EECS 16B Designing Information Devices and Systems II
Fall 2021 Final Exam Solutions Final Exam

Section 0: Pre-exam questions (4 points)

1. **Honor Code: 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.

2. **How are you hoping to relax during winter break? (2 pts)**

   Solution: Any answer that actually answers the question gets full credit.

3. **Think about something that you know how to do well and enjoy doing. What is it? (2 pts)**

   Solution: Any answer that actually answers the question gets full credit.

4. SVD Calculation (11 points)

(a) (7pts) Let $A = \begin{bmatrix} 0 & 2 \\ \sqrt{2} & 0 \\ 0 & 1 \end{bmatrix}$. The eigenvalues of $AA^\top = \begin{bmatrix} 4 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 1 \end{bmatrix}$ are $5, 2, 0$ with corresponding unnormalized eigenvectors $\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$.

In addition, we know that:

$$A^\top \begin{bmatrix} 2 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & \sqrt{2} & 0 \\ 5 & 0 & 0 \end{bmatrix} \tag{1}$$

Solution: Observe that this problem is just asking you to carry out a process that you studied in HW11's question entitled "SVD from the other side." You had practice doing this in discussion 10B as well.

Write out the singular value decomposition (SVD) of A in any form you choose (outer product form, compact, or full). (No need to show work.)

Solution: The normalized eigenvectors of $AA^\top$ are the columns of U:

$$U = \begin{bmatrix} \frac{2\sqrt{5}}{5} & 0 & \frac{-\sqrt{5}}{5} \\ 0 & 1 & 0 \\ \frac{\sqrt{5}}{5} & 0 & \frac{2\sqrt{5}}{5} \end{bmatrix} \tag{2}$$

Also, the singular values of A are the square roots of the eigenvalues of $AA^\top$. Since Σ is the same shape as A and contains the singular values along its diagonal (with zeros elsewhere):

$$\Sigma = \begin{bmatrix} \sqrt{5} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \tag{3}$$

Finally, we can find the columns of V from the columns of U and the nonzero singular values:

$$\vec{v}_1 = \frac{1}{\sigma_1} A^\top \vec{u}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{4}$$

$$\vec{v}_2 = \frac{1}{\sigma_2} A^\top \vec{u}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{5}$$

$$V = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{6}$$

Writing the SVD in outer product form yields:

$$A = \sum_{i=1}^{2} \sigma_i \vec{u}_i \vec{v}_i^\top = \sqrt{5} \begin{bmatrix} \frac{2\sqrt{5}}{5} \\ 0 \\ \frac{\sqrt{5}}{5} \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} + \sqrt{2} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{7}$$

where $\vec{u}_i$ and $\vec{v}_i$ are the columns of U and V, respectively. We can also write the full SVD:

$$A = U\Sigma V^\top = \begin{bmatrix} \frac{2\sqrt{5}}{5} & 0 & \frac{-\sqrt{5}}{5} \\ 0 & 1 & 0 \\ \frac{\sqrt{5}}{5} & 0 & \frac{2\sqrt{5}}{5} \end{bmatrix} \begin{bmatrix} \sqrt{5} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{8}$$

or the compact SVD:

$$A = U\Sigma V^\top = \begin{bmatrix} \frac{2\sqrt{5}}{5} & 0 \\ 0 & 1 \\ \frac{\sqrt{5}}{5} & 0 \end{bmatrix} \begin{bmatrix} \sqrt{5} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{9}$$

For these specific numbers, once you had the first two columns of U and know the Σ, you could also just read off what the $\vec{v}_i$ have to be since you knew that the first column of A is a scaled multiple of $\vec{u}_2$ and the second column of A is a scaled multiple of $\vec{u}_1$. The numbers were made simple enough to let you do this as a valid way of solving the problem.

(b) (4pts) **What is the best rank 1 approximation of A (i.e., what is the rank 1 matrix B that minimizes $\|A - B\|_F$)? Write your answer as a 3×2 dimensional matrix.** (No need to show work)

Solution: The best rank 1 approximation is $B = \sigma_1 \vec{u}_1 \vec{v}_1^\top$, or the outer product of the vectors $\vec{u}_i$ and $\vec{v}_i$ that correspond to the largest singular value σ_1.

In this case:

$$B = \sqrt{5} \begin{bmatrix} \frac{2\sqrt{5}}{5} \\ 0 \\ \frac{\sqrt{5}}{5} \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{10}$$

For this specific problem, you could also observe that the two columns of A are clearly orthogonal to each other. Consequently, the best rank 1 approximation can only capture one of them, which means that you just want the heavier one which is the second one. From this, you could have immediately seen that B must have this form. This is also a valid alternative way to solve this problem.

5. Gram-Schmidt with Complex Vectors (11 points)

Suppose you are given two complex vectors $\vec{v}_1 = \begin{bmatrix} \mathrm{j} \\ 1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. In addition, the complex inner products are given as:

$$\langle \vec{v}_1, \vec{v}_1 \rangle = \vec{v}_1^* \vec{v}_1 = 2 \tag{11}$$

$$\langle \vec{v}_2, \vec{v}_2 \rangle = \vec{v}_2^* \vec{v}_2 = 1 \tag{12}$$

$$\langle \vec{v}_1, \vec{v}_2 \rangle = \vec{v}_2^* \vec{v}_1 = +\mathrm{j} \tag{13}$$

$$\langle \vec{v}_2, \vec{v}_1 \rangle = \vec{v}_1^* \vec{v}_2 = -\mathrm{j} \tag{14}$$

Use the Gram-Schmidt algorithm to generate an orthonormal sequence of vectors $(\vec{u}_1, \vec{u}_2)$ from the list of vectors $(\vec{v}_1, \vec{v}_2)$, starting with $\vec{v}_1$. **If you start the Gram-Schmidt algorithm with $\vec{v}_2$, you will receive zero credit for this entire problem.**

Solution: Observe that this problem is essentially the same as what you did in discussion 14B, except that we flipped the sign of $\vec{v}_2$ and simplified it further by just making things two dimensional instead of adding one more $\vec{v}_3$.

(a) (3pts) **What is the first vector $\vec{u}_1$?** (No need to show work)

Solution: The first step of the Graham-Schmidt algorithm is to normalize the first vector $\vec{v}_1$ to get $\vec{u}_1$. Remember that $\langle \vec{v}_1, \vec{v}_1 \rangle = \|\vec{v}_1\|^2$, so $\|\vec{v}_1\| = \sqrt{2}$.

$$\vec{u}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathrm{j} \\ 1 \end{bmatrix} \tag{15}$$

(b) (8pts) **What is the second vector $\vec{u}_2$?** (Show work)

Solution: The second unnormalized basis vector $\vec{r}_2$ is found by projecting $\vec{v}_2$ onto $\vec{u}_1$ and subtracting the projection from $\vec{v}_2$.

$$\vec{r}_2 = \vec{v}_2 - P_{\vec{u}_1} \vec{v}_2 \tag{16}$$

$$= \vec{v}_2 - \frac{\langle \vec{v}_2, \vec{u}_1 \rangle}{\langle \vec{u}_1, \vec{u}_1 \rangle} \vec{u}_1 \tag{17}$$

$$= \vec{v}_2 - \frac{\langle \vec{v}_2, \vec{v}_1 \rangle}{\|\vec{v}_1\|} \frac{\vec{v}_1}{\|\vec{v}_1\|} \tag{18}$$

$$= \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \left(\frac{-\mathrm{j}}{2} \right) \begin{bmatrix} \mathrm{j} \\ 1 \end{bmatrix} \tag{19}$$

$$= \frac{1}{2} \begin{bmatrix} 1 \\ \mathrm{j} \end{bmatrix} \tag{20}$$

The normalized basis vector is

$$\vec{u}_2 = \frac{\vec{r}_2}{\|\vec{r}_2\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ \mathrm{j} \end{bmatrix} \tag{21}$$

6. Choosing cost functions for learning classification (9 points)

We have labeled data $\{(\vec{x}_i, y_i)\}$ where the labels y_i are either '+' or '-'. We want to learn a vector $\vec{w}$ so that we can use the sign of $\vec{w}^\top \vec{x}$ to classify $\vec{x}$. To do this, we will minimize a sum $c_{\text{total}}(\vec{w}) = \sum_i c^{y_i}(\vec{x}_i^\top \vec{w})$. We consider cost functions:

(a) Squared loss: $c^+(p) = (p-1)^2$ and $c^-(p) = (p-(-1))^2$

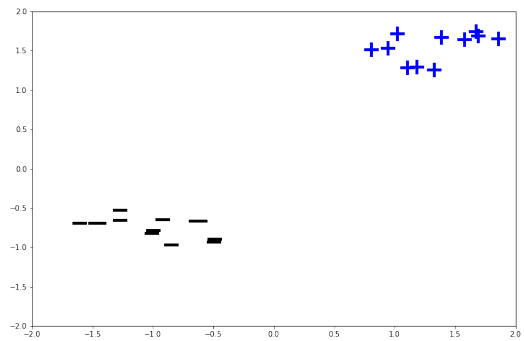
(b) Exponential loss: $c^+(p) = e^{-p}$ and $c^-(p) = e^{+p}$

(c) Logistic loss: $c^+(p) = \ln\left(1 + e^{-p}\right)$ and $c^-(p) = \ln\left(1 + e^{+p}\right)$

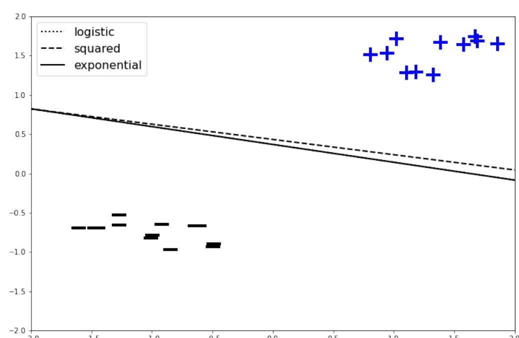
Solution: This problem engages exactly with the ideas in HW13's problem entitled "Linearization to help classification: discovering logistic regression and how to solve it" as well as what you saw in Discussion 13A. The jupyter notebooks associated with those is where you saw examples like these.

