{-6}\text{F} \tag{83}$$

Thus, the overall circuit will be the three filters with op-amp buffers in between to prevent loading (which allows the transfer functions to be multiplied together):



10. Nonlinear System ID and Feedback (20 pts.)

Suppose we know that a scalar, discrete-time system follows some nonlinear dynamics equation of the form

$$x[i + 1] = f(x[i], u[i]) \tag{84}$$

for some nonlinear function $f(\cdot, \cdot)$ that is unknown. In this problem, we will make guesses about the form of f and use system ID to estimate the exact function from data.

(a) (3 pts.) Your friend Vik believes that f has the following form:

$$f(x[i], u[i]) = a_3(x[i])^3 + a_2(x[i])^2 + a_1x[i] + a_0 + u[i] \tag{85}$$

He doesn't know the values of a_3, a_2, a_1, a_0 , but he wants to estimate them using system ID. He provides a sequence of control inputs $u[0], u[1], \dots, u[n - 1]$ to the system, and he observes $x[0], x[1], x[2], \dots, x[n]$. **Set up a system of linear equations of the form $D\vec{p} = \vec{s}$ to estimate**

$$\vec{p} = \begin{bmatrix} a_3 \\ a_2 \\ a_1 \\ a_0 \end{bmatrix}.$$

(HINT: Isolate all the known quantities to one side of the equation, and the unknowns to the other.)

Solution: From the provided form of f , we can write

$$x[i + 1] = a_3(x[i])^3 + a_2(x[i])^2 + a_1x[i] + a_0 + u[i] \tag{86}$$

$$x[i + 1] - u[i] = a_3(x[i])^3 + a_2(x[i])^2 + a_1x[i] + a_0 \tag{87}$$

We can arrange this into a system ID problem of the following form:

$$\underbrace{\begin{bmatrix} (x[0])^3 & (x[0])^2 & x[0] & 1 \\ (x[1])^3 & (x[1])^2 & x[1] & 1 \\ \vdots & \vdots & \vdots & \vdots \\ (x[n-1])^3 & (x[n-1])^2 & x[n-1] & 1 \end{bmatrix}}_D \underbrace{\begin{bmatrix} a_3 \\ a_2 \\ a_1 \\ a_0 \end{bmatrix}}_{\vec{p}} = \underbrace{\begin{bmatrix} x[1] - u[0] \\ x[2] - u[1] \\ \vdots \\ x[n] - u[n-1] \end{bmatrix}}_{\vec{s}} \tag{88}$$

(b) (3 pts.) Your other friend Prakash believes that f has the following form:

$$f(x[i], u[i]) = e^{a_1 x[i] + a_0} \cdot u[i] \tag{89}$$

He doesn't know the values of a_1, a_0 , but he wants to estimate them using system ID. He provides a sequence of positive-valued control inputs $u[0], u[1], \dots, u[n-1]$ to the system and he observes $x[0], x[1], x[2], \dots, x[n]$. **Set up a system of linear equations of the form $D\vec{p} = \vec{s}$ to estimate $\vec{p} = \begin{bmatrix} a_1 \\ a_0 \end{bmatrix}$.**

(HINT: Consider writing $\ln(x[i+1]) = \ln(f(x[i], u[i]))$.)

Solution: From the provided form of f , we can write

$$x[i+1] = e^{a_1 x[i] + a_0} \cdot u[i] \tag{90}$$

$$\ln(x[i+1]) = a_1 x[i] + a_0 + \ln(u[i]) \tag{91}$$

$$\ln(x[i+1]) - \ln(u[i]) = a_1 x[i] + a_0 \tag{92}$$

We can arrange this into a system ID problem of the following form:

$$\underbrace{\begin{bmatrix} x[0] & 1 \\ x[1] & 1 \\ \vdots & \vdots \\ x[n-1] & 1 \end{bmatrix}}_D \underbrace{\begin{bmatrix} a_1 \\ a_0 \end{bmatrix}}_{\vec{p}} = \underbrace{\begin{bmatrix} \ln(x[1]) - \ln(u[0]) \\ \ln(x[2]) - \ln(u[1]) \\ \vdots \\ \ln(x[n]) - \ln(u[n-1]) \end{bmatrix}}_{\vec{s}} \tag{93}$$

- (c) (6 pts.) After all their testing, Vik and Prakash conclude that Prakash's guess is correct. In fact, the system follows the dynamics equation given by

$$x[i + 1] = e^{-x[i]} \cdot u[i] \tag{94}$$

This system is not BIBO stable.

- i. **Construct a sequence of non-zero control inputs such that $x[i]$ remains bounded. Explain what value $x[i]$ is bounded by (i.e., find a α such that $|x[i]| \leq \alpha$).**
- ii. **Construct a sequence of non-zero control inputs such that $x[i]$ becomes unbounded. Explain your answer intuitively.**

Assume that $x[0] = 0$.

Solution: A sequence of non-zero control inputs that keep $x[i]$ bounded is $1, 1, 1, \dots$. Notice that $x[i]$ will be always non-negative in this case. Hence, $0 \leq e^{-x[i]} \leq 1$. As a result, $|x[i + 1]| = |e^{-x[i]} \cdot u[i]| = e^{-x[i]} \cdot |u[i]| \leq 1 \cdot 1 = 1$. Thus, $\alpha = 1$.

A sequence of non-zero control inputs that makes $x[i]$ unbounded is $-1, -1, -1, \dots$. Notice that $e^{-x[i]}$ is always non-negative, so $x[i + 1] = e^{-x[i]}u[i] < 0$. Hence, $|e^{-x[i]}| \geq 1$, and the state will grow unbounded.

- (d) (8 pts.) Consider again the system in the previous part and the feedback law $u[i] = e^{kx[i]}$ where k is some constant. In this part, we will prove that the system will only be stable if and only if $k \leq 1$. To do this, accomplish the following:

- i. **Assume $k > 1$. Show that there exists an N such that, for all $i \geq N$, $x[i] < x[i + 1]$.**
 (HINT: It is the case that there exists an N such that, for all $i \geq N$, $\ln(x[i]) < \alpha x[i]$ for $\alpha > 0$.)
- ii. **Assume $k \leq 1$. Find a constant ε such that $|x[i]| \leq \varepsilon$ for every $i \geq 0$.**
 (HINT: Remember that $e^{\text{anything non-positive}}$ is bounded between 0 and 1.)

Again, assume $x[0] = 0$.

Solution: For the first part, we can write $\alpha = k - 1 > 0$. This means that

$$x[i + 1] = e^{\alpha x[i]} \tag{95}$$

$$\ln(x[i + 1]) = \alpha x[i] \tag{96}$$

From the hint, there exists an N such that for all $i \geq N$, $\alpha x[i] > \ln(x[i]) \implies \ln(x[i + 1]) > \ln(x[i])$. Raising both sides to the power e , we have $x[i + 1] > x[i]$.

For the second part, we can again write $\alpha = k - 1 \leq 0$. We have that

$$|x[i + 1]| = |e^{\alpha x[i]}| = e^{\alpha x[i]} \tag{97}$$

Since $x[i]$ will always be non-negative and α is non-positive, $0 \leq e^{\alpha x[i]} \leq 1$. Thus, $|x[i]| \leq 1$ for all values of i .

11. Motivating Ridge Regression From Least Squares (30 pts.)

Suppose we have the linear system:

$$A\vec{x} = \vec{y} \tag{98}$$

As we have seen in this class, Least Squares (LS) is a great tool for performing System ID to determine \vec{x} in such a system. However, there are a few problems and limitations with the method.

- 1 When the singular values of A are small, our answers can be "unstable" (we'll see what this means soon).
- 2 A may not always be full column rank.

In this problem, we will introduce and motivate **Ridge Regression**, which will hopefully improve upon Least Squares by solving the above issues. For the sake of simplicity, assume $A \in \mathbb{R}^{n \times n}$, $\vec{x} \in \mathbb{R}^{n \times 1}$, and $\vec{y} \in \mathbb{R}^{n \times 1}$.