For the plotted data, which cost functions will result in learning reasonable classifiers? (Multiple options might be correct and should be marked for full credit, but every plot has at least one correct option.)

Cost Function	
Squared	☐
Exponential	☐
Logistic	☐



Solution:



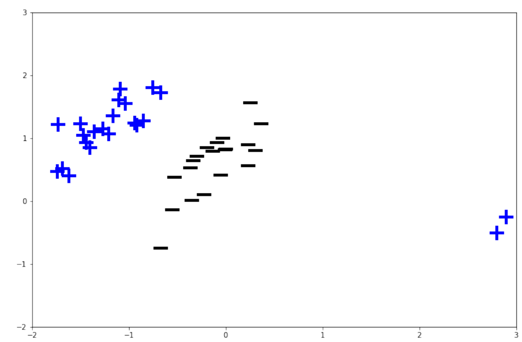
	Squared	**Exponential**	**Logistic**
Reasonable Choice	●	●	●

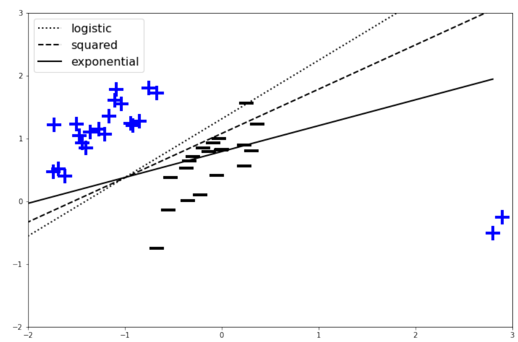
Why? This data is pretty balanced and so any of these approaches will work. There are no outliers of any kind to confuse least-squares.

Solution:

Cost Function	
Squared	☐
Exponential	☐
Logistic	☐

Note: You want to make sure that the two right-most '+' points don't influence your learned classifier by too much.

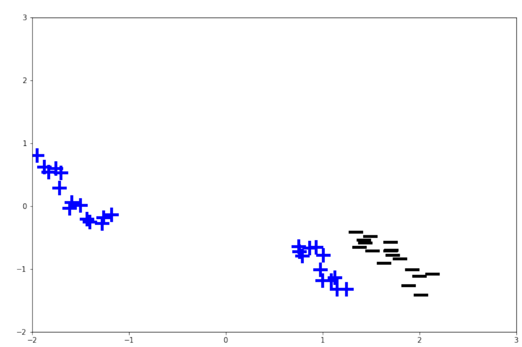




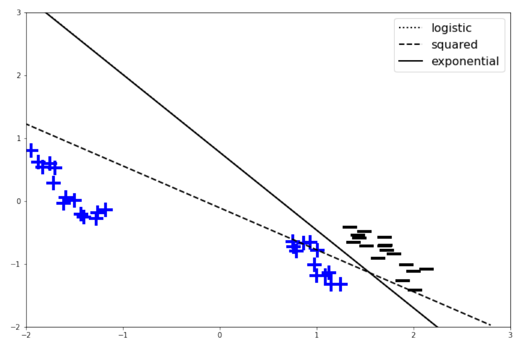
	Squared	**Exponential**	**Logistic**
Reasonable Choice	☐	☐	●

This data clearly has a "wrongly classified" outlier that is a "+" point deep in what should be "-" territory. Both the squared loss and (especially) exponential loss will be quite sensitive to this while logistic loss will end up largely ignoring this point. In such cases, we really need to be using logistic loss of the choices offered here.

Cost Function	
Squared	☐
Exponential	☐
Logistic	☐



Solution:



	Squared	Exponential	Logistic
Reasonable Choice	☐	●	●

In this case, we clearly have two masses of points in each category. The + category is particularly striking because the second mass of points is deep within what we would consider "+" category. These type of points will confuse squared loss quite a bit while both exponential and logistic loss don't care about points that are deeply within their own proper territories. They focus on things closer to the boundaries.

7. Nonlinear Feedback Control (14 points)

Consider the following differential equation model for a control system:

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{x}(t) = \vec{f}(\vec{x}(t)) + \vec{g}(\vec{x}(t))u(t) \tag{22}$$

Assume that our state $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is two-dimensional and our input $u \in \mathbb{R}$ is a scalar. Let our dynamics functions be

$$\vec{f}(\vec{x}) = \begin{bmatrix} x_1^3 \\ x_2^2 \end{bmatrix}, \ \vec{g}(\vec{x}) = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} \tag{23}$$

Solution: This problem is in the same style as what students did in the HW 13 problem entitled "Latch," except that there's no surrounding motivation here. It also follows what was done in Discussion 12A closely.

(a) (4pts) **Show that** $x_1^* = -1, x_2^* = 1, u^* = 1$ **is an equilibrium point.** (i.e. if we find ourselves exactly in this state with exactly this input and no disturbances, we will not move from here.)

Solution:

$$\vec{h}(\vec{x}, u) = \vec{f}(\vec{x}) + \vec{g}(\vec{x})u = \begin{bmatrix} x_1^3 + x_2 u \\ x_2^2 + x_1 u \end{bmatrix} \tag{24}$$

$$\vec{h}(\vec{x}^*, u^*) = \begin{bmatrix} (-1)^3 + (1)(1) \\ (1)^2 + (-1)(1) \end{bmatrix} = \begin{bmatrix} -1+1 \\ 1-1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{25}$$

Since $\frac{\mathrm{d}x}{\mathrm{d}t} = \vec{h}(\vec{x}^*, u^*) = \vec{0}$, this is a valid equilibrium point.

(b) (8pts) Recall that our nonlinear controlled differential equation for the state $\vec{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$ is given by:

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{x}(t) = \begin{bmatrix} x_1^3(t) + x_2(t)u(t) \\ x_2^2(t) + x_1(t)u(t) \end{bmatrix}.$$

The equilibrium of interest has $\vec{x}^* = \begin{bmatrix} -1 \\ +1 \end{bmatrix}$ and $u^* = +1$.

We define the variables $\vec{\widetilde{x}} = \vec{x} - \vec{x}^*$ and $\widetilde{u} = u - u^*$ as the deviations from the equilibrium point $(\vec{x}^*, u^*)$.

What should the matrix A and vector $\vec{b}$ be in our locally linearized model $\frac{\mathrm{d}}{\mathrm{d}t}\vec{\widetilde{x}}(t) \approx A\vec{\widetilde{x}}(t) + \vec{b}\widetilde{u}(t)$? (Show work and give specific numbers for A and $\vec{b}$.)

Solution:

$$\vec{h}(\vec{x}, u) = \vec{f}(\vec{x}) + \vec{g}(\vec{x})u = \begin{bmatrix} x_1^3 + x_2 u \\ x_2^2 + x_1 u \end{bmatrix} \tag{26}$$

$$\frac{\partial \vec{h}}{\partial \vec{x}} = \begin{bmatrix} 3x_1^2 & u \\ u & 2x_2 \end{bmatrix} \tag{27}$$

$$A = \left.\frac{\partial \vec{h}}{\partial \vec{x}}\right|_{\vec{x}=\vec{x}^*, u=u^*} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \tag{28}$$

$$\frac{\partial \vec{h}}{\partial u} = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} \tag{29}$$

$$\vec{b} = \left.\frac{\partial \vec{h}}{\partial u}\right|_{\vec{x}=\vec{x}^*, u=u^*} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{30}$$

(c) (2pts) **Is the resulting linearized system from the above part controllable?**

Solution:

$$\mathcal{C} = \begin{bmatrix} \vec{b} & A\vec{b} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix} \tag{31}$$

The controllability matrix $\mathcal{C}$ has full rank of 2 so the system is controllable.

8. Bode Plot Analysis of Band Stop and Notch Filters (35 points)

Consider a system with an input that consists of three signals, i.e. $v_i(t) = s_1(t) + s_2(t) + n(t)$, where $s_1(t) = 2\cos\left(10^2 t\right)$ and $s_2(t) = 2\cos\left(10^7 t\right)$ are the two desired signals, and $n(t) = 10\cos\left(10^4 t\right)$ is an interference signal. We wish to construct a filter so that the amplitude of the interference signal is attenuated to be at most $\frac{1}{10}$ times the amplitude of the desired signals after filtering.

Solution: This problem largely follows what you've seen in Discussion 5B as well as HW 6's problem entitled: "Bandpass Filter: Lowpass and Highpass Cascade." This is was further reviewed for you a bit in Discussion 15A where we went over a related problem from the Spring 21 final.

We will attempt two different ways of achieving this behavior. In parts (a), (b), and (c), we will analyze the first approach, which is to use a band stop filter that attenuates the interference signal and passes the desired signals. In part (d), we will analyze the second approach, which is to create a notch at the interference signal frequency.