- (a) (6 pts.) Suppose the compact SVD of A is given by $A = U_r \Sigma_r V_r^T$. **Rewrite the least squares solution \vec{x}_{LS}^* in its simplest form, replacing A with its SVD.**

Solution: Plugging in the SVD to the least-square solution should yield the pseudoinverse as we saw in discussion:

$$\vec{x}_{LS}^* = (A^T A)^{-1} A^T \vec{y} \tag{99}$$

$$= ((V_r \Sigma_r U_r^T)(U_r \Sigma_r V_r^T))^{-1} (V_r \Sigma_r U_r^T) \vec{y} \tag{100}$$

$$= (V_r \Sigma_r^2 V_r^T)^{-1} V_r \Sigma_r U_r^T \vec{y} \tag{101}$$

$$= V_r \Sigma_r^{-2} V_r^T V_r \Sigma_r U_r^T \vec{y} \tag{102}$$

$$= V_r \Sigma_r^{-1} U_r^T \vec{y} \tag{103}$$

Note the

- (b) (4 pts.) In order to see how small singular values can make \vec{x}_{LS}^* unstable, let us simplify and consider the scalar case. Let us suppose our SVD is now completely scalar (i.e. $U_r = u$, $\Sigma_r = \sigma$, and $V_r^T = v$).

After rewriting x_{LS}^* in the scalar case, what does $\|x_{LS}^*\|$ approach as $\sigma \rightarrow 0$?

Solution: The scalar version of the Least Squares solution is given by:

$$x_{LS}^* = \frac{uv}{\sigma} y \tag{104}$$

Thus as $\sigma \rightarrow 0$, $\|x_{LS}^*\| \rightarrow \infty$ which is quite an unstable solution!

Oh no! It seems that our solution \vec{x}_{LS}^* can really get unstable when the singular values are small. Let us take a look at Ridge Regression, and see if it is any better.

One way to think of Ridge Regression is to see it as exactly like Least Squares, except with a slight modification to A and \vec{y} . Thus, we formulate Ridge Regression as:

$$A_{\text{ridge}} \vec{x} = \vec{y}_{\text{ridge}} \tag{105}$$

where $A_{\text{ridge}} = \begin{bmatrix} A \\ I_{n \times n} \end{bmatrix} \in \mathbb{R}^{2n \times n}$, $\vec{x} \in \mathbb{R}^{n \times 1}$, and $\vec{y}_{\text{ridge}} = \begin{bmatrix} \vec{y} \\ \vec{0}_n \end{bmatrix} \in \mathbb{R}^{2n \times 1}$.

Well, for starters, we've already solved Problem 2. With one small change, we've ensured that our data matrix, A_{ridge} will always be full column rank. In the following parts, we'll try and see if we've solved Problem 1.

For the rest of the problem, assume that A is full row rank.

- (c) (8 pts.) **Show that the solution to the Ridge Regression problem \vec{x}_{ridge}^* is given by $\vec{x}_{\text{ridge}}^* = (A^T A + I)^{-1} A^T \vec{y}$**

Solution: Because our linear system is still in the same format as before, we can just plug in A_{ridge} and \vec{y}_{ridge} into the Least Squares solution as follows:

$$\vec{x}_{\text{ridge}}^* = (A_{\text{ridge}}^T A_{\text{ridge}})^{-1} A_{\text{ridge}}^T \vec{y}_{\text{ridge}} \tag{106}$$

$$= \left(\left(\begin{bmatrix} A \\ I_{n \times n} \end{bmatrix} \right)^T \begin{bmatrix} A \\ I_{n \times n} \end{bmatrix} \right)^{-1} \begin{bmatrix} A \\ I_{n \times n} \end{bmatrix}^T \begin{bmatrix} \vec{y} \\ \vec{0}_n \end{bmatrix} \tag{107}$$

$$= \left(\begin{bmatrix} A^T & I_{n \times n} \end{bmatrix} \begin{bmatrix} A \\ I_{n \times n} \end{bmatrix} \right)^{-1} \begin{bmatrix} A^T & I_{n \times n} \end{bmatrix} \begin{bmatrix} \vec{y} \\ \vec{0}_n \end{bmatrix} \tag{108}$$

$$= (A^T A + I)^{-1} A^T \vec{y} \tag{109}$$

where the final result comes as a result of block matrix multiplication.