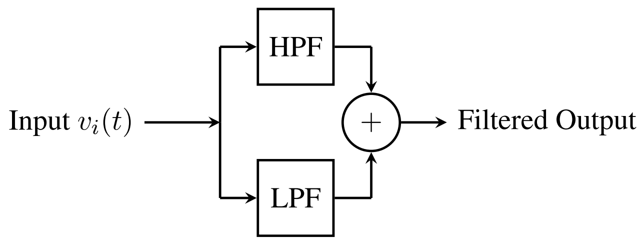


Figure 1: Block diagram of a Band Stop Filter. The HPF and LPF blocks represent a high-pass filter and a low-pass filter respectively. *NOTE*: Remember, you can always turn block diagrams into circuits by using the corresponding circuit for the transfer function in the block diagram, and isolating transfer functions with buffers.

(a) (10pts) Let's start by analyzing the HPF path in fig. 1, which should pass $s_2(t) = 2\cos(10^7 t)$ and attenuate $n(t) = 10\cos(10^4 t)$. Assume that the high-pass filter transfer function is $H_{\text{HPF}}(\text{j}\omega) = \frac{\text{j}\frac{\omega}{10^6}}{1+\text{j}\frac{\omega}{10^6}}$.

i. **Draw the magnitude and phase Bode plots of the high-pass filter $H_{\text{HPF}}(\text{j}\omega) = \frac{\text{j}\frac{\omega}{10^6}}{1+\text{j}\frac{\omega}{10^6}}$.**

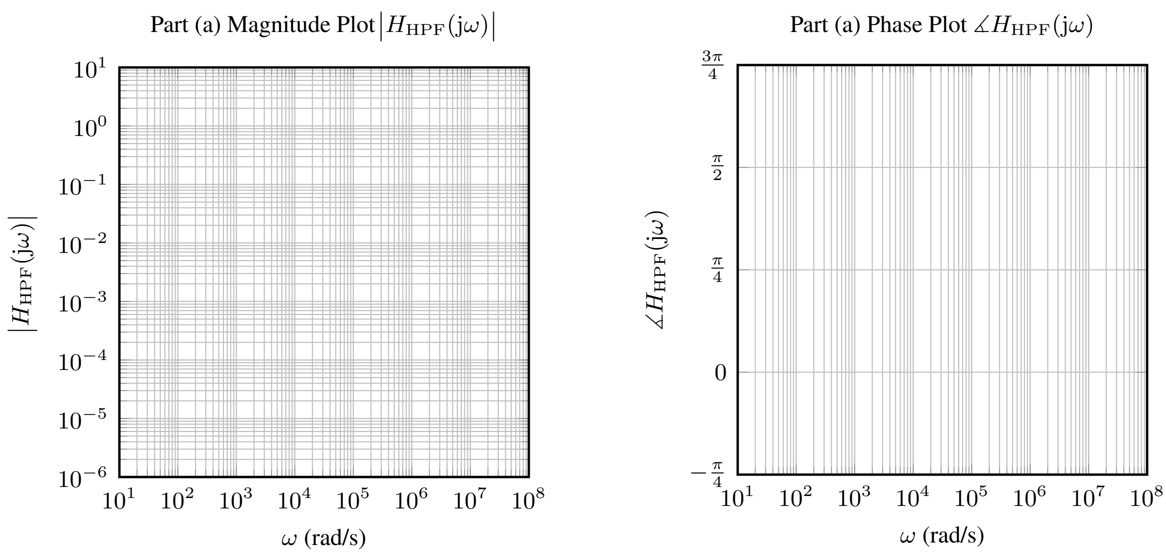


Figure 2: Part (a) Magnitude and Phase Bode Plots for the transfer function $H_{\text{HPF}}(\text{j}\omega)$.

Solution: The magnitude and phase Bode plots are shown in fig. 3. **Students only need to draw the Bode approximation straight lines (dashed lines in the plots) for full credit.** The true curves are drawn in solid lines just for reference.

This is a classic high-pass filter plot where we know that the corner frequency is 10^6 based on the form of the transfer function. We see the classic drop of a factor of 10 in magnitude for every factor of 10 in frequency to the left of this corner frequency and holding constant at $1 = 10^0$ to the right of this corner frequency. This is because when ω is small, the numerator is shrinking while the denominator is not changing since the denominator is dominated by the 1.

For phase, at the corner frequency we are at $\frac{\pi}{4}$ since $\frac{\text{j}}{1+\text{j}} = \frac{1+\text{j}}{2}$ clearly has equal and positive real and imaginary parts. Under the Bode approximation, moving a factor of 10 or more in frequency to the left makes everything dominated by the real term in the denominator, which makes the overall transfer function almost entirely imaginary due to the $+\text{j}$ in the numerator. This gives a phase of $\frac{\pi}{2}$. When the frequency is high — more than a factor of 10 above the cutoff frequency, everything is dominated by the imaginary term in the denominator, which makes the overall transfer function essentially real and positive. This gives a phase of 0. Connecting these regions with straight lines in between gives the classic phase plot approximation for a high-pass filter.

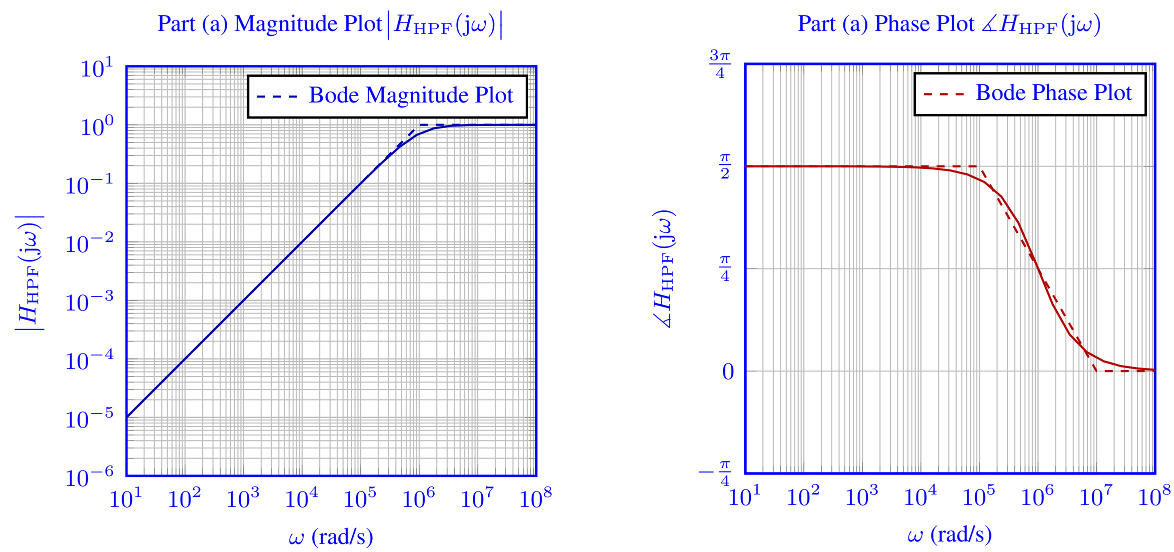


Figure 3: Part (a) Magnitude and Phase Bode Plots for the transfer function $H_{\mathrm{HPF}}(\mathrm{j}\omega)$.

ii. **By reading the corresponding values from the Bode plots, write down the approximate output signal expressions corresponding to $s_1(t) = 2\cos\left(10^2 t\right)$, $s_2(t) = 2\cos\left(10^7 t\right)$ and $n(t) = 10\cos\left(10^4 t\right)$ after high-pass filtering.** *NOTE*: You do not need to compute exact magnitude and phase values using the transfer function. Just use the Bode approximation by reading from the plots.

Solution: The outputs corresponding to $s_1(t)$, $s_2(t)$ and $n(t)$ are given by

$$\widetilde{s_1}(t) = 2 \times 10^{-4} \cos\left(10^2 t + \frac{\pi}{2}\right) = -2 \times 10^{-4} \sin\left(10^2 t\right) \tag{32}$$

$$\widetilde{n}(t) = 0.1 \cos\left(10^4 t + \frac{\pi}{2}\right) = -0.1 \sin\left(10^4 t\right) \tag{33}$$

$$\widetilde{s_2}(t) = 2\cos\left(10^7 t\right) \tag{34}$$

Full credit is given for any of the equivalent trigonometric expressions.
Here, we can read off the magnitude 10^{-4} and phase $\frac{\pi}{2}$ from the plot for frequency 10^2.
Similarly, we can see the magnitude 10^{-2} and phase $\frac{\pi}{2}$ from the plot for frequency 10^4. Multiplying 10^{-2} with the initial amplitude 10 gives us the 0.1.
Finally, we can see the magnitude $1 = 10^0$ and phase 0 from the plot for frequency 10^7. Consequently, s_2 is basically unchanged after passing through this filter.

(b) (17pts) Now let's analyze the LPF path in fig. 1, which should pass $s_1(t) = 2\cos(10^2 t)$ and attenuate $n(t) = 10\cos(10^4 t)$. Assume that you are allowed to use multiple copies of a low-pass filter given by $H_{\text{LPF}}(\text{j}\omega) = \frac{1}{1+\text{j}\frac{\omega}{10^3}}$ and a unity gain buffer given by $H_{\text{buf}}(\text{j}\omega) = 1$.

i. Remember that we wish to attenuate the amplitude of $n(t) = 10\cos(10^4 t)$ to be at most $\frac{1}{10}$ times the amplitude of $s_1(t) = 2\cos(10^2 t)$ after filtering. **At least how many copies of $H_{\text{LPF}}(\text{j}\omega)$ and $H_{\text{buf}}(\text{j}\omega)$ do we need to cascade to achieve this goal? What is the overall transfer function $H_1(\text{j}\omega)$ of this cascaded low-pass filter?** *NOTE*: You do not have to simplify the expression.