(d) (8 pts.) **Now, just like in part A, show that the Ridge Regression solution can be written as**
 $\vec{x}_{\text{ridge}}^* = V_r(\Sigma_r^2 + I)^{-1}\Sigma_r U_r^\top \vec{y}$ **by replacing A in terms of its SVD $A = U_r \Sigma_r V_r^\top$.**

(HINT: How can you rewrite I in the Ridge Regression solution as a product of two convenient orthonormal matrices?)

(HINT: For arbitrary matrices, $A, B, C,$ and D $ABC + ADC = A(B + D)C$.)

Solution: Once again, we mirror our answer from what we did in part (a). We start with the Ridge Regression solution and continue by plugging in the SVD for A .

$$\vec{x}_{\text{ridge}}^* = (A^\top A + I)^{-1} A^\top \vec{y} \tag{110}$$

$$= (V_r \Sigma_r U_r^\top U_r \Sigma_r V_r^\top + I)^{-1} (V_r \Sigma_r U_r^\top) \vec{y} \tag{111}$$

$$= (V_r \Sigma_r^2 V_r^\top + I)^{-1} (V_r \Sigma_r U_r^\top) \vec{y} \tag{112}$$

We can use the hints and the fact that A is full row rank to rewrite I as $I = V_r V_r^\top$ and do the following manipulation:

$$(V_r \Sigma_r^2 V_r^\top + I)^{-1} (V_r \Sigma_r U_r^\top) \vec{y} = (V_r \Sigma_r^2 V_r^\top + V_r V_r^\top)^{-1} (V_r \Sigma_r U_r^\top) \vec{y} \tag{113}$$

$$= (V_r (\Sigma_r^2 + I) V_r^\top)^{-1} (V_r \Sigma_r U_r^\top) \vec{y} \tag{114}$$

$$= V_r (\Sigma_r^2 + I)^{-1} \Sigma_r U_r^\top \vec{y} \tag{115}$$

(e) (4 pts.) Once again, to understand the effect of A having small singular values, let us simplify and consider the scalar case. Suppose once again that the SVD of A is now completely scalar (i.e. $U = u, \Sigma = \sigma,$ and $V^\top = v$).

After rewriting x_{ridge}^* in the scalar case, what does $\|x_{\text{ridge}}^*\|$ approach as $\sigma \rightarrow 0$?

Solution: The scalar version of the Ridge Regression solution is given by:

$$x_{\text{ridge}}^* = uv \left(\frac{\sigma}{\sigma^2 + 1} \right) y \tag{116}$$

Thus as $\sigma \rightarrow 0, \|x_{\text{ridge}}^*\| \rightarrow 0$ which is much better stability for small singular values of A !

12. Bounding Linearization Error (20 pts.)

For the entire problem, we will consider the nonlinear, continuous-time system

$$\frac{dx}{dt} = f(x(t), u(t)) = \sin(x(t)) + u(t) \tag{117}$$

where $f(x, u) := \sin(x) + u$. We can find an approximate linear system to represent the original nonlinear system, using a process called linearization. We choose a linearization point such that $f(x^*, u^*) = 0$. As a concrete example, let's choose $x^* = 0$ and $u^* = 0$. When we linearize about this point, we obtain the following linear system approximation:

$$\frac{dx}{dt} = \hat{f}(x(t), u(t)) = x(t) + u(t) \tag{118}$$

where $\hat{f}(x, u) = x + u$. In this problem, we will derive a control law that bounds the approximation error.

- (a) (8 pts.) Recall the discretization formula for a one-dimensional system:

$$\frac{dx}{dt} = \lambda x(t) + bu(t) \tag{119}$$

$$x_d[i + 1] = e^{\lambda\Delta} x_d[i] + b \cdot \frac{e^{\lambda\Delta} - 1}{\lambda} u_d[i] \tag{120}$$

where Δ is the discretization sampling period. **Discretize the linearized system in eq. (118) for a given sampling period $\Delta > 0$. Is the discretized system stable?**

Solution: Pattern matching, we have $\lambda = 1$ and $b = 1$. Applying the discretization formula, we have

$$x_d[i + 1] = e^{\Delta} x_d[i] + (e^{\Delta} - 1) u_d[i] \tag{121}$$

Since $\Delta > 0$, $e^{\Delta} > 1$ which implies that the discretized system is unstable.