Solution: We require 2 first-order low-pass filters cascaded with a buffer.
We can see this because a single first-order low-pass filter wouldn't work. Why? Because the spacing of the frequencies is such that the cutoff 10^3 is only a factor of 10 away from the interference $n(t)$'s frequency of 10^4. This means that the simple low-pass filter can only attenuate by a factor of 10, and that would bring the amplitude from 10 to just 1. But we need the amplitude to be below $\frac{1}{10} \times 2 = 0.2$. This means we need another factor of 10 attenuation, which we can get by cascading two of the first-order low-pass filters through a buffer to prevent loading effects.
The overall transfer function is given by

$$H_1(\text{j}\omega) = H_{\text{LPF}}(\text{j}\omega) \cdot H_{\text{buf}}(\text{j}\omega) \cdot H_{\text{LPF}}(\text{j}\omega) = \frac{1}{\left(1 + \text{j}\frac{\omega}{10^3}\right)^2} \tag{35}$$

ii. **Draw the magnitude and phase Bode plots of the overall low-pass filter $H_1(\text{j}\omega)$.**

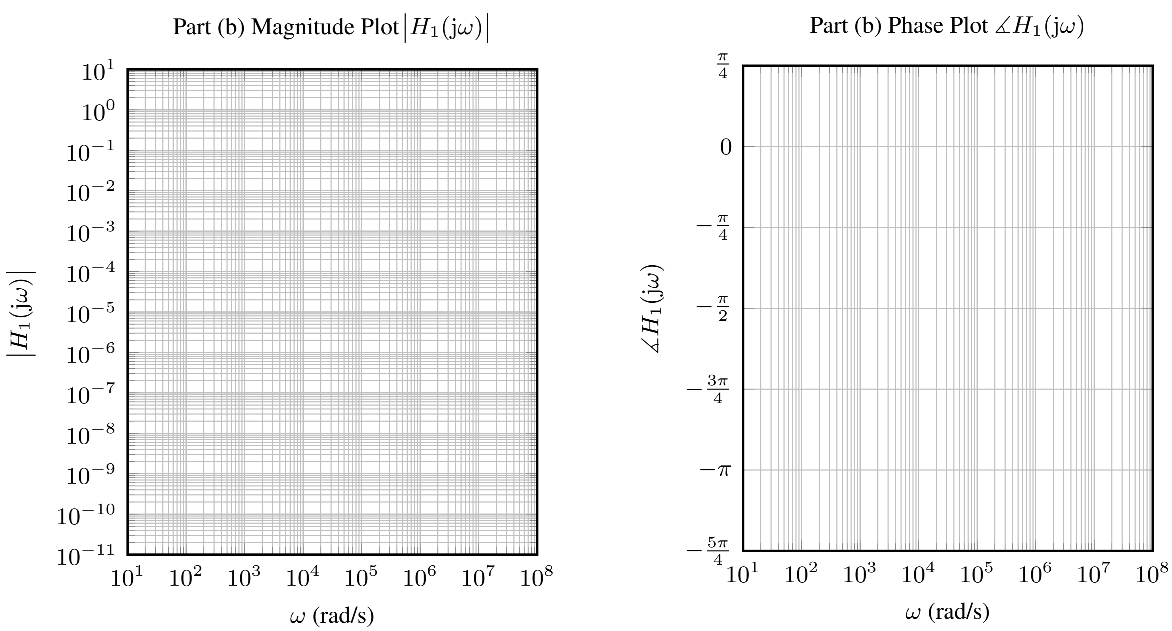


Figure 4: Part (b) Magnitude and Phase Bode Plots for the transfer function $H_1(\text{j}\omega)$.

Solution: The magnitude and phase Bode plots are shown in fig. 5. **Students only need to draw the Bode approximation straight lines (dashed lines in the plots) for full credit.** The true curves are drawn in solid lines just for reference.

This is a composition of two classic first-order low-pass filters. To the left of the cutoff frequency, the magnitude is 1. Since $1 \times 1 = 1$, the same magnitude exists for H_1 to the left of the cutoff frequency in the Bode approximation. For frequencies to the right of the cutoff frequency, the classic first-order low-pass filter drops the magnitude by a factor of 10 for every factor of 10 increase in the frequency. Since $10 \times 10 = 100$, this means that H_1 drops by a factor of 100 for every factor of 10 increase in the frequency.

The same argument works for phase. Sufficiently to the left (a factor of 10 or more) of the cutoff frequency, a classic first-order low-pass filter has a phase of 0 since it is just dominated by a positive real number. More than a factor of 10 above the cutoff frequency, the imaginary term in the denominator dominates, which turns into a phase of $-\frac{\pi}{2}$ for a classic first-order low-pass filter. Because of how multiplication of complex numbers works in polar coordinates, phases add when we cascade two filters. This means that for H_1, the phase becomes $-\pi$ once you are right of a factor of 10 above the cutoff frequency. Connecting these with a straight line gives the Bode approximation for phase, which does the right thing at the cutoff frequency since $\left(\frac{1}{1+\mathrm{j}}\right)^2 = \frac{1}{2\mathrm{j}} = -\frac{1}{2}\mathrm{j}$ giving a phase of $-\frac{\pi}{2}$.

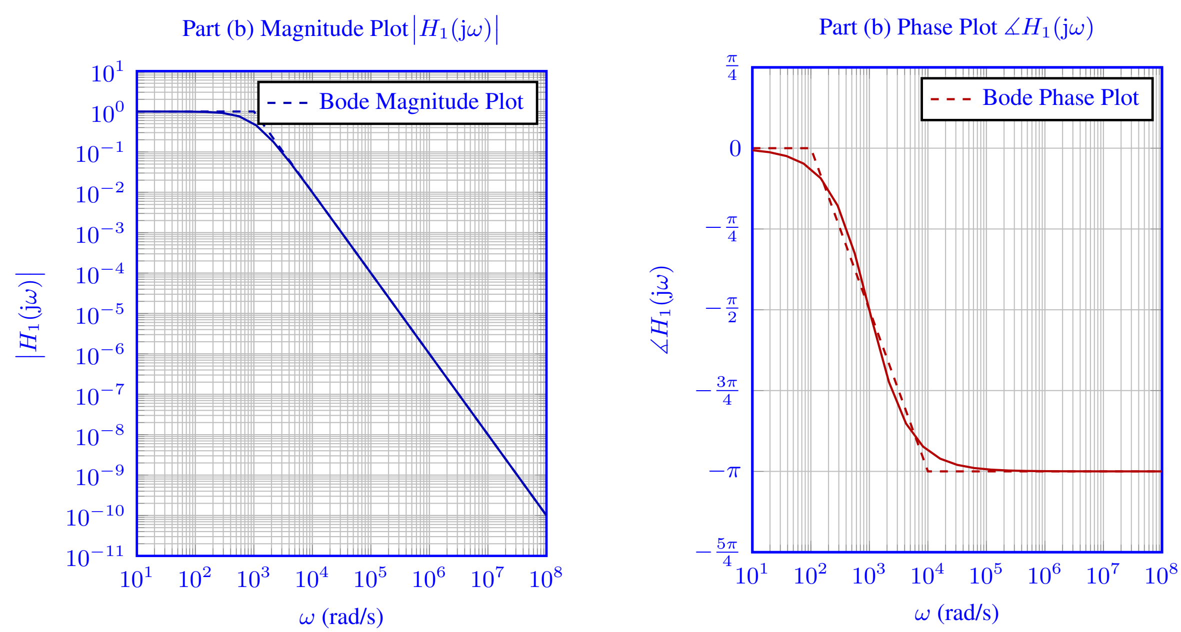


Figure 5: Part (b) Magnitude and Phase Bode Plots for the transfer function $H_1(\mathrm{j}\omega)$.

iii. **By reading the corresponding values from the Bode plots, write down the approximate output signal expressions corresponding to $s_1(t) = 2\cos(10^2 t)$, $s_2(t) = 2\cos(10^7 t)$ and $n(t) = 10\cos(10^4 t)$ after the overall low-pass filtering using $H_1(\mathrm{j}\omega)$.** *NOTE*: You do not need to compute exact magnitude and phase values using the transfer function. Just use the Bode approximation by reading from the plots.

Solution: The outputs corresponding to $s_1(t)$, $s_2(t)$ and $n(t)$ are given by

$$\widetilde{s_1}(t) = 2\cos\left(10^2 t\right) \tag{36}$$

$$\widetilde{n}(t) = 0.1\cos\left(10^4 t - \pi\right) = -0.1\cos\left(10^4 t\right) \tag{37}$$

$$\widetilde{s_2}(t) = 2 \times 10^{-8}\cos\left(10^7 t - \pi\right) = -2 \times 10^{-8}\cos\left(10^7 t\right) \tag{38}$$

Full credit is given for any of the equivalent trigonometric expressions.
For s_1, we can read off the magnitude and phase and the result is that this survives the filter entirely unchanged.
For n, we can read off the magnitude as 10^{-2} for the frequency 10^4 with a phase of $-\pi$. Since $10 \times 10^{-2} = 0.1$, we get the expression above.
For s_2, we can read off the magnitude as 10^{-8} for the frequency 10^7 with a phase of $-\pi$.

We have redrawn fig. 1 here for your convenience with the cutoff/corner frequencies labeled.

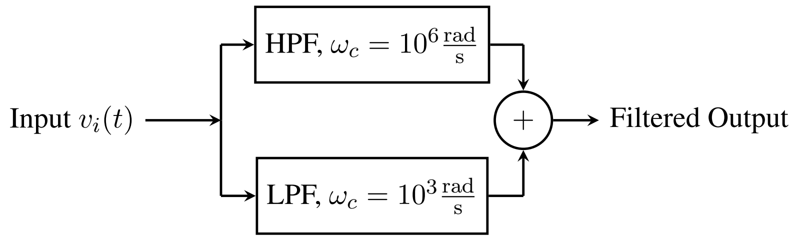


Figure 6: Block diagram of a Band Stop Filter.

(c) (3pts) **Explain in words what is the effect of the complete filter above on the three signals $s_1(t) = 2\cos(10^2 t)$, $s_2(t) = 2\cos(10^7 t)$ and $n(t) = 10\cos(10^4 t)$?** You do not have to provide numerical answers for this part.

Solution: $s_1(t)$ is passed by LPF but severely attenuated by HPF, so the final result is approximately $s_1(t)$.

$s_2(t)$ is passed by HPF but severely attenuated by LPF, so the final result is approximately $s_2(t)$.

$n(t)$ is attenuated by both the LPF and HPF.

Hence the overall filter in fig. 1 exhibits band stop behavior, i.e. it passes low and high frequencies but attenuates frequencies in the middle band.

Students get full credit for just the text description above. Detailed quantitative explanation is shown below just for reference.

By adding the results from parts (a) (iii) and (b) (iii), the final output signal expressions are

$$\widetilde{s_1}(t) = 2 \times 10^{-4} \cos\left(10^2 t + \frac{\pi}{2}\right) + 2\cos\left(10^2 t\right) \approx 2\cos\left(10^2 t\right) = s_1(t) \tag{39}$$

$$\widetilde{n}(t) = 0.1\cos\left(10^4 t + \frac{\pi}{2}\right) + 0.1\cos\left(10^4 t - \pi\right) = 0.1\sqrt{2}\cos\left(10^4 t + \frac{3\pi}{4}\right) \tag{40}$$

$$\widetilde{s_2}(t) = 2\cos\left(10^7 t\right) + 2 \times 10^{-8}\cos\left(10^7 t - \pi\right) \approx 2\cos\left(10^7 t\right) = s_2(t) \tag{41}$$

Hence the magnitude of the filtered interference signal is indeed less than $\frac{1}{10}$ of the magnitude of the desired signals. Notice that this is also true if we consider the "self-interference" that the phase-shifted and attenuated signals might be construed as contributing through the other filter branch.

The overall transfer function of the band stop filter is

$$H_{\text{BSF}}(\mathrm{j}\omega) = H_{\text{HPF}}(\mathrm{j}\omega) + H_1(\mathrm{j}\omega) \tag{42}$$

$$= \frac{\mathrm{j}\frac{\omega}{10^6}}{1 + \mathrm{j}\frac{\omega}{10^6}} + \frac{1}{\left(1 + \mathrm{j}\frac{\omega}{10^3}\right)^2} \tag{43}$$

The band stop behavior can be easily visualized from the magnitude and phase plots in fig. 7.

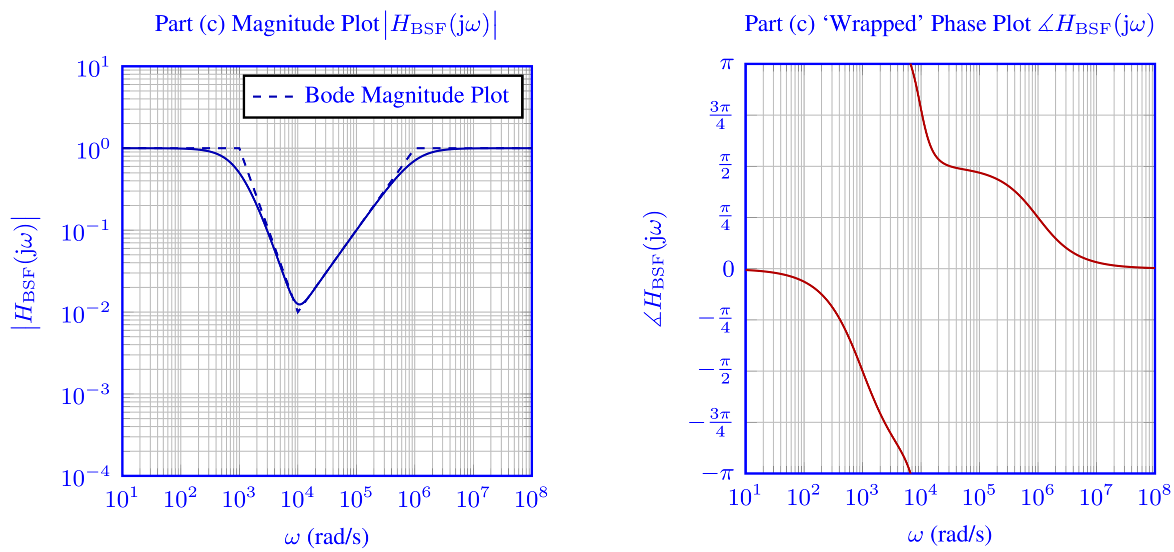


Figure 7: Part (c) Magnitude and 'wrapped' Phase Bode Plots for the transfer function $H_{\mathrm{BSF}}(\mathrm{j}\omega)$.

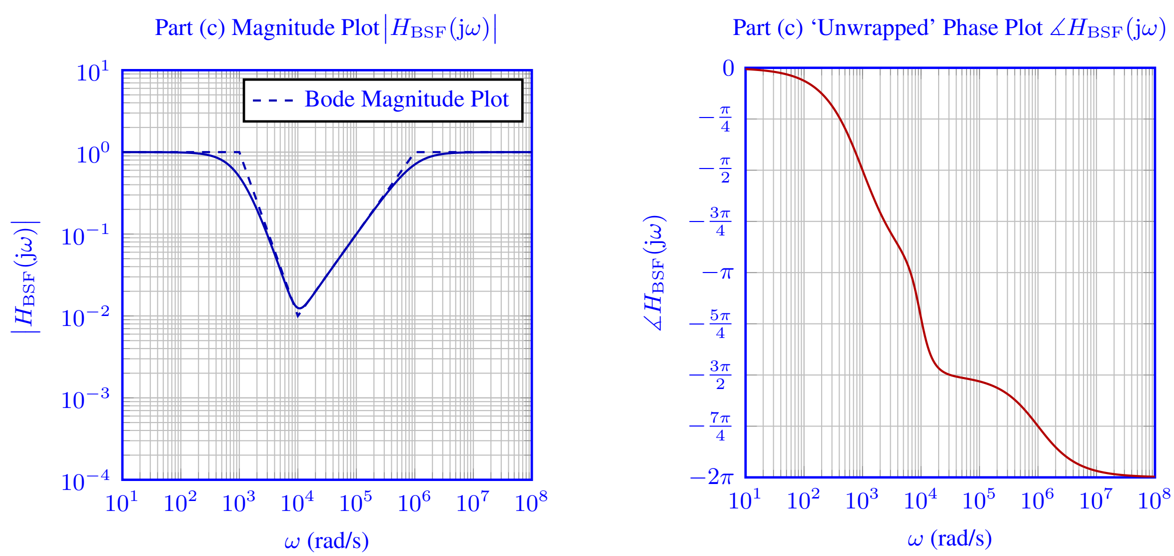


Figure 8: Part (c) Magnitude and 'unwrapped' Phase Bode Plots for the transfer function $H_{\mathrm{BSF}}(\mathrm{j}\omega)$.

Note that both these magnitude and phase plots in fig. 7 follow the low-pass filter plots in fig. 5 at low frequencies and the high-pass filter plots in fig. 3 at high frequencies. Also note that the phase plots in fig. 7 and fig. 8 are equivalent.

If you think carefully about the kinds of approximations that we are doing, you can see how these plots also reflect those approximations. There's just one segment in the phase plot where our style of approximations break down, and that is in the neighborhood of 10^4 frequency. Here, the amplitude is so small that the phase can move around very easily when things get added together.

(d) (5pts) Finally, let's consider the second approach. Since we know the exact frequency of the interference signal $n(t)$, we can use a Notch filter, as shown in fig. 9, to create a notch at that frequency, instead of using the previous band stop filter. This should completely attenuate $n(t) = 10\cos\left(10^4 t\right)$ and pass signals of all other frequencies, including $s_1(t)$ and $s_2(t)$. **If $C = 10\,\mu\text{F}$ in the Notch filter in fig. 9, calculate the inductance L needed to completely block signals at $10^4\,\frac{\text{rad}}{\text{s}}$.** *NOTE*: You do not need to know the value of R to solve this question.

Solution: This part can be viewed as a natural follow-on to what you saw in both lecture and in Discussion 5A on how to think about a filter that includes both an inductor and a capacitor.