- (b) (7 pts.) Regardless of your answer to the previous part, suppose we have the following discrete-time system

$$x_d[i + 1] = ax_d[i] + bu_d[i] \tag{122}$$

where $b \neq 0$. Suppose we want $|x_d[i + 1] - x_d[i]| \leq \delta$ for some given value of δ . In order to achieve this, we impose the following condition on $u_d[i]$:

$$\frac{-\delta + (1 - a)x_d[i]}{b} \leq u_d[i] \leq \frac{\delta + (1 - a)x_d[i]}{b} \tag{123}$$

Generally, we want to choose a single value for $u_d[i]$ in our feedback law. Given the condition above, we may want to choose the $u_d[i]$ with the lowest energy (i.e., minimize $(u_d[i])^2$ subject to the constraint above). **For each of the following cases, give a single expression for $u_d[i]$ that minimizes $(u_d[i])^2$ subject to the constraint above:**

- i. $\frac{-\delta + (1 - a)x_d[i]}{b} \leq 0$ and $\frac{\delta + (1 - a)x_d[i]}{b} \geq 0$
- ii. $\frac{-\delta + (1 - a)x_d[i]}{b} \geq 0$ and $\frac{\delta + (1 - a)x_d[i]}{b} \geq 0$
- iii. $\frac{-\delta + (1 - a)x_d[i]}{b} \leq 0$ and $\frac{\delta + (1 - a)x_d[i]}{b} \leq 0$

You may not use min or max in your final answer.

Solution:

- i. In this case, $u_d[i] = 0$ is a valid control action that satisfies the constraints, and it minimizes $(u_d[i])^2$ globally. Hence, we choose $u_d[i] = 0$.
 - ii. Since $u_d[i] \geq 0$ and $(u_d[i])^2$ is monotonically increasing for non-negative values of $u_d[i]$, we choose $u_d[i] = \frac{-\delta + (1-a)x_d[i]}{b}$, the left endpoint of the interval.
 - iii. Since $u_d[i] \leq 0$ and $(u_d[i])^2$ is monotonically decreasing for non-positive values of $u_d[i]$, we choose $u_d[i] = \frac{\delta + (1-a)x_d[i]}{b}$, the right endpoint of the interval.
- (c) (5 pts.) Now, we can return to the nonlinear system in eq. (117). Suppose we want to impose a condition on linearization error as follows:

$$\left| f(x(t), u(t)) - \hat{f}(x(t), u(t)) \right| \leq \pi \tag{124}$$

where $f(x(t), u(t)) = \sin(x(t)) + u(t)$ from eq. (117) and $\hat{f}(x(t), u(t)) = x(t) + u(t)$ from eq. (118).

Find the least restrictive condition on $x(t)$, of the form $x_\ell \leq x(t) \leq x_u$, for concrete values of $x_\ell, x_u \in \mathbb{R}$, such that the condition on linearization error in eq. (124) is satisfied.

(HINT: In this case, $\left| f(x(t), u(t)) - \hat{f}(x(t), u(t)) \right|$ increases monotonically as $x(t)$ moves further from the origin.)

(HINT: Your expression for $\left| f(x(t), u(t)) - \hat{f}(x(t), u(t)) \right|$ shouldn't depend on $u(t)$.)

(HINT: What is $\sin(\pi)$?)

From here, you can set δ from part (b) to $\delta = \min(x_d[i] - x_\ell, x_u - x_d[i])$. This gives you the optimal control law to ensure your linearization error is bounded.

Solution: From part (b), $\hat{f}(x(t), u(t)) = x(t) + u(t)$. Hence,

$$\left| f(x(t), u(t)) - \hat{f}(x(t), u(t)) \right| = |\sin(x(t)) + u(t) - (x(t) + u(t))| \tag{125}$$

$$= |\sin(x(t)) - x(t)| \leq \pi \tag{126}$$

From the hint, $\sin(\pi) = 0$, and so $|\sin(\pi) - \pi| = \pi$. Since $|\sin(x(t)) - x(t)|$ is monotonically increasing in $|x(t)|$, we can choose $x_\ell = -\pi$ and $x_u = \pi$.