(*HINT: At $\omega = 10^4\,\frac{\text{rad}}{\text{s}}$, what do you want the impedance of the series connection of L and C to be?*)

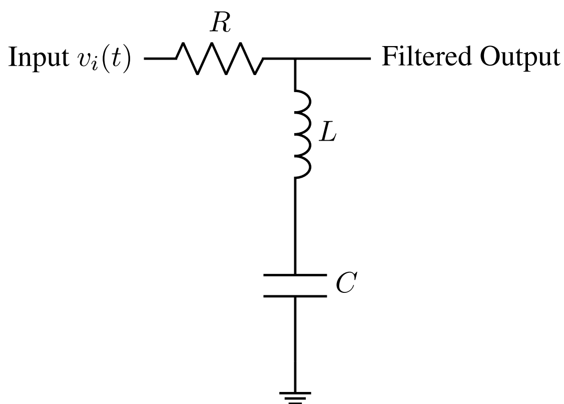


Figure 9: LC Notch filter

Solution: The inductance value can be calculated as

$$L = \frac{1}{\omega_n^2 C} = \frac{1}{\left(10^4\right)^2 \times 10 \times 10^{-6}} = 1\,\text{mH} \tag{44}$$

The above calculation uses the formula that the resonant frequency is $\frac{1}{\sqrt{LC}}$. If a student didn't already have that on their cheat sheet, it is possible to simply say that you want the combined series impedance to be zero at $\omega_n = 10^4$ so that the voltage divider doesn't let the interference signal through. Since the capacitor impedance is $-\frac{\text{j}}{\omega_n C} = -\text{j}10$, we need the inductor impedance to be $+\text{j}10 = \text{j}\omega_n L$. This means that $L = 10^{-3}\text{H}$ is the solution.

Students get full credit for just the computations above. Detailed explanation is shown below.

The transfer function is given by

$$H_{\text{notch}}(\text{j}\omega) = \frac{\text{j}\left(\omega L - \frac{1}{\omega C}\right)}{R + \text{j}\left(\omega L - \frac{1}{\omega C}\right)} \tag{45}$$

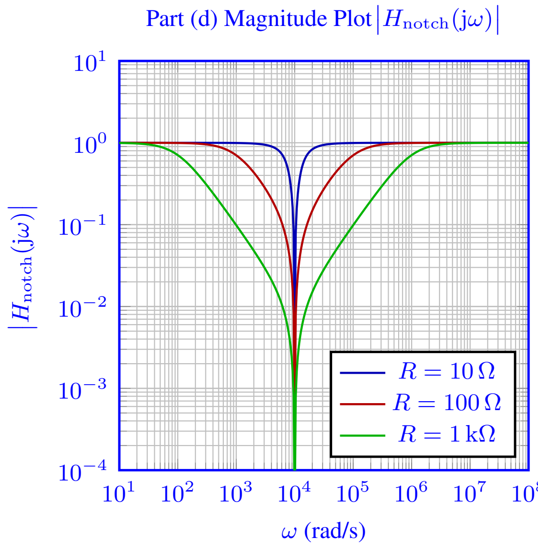
Since we wish to create a notch at the interference frequency $\omega_n = 10^4\,\frac{\text{rad}}{\text{s}}$, we have

$$H_{\text{notch}}(\text{j}\omega_n) = 0 \tag{46}$$

$$\implies \omega_n L = \frac{1}{\omega_n C} \tag{47}$$

$$\implies L = \frac{1}{\omega_n^2 C} = 1\,\text{mH} \tag{48}$$

The notch behavior can be seen in the magnitude and phase plots in fig. 10 for different values of R. Higher values of R increase the width of the notch.



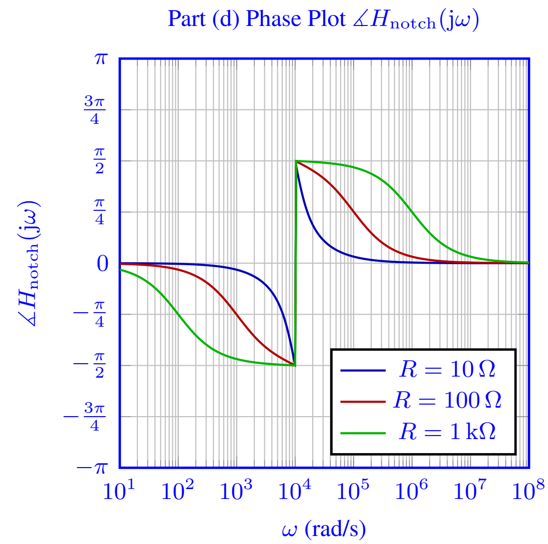


Figure 10: Part (d) Magnitude and Phase Bode Plots for the transfer function $H_{\text{notch}}(j\omega)$.

If $R = 10\,\Omega$ for example, then by reading from the plots in fig. 10, the final output signal expressions are

$$\widetilde{s_1}(t) = 2\cos\left(10^2 t\right) = s_1(t) \tag{49}$$

$$\widetilde{n}(t) \approx 0 \tag{50}$$

$$\widetilde{s_2}(t) = 2\cos\left(10^7 t\right) = s_2(t) \tag{51}$$

9. Using a Nonlinear NMOS Transistor for Amplification (35 points)

Consider the following schematic where $V_{DD} = 1.5\,V$, $R_L = 400\,\Omega$ and the NMOS transistor has threshold voltage $V_{th} = 0.2\,V$. We are interested in analyzing the response of this circuit to input voltages of the form $V_{in}(t) = V_{in,DC} + v_{in,AC}(t)$, where $V_{in,DC}$ is some constant voltage and $v_{in,AC}(t) = 0.001\cos(\omega t)V$ is a sinusoidal signal whose magnitude is much smaller than $V_{in,DC}$.

Solution: This problem is most directly related to HW 12's problem "Linearizing for understanding amplification," where you had to go through a nearly identical process, just with a different equation governing the relationship between the input voltage and the current.

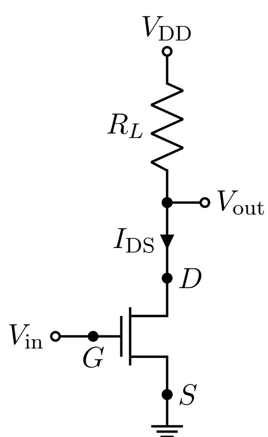
The I-V relationship of an NMOS can be modeled as non-linear functions over different regions of operation. For simplicity, let's just focus on the case when $0 \leq V_{GS} - V_{th} < V_{DS}$. In this regime of interest, the relevant I-V relationship is given by

$$I_{DS}(V_{GS}) = \frac{K}{2}(V_{GS} - V_{th})^2 \tag{52}$$

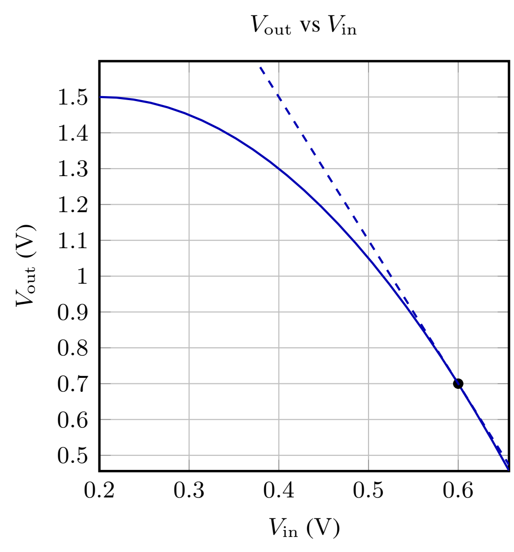
where K is a constant that depends on the NMOS transistor size and properties.

Solution: You might be wondering why this model has the source-drain connection modeled by a current source instead of a resistor like you saw at the beginning of the course when we were discussing models for NMOS and PMOS transistors. This is a consequence of the underlying physics of the transistor in this regime. When the drain voltage is higher than the gate voltage, then the "puddle" of charge carriers that form under the gate thins out dramatically in the neighborhood of the drain. This pinching off creates a limit on how much current can flow because as the drain voltage goes higher this is offset by how increasingly restricted the charge puddle is near the drain. These two effects end up essentially canceling each other and instead of getting a resistor-like behavior where more drain-source voltage results in more drain-source current, the current just saturates. You can learn more about this in 130 where you will learn in more detail why and how this happens, and why the resulting relationship with the gate current has this quadratic form.

For our purposes here, this is just some equation governing the I-V relationship that we can differentiate if we want to.



(a) NMOS Transistor circuit



(b) V_{out} vs V_{in} in the regime of interest. Tangent is drawn at the operating point $V_{in,DC} = 0.6\,V$, $V_{out,DC} = 0.7\,V$

Figure 11: NMOS figures.

From Ohm's law and KCL, we know that

$$V_{\text{out}}(t) = V_{\text{DD}} - R_L I_{\text{DS}}(t). \tag{53}$$

Note from fig. 11a that $V_{\text{in}} = V_{\text{GS}}$ and $V_{\text{out}} = V_{\text{DS}}$. In fig. 11b, we can see the curve of V_{out} vs V_{in} in the transistor operating regime of interest.

(a) (4 pts) **Using eq. (52) and eq. (53), express $V_{\text{out}}(t)$ as a function of $V_{\text{in}}(t)$ symbolically.** (You can use $V_{\text{DD}}, R_L, V_{\text{in}}, K, V_{\text{th}}$ in your answer.)

Solution:

$$V_{\text{out}}(t) = V_{\text{DD}} - R_L I_{\text{DS}}(t)\Big|_{V_{\text{GS}}=V_{\text{in}}(t)} \tag{54}$$

$$= V_{\text{DD}} - R_L \frac{K}{2}\left(V_{\text{in}}(t) - V_{\text{th}}\right)^2 \tag{55}$$

The input $V_{\text{in}}(t) = V_{\text{in,DC}} + v_{\text{in,AC}}(t)$ results in an output of the form $V_{\text{out}}(t) = V_{\text{out,DC}} + v_{\text{out,AC}}(t)$. Since V_{DD} is constant, we can linearize $V_{\text{out}}(t)$ around $V_{\text{out,DC}}$ from eq. (53), as illustrated in fig. 11b.

$$A_v(V_{\text{in}}, V_{\text{out}})\bigg|_{V_{\text{in}}^\star = V_{\text{in,DC}}} = \frac{\mathrm{d}V_{\text{out}}}{\mathrm{d}V_{\text{in}}}\bigg|_{V_{\text{in}}^\star = V_{\text{in,DC}}} = -R_L \frac{\mathrm{d}I_{\text{DS}}}{\mathrm{d}V_{\text{in}}}\bigg|_{V_{\text{in}}^\star = V_{\text{in,DC}}} \tag{56}$$

$$= -R_L \frac{\mathrm{d}I_{\text{DS}}}{\mathrm{d}V_{\text{GS}}}\bigg|_{V_{\text{GS}}^\star = V_{\text{in,DC}}} \tag{57}$$

$$= -R_L g_m(V_{\text{GS}})\Big|_{V_{\text{GS}}^\star = V_{\text{in,DC}}} \tag{58}$$

Here A_v is defined as the linearized voltage gain, which is illustrated by the slope of the tangent to the V_{out} vs V_{in} curve in fig. 11b at the point $\left(V_{\text{in,DC}},\ V_{\text{out,DC}}\right)$, and $g_m = \frac{\mathrm{d}I_{\text{DS}}}{\mathrm{d}V_{\text{GS}}}$ is the transistor transconductance linearized around the point $V_{\text{GS}}^\star = V_{\text{in,DC}}$.

(b) (10pts) To linearize the whole circuit around the operating point $V_{\text{in}}^\star = V_{\text{in,DC}}$, as shown in eq. (56), eq. (57), eq. (58), we need to linearize the non-linear transistor I-V curve, given by $I_{\text{DS}}(V_{\text{GS}}) = \frac{K}{2}(V_{\text{GS}} - V_{\text{th}})^2$ in eq. (52) to find the linearized transconductance gain $g_m = \frac{\mathrm{d}I_{\text{DS}}}{\mathrm{d}V_{\text{GS}}}$.

i. **Using eq. (52), derive the linearized transconductance gain $g_m = \frac{\mathrm{d}I_{\text{DS}}}{\mathrm{d}V_{\text{GS}}}$ symbolically.**
NOTE: Please simplify your answer.

Solution: Using eq. (52), the small signal transconductance gain is given by

$$g_m = \frac{\mathrm{d}I_{\text{DS}}}{\mathrm{d}V_{\text{GS}}} = K(V_{\text{GS}} - V_{\text{th}}) \tag{59}$$

ii. **From the following options, choose which circuit element can be used to represent the transistor in a linearized circuit with $\Delta I_{\text{DS}} = g_m \Delta V_{\text{GS}}$,** where ΔI_{DS} and ΔV_{GS} are small deviations around I_{DS} and V_{GS} respectively.

Select one	Choices
◯	resistor between G and S terminals
◯	resistor between D and S terminals
◯	voltage controlled current source between D and S terminals

Solution: The NMOS transistor behaves as a voltage controlled current source between D and S terminals in this regime, with ΔV_{GS} being the control voltage, as shown in fig. 12. (Yes, we gave away the answer in the next part of the question.)

Select one	Choices
○	resistor between G and S terminals
○	resistor between D and S terminals
●	voltage controlled current source between D and S terminals

iii. Using $K = 0.025\,\frac{1}{\Omega\,\mathrm{V}}$, and $V_{\mathrm{th}} = 0.2\,\mathrm{V}$ and $R_L = 400\,\Omega$, **calculate the numerical values of the following quantities at the operating point** $\left(V_{\mathrm{GS}}^{\star} = 0.6\,\mathrm{V},\ V_{\mathrm{DS}}^{\star} = 0.7\,\mathrm{V}\right)$**:**

- **linearized transconductance gain** g_m
- **linearized voltage gain** $A_v = -R_L g_m$ **from eq. (58)**

NOTE: Please simplify the numerical answers — this will also help you check your answer graphically against fig. 11b if you want.

Solution: The numerical values at the given operating point are:

$$g_m = 0.025 \times (0.6 - 0.2) = 0.01\,\frac{1}{\Omega} \tag{60}$$

$$A_v = -400 \times 0.01 = -4 \tag{61}$$

(c) (5pts) The circuit in fig. 12 below is a linearized model for the transistor circuit in fig. 11a, according to eq. (56), eq. (57), eq. (58). **Find the transfer function $H_1(\mathrm{j}\omega) = \frac{\widetilde{v}_{\text{out,AC}}(\mathrm{j}\omega)}{\widetilde{v}_{\text{in,AC}}(\mathrm{j}\omega)}$ in terms of g_m and R_L.**

Solution: These latter parts of this problem are asking you to do a phasor-based analysis of the kinds of circuits that you saw in HWs 2 and 3 when we were looking inside of an op-amp — except that instead of a voltage-controlled voltage source there, we have a voltage-controlled current source here. This kind of phasor-based analysis of circuits is something that you have seen many times in HWs as well as discussion.

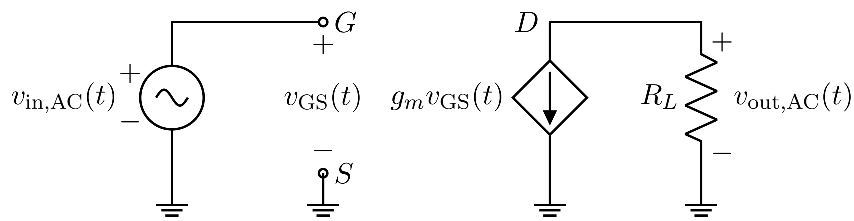


Figure 12: Small signal model for NMOS circuit in fig. 11a, according to eq. (56), eq. (57), eq. (58).

Solution: In phasor domain, the voltage drop across the resistor is given by

$$\widetilde{v}_{\text{out,AC}}(\mathrm{j}\omega) = -R_L g_m \widetilde{v}_{\text{GS}}(\mathrm{j}\omega) \tag{62}$$

$$= -R_L g_m \widetilde{v}_{\text{in,AC}}(\mathrm{j}\omega) \tag{63}$$

Hence the transfer function is given by

$$\frac{\widetilde{v}_{\text{out,AC}}(\mathrm{j}\omega)}{\widetilde{v}_{\text{in,AC}}(\mathrm{j}\omega)} = -g_m R_L \tag{64}$$

We gave away this answer in eq. (58).
There is no frequency-dependence here.

(d) (8pts) Consider the following modified model of the transistor circuit in fig. 11a where the transistor has a drain capacitance C_D as shown in fig. 13 below. **Find the transfer function $H_2(\mathrm{j}\omega) = \frac{\widetilde{v}_{\text{out,AC}}(\mathrm{j}\omega)}{\widetilde{v}_{\text{in,AC}}(\mathrm{j}\omega)}$ in terms of g_m, R_L, C_D and $\mathrm{j}\omega$. What type of filter is implemented by this circuit model?**

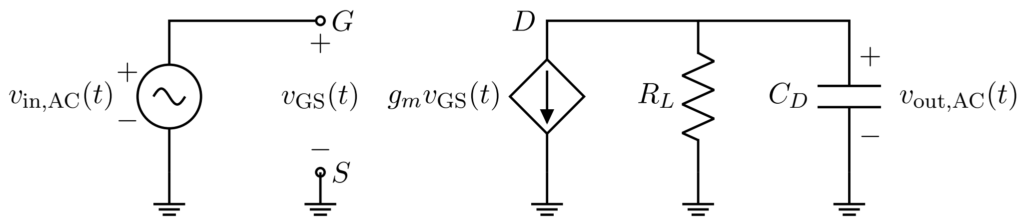


Figure 13: Small signal model for NMOS circuit in fig. 11a with drain capacitance.

Solution: Note that the voltage drop across the resistor and capacitor are the same, and so they are in parallel. In phasor domain, we can just find the equivalent impedance of this parallel combination and then use the definition of impedance to get the voltage drop.

$$Z_{eq} = (Z_R || Z_C) = \left(R_L || \frac{1}{\text{j}\omega C_D}\right) \tag{65}$$

$$= \frac{\frac{R_L}{\text{j}\omega C_D}}{R_L + \frac{1}{\text{j}\omega C_D}} = \frac{R_L}{1 + \text{j}\omega R_L C_D} \tag{66}$$

$$\widetilde{v}_{\text{out,AC}}(\text{j}\omega) = Z_{eq}\widetilde{I} = Z_{eq}(-g_m \widetilde{v}_{\text{GS}}(\text{j}\omega)) \tag{67}$$

$$= -\frac{R_L}{1 + \text{j}\omega C_D R_L} g_m \widetilde{v}_{\text{in,AC}}(\text{j}\omega) \tag{68}$$

Therefore, the transfer function is given by

$$\frac{\widetilde{v}_{\text{out,AC}}(\text{j}\omega)}{\widetilde{v}_{\text{in,AC}}(\text{j}\omega)} = -\frac{g_m R_L}{1 + \text{j}\omega C_D R_L} \tag{69}$$

This is a first order low pass filter.
You could also solve this through KCL but this is the most direct approach.

(e) (8pts) Consider the following modified model of the transistor circuit in fig. 11a where V_{in} is not ideal and has some resistance in series, R_{in}, and the transistor has a gate capacitance C_{GS} as shown in fig. 14 below. **Find the transfer function** $H_3(\mathrm{j}\omega) = \frac{\widetilde{v}_{\text{out,AC}}(\mathrm{j}\omega)}{\widetilde{v}_{\text{in,AC}}(\mathrm{j}\omega)}$ **in terms of** $R_{\text{in}}, C_{\text{GS}}, g_m, R_L, C_D$ **and** $\mathrm{j}\omega$**.** (*HINT: First analyze* $v_{\text{GS}}(t)$ *in phasor domain. Then try to re-use the result from the previous part.*)

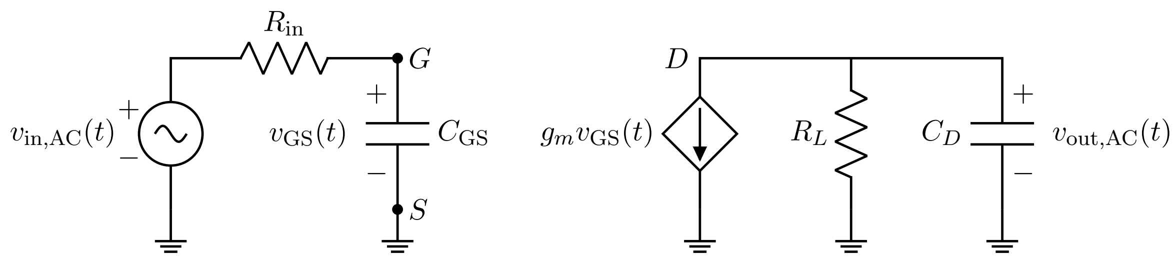


Figure 14: Small signal model for NMOS circuit in fig. 11a with non-ideal source and gate capacitance, in addition to drain capacitance.

Solution: First we analyze the voltage divider to get the transfer function from $\widetilde{v}_{\text{in,AC}}(\mathrm{j}\omega)$ to $\widetilde{v}_{\text{GS}}(\mathrm{j}\omega)$:

$$\frac{\widetilde{v}_{\text{GS}}(\mathrm{j}\omega)}{\widetilde{v}_{\text{in,AC}}(\mathrm{j}\omega)} = \frac{1}{1 + \mathrm{j}\omega C_{\text{GS}} R_{\text{in}}} \tag{70}$$

From the previous part, we know that

$$\frac{\widetilde{v}_{\text{out,AC}}(\mathrm{j}\omega)}{\widetilde{v}_{\text{GS}}(\mathrm{j}\omega)} = -\frac{g_m R_L}{1 + \mathrm{j}\omega C_D R_L} \tag{71}$$

Cascading these 2 transfer functions, we get the overall transfer function given by

$$\frac{\widetilde{v}_{\text{out,AC}}(\mathrm{j}\omega)}{\widetilde{v}_{\text{in,AC}}(\mathrm{j}\omega)} = -\frac{g_m R_L}{(1 + \mathrm{j}\omega C_D R_L)(1 + \mathrm{j}\omega C_{\text{GS}} R_{\text{in}})} \tag{72}$$

10. Minimum Norm Solutions for Circuits involving Resistors (32 points)

Consider a current i_s flowing into a network of two parallel resistors R_1 and R_2, as shown in fig. 15 below.

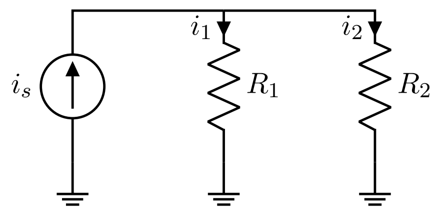


Figure 15: Current i_s dividing into i_1 and i_2.

From EECS 16A, we know that we can equate the voltage drops across the parallel resistors to derive $i_1 = \frac{R_2}{R_1+R_2} i_s$ and $i_2 = \frac{R_1}{R_1+R_2} i_s$. In this problem, we will try to derive the same current division result using the concept of minimum norm instead of voltage analysis.

Solution: This problem is directly testing ideas from Discussion 10A with a connection to HW 10's problem: weighted minimum norm.

It turns out that the current i_s will divide into two parts i_1 and i_2 in such a way that minimizes the total power dissipation $P = i_1^2 R_1 + i_2^2 R_2$ in the resistors.

(a) (8pts) **Argue that the current division result given by** $i_1 = \frac{R_2}{R_1+R_2} i_s$ **and** $i_2 = \frac{R_1}{R_1+R_2} i_s$ **minimizes the total power dissipation** $P = i_1^2 R_1 + i_2^2 R_2$ **using calculus.** Use the fact that KCL gives $i_2 = i_s - i_1$ to express P as a function of i_1 only. (*HINT: Once you solve for the optimal* i_1*, you don't have to do calculus again for* i_2*. Just use KCL.*)

Solution: We can minimize P as follows:

$$P = i_1^2 R_1 + i_2^2 R_2 = i_1^2 R_1 + (i_s - i_1)^2 R_2 \tag{73}$$

$$\implies \frac{\mathrm{d}P}{\mathrm{d}i_1} = 2 i_1 R_1 - 2(i_s - i_1) R_2 = 0 \tag{74}$$

$$\implies i_1 = \frac{R_2}{R_1 + R_2} i_s \tag{75}$$

$$\implies i_2 = i_s - i_1 = \frac{R_1}{R_1 + R_2} i_s. \tag{76}$$

Here, we notice that the quadratic has a positive constant multiplying the squared term, and so this must be the unique minimum.

Hence it is proved that $i_1 = \frac{R_2}{R_1+R_2} i_s$ and $i_2 = \frac{R_1}{R_1+R_2} i_s$ minimize P.

(b) (12pts) Instead of using calculus to minimize the total power dissipation P, we can represent the current division problem as a minimum norm problem. Consider the vector $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ where $x_1 = i_1\sqrt{R_1}$ and $x_2 = i_2\sqrt{R_2}$. Notice that $P = i_1^2 R_1 + i_2^2 R_2 = x_1^2 + x_2^2 = \|\vec{x}\|^2$.

i. **Find the row vector A so that the KCL constraint $i_1 + i_2 = i_s$ can be written as $A\vec{x} = i_s$.**

Solution: The total power dissipation given by $P = \|\vec{x}\|^2$ has to be minimized. KCL gives us

$$i_1 + i_2 = i_s \tag{77}$$

$$\implies \frac{x_1}{\sqrt{R_1}} + \frac{x_2}{\sqrt{R_2}} = i_s \tag{78}$$

$$\implies \begin{bmatrix} \frac{1}{\sqrt{R_1}} & \frac{1}{\sqrt{R_2}} \end{bmatrix} \vec{x} = i_s \tag{79}$$

$$\implies A\vec{x} = i_s \tag{80}$$

So $A = \begin{bmatrix} \frac{1}{\sqrt{R_1}} & \frac{1}{\sqrt{R_2}} \end{bmatrix}$.

ii. Using the A matrix you found above, **what is the minimum norm solution to $A\vec{x} = i_s$?** Show your work.

To help you save computation, the compact SVD of a general 1×2 row vector is given by

$$\begin{bmatrix} a & b \end{bmatrix} = \underbrace{\begin{bmatrix} 1 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sqrt{a^2+b^2} \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} \frac{a}{\sqrt{a^2+b^2}} & \frac{b}{\sqrt{a^2+b^2}} \end{bmatrix}}_{V^\top} \tag{81}$$

Solution: We know that the minimum norm solution for the system $A\vec{x} = i_s$ is given by $\widetilde{\vec{x}} = A^\dagger i_s$, where $A^\dagger$ is the pseudo-inverse of A. Using eq. (81), the compact SVD of A is given by

$$A = U\Sigma V^\top \tag{82}$$

$$\begin{bmatrix} \frac{1}{\sqrt{R_1}} & \frac{1}{\sqrt{R_2}} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sqrt{\frac{R_1+R_2}{R_1 R_2}} \end{bmatrix}}_{\Sigma} \underbrace{\frac{1}{\sqrt{R_1+R_2}} \begin{bmatrix} \sqrt{R_2} & \sqrt{R_1} \end{bmatrix}}_{V^\top} \tag{83}$$

The pseudo-inverse is given by

$$A^\dagger = V\Sigma^{-1}U^\top \tag{84}$$

$$= \underbrace{\frac{1}{\sqrt{R_1+R_2}} \begin{bmatrix} \sqrt{R_2} \\ \sqrt{R_1} \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} \sqrt{\frac{R_1 R_2}{R_1+R_2}} \end{bmatrix}}_{\Sigma^{-1}} \underbrace{\begin{bmatrix} 1 \end{bmatrix}}_{U^\top} \tag{85}$$

$$= \begin{bmatrix} \frac{R_2\sqrt{R_1}}{R_1+R_2} \\ \frac{R_1\sqrt{R_2}}{R_1+R_2} \end{bmatrix} \tag{86}$$

Hence the minimum norm solution is

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{R_2\sqrt{R_1}}{R_1+R_2} \\ \frac{R_1\sqrt{R_2}}{R_1+R_2} \end{bmatrix} i_s \tag{87}$$

iii. **Transform the minimum norm solution of $A\vec{x} = i_s$ to the original variables i_1 and i_2, and confirm that the result is $i_1 = \frac{R_2}{R_1+R_2} i_s$ and $i_2 = \frac{R_1}{R_1+R_2} i_s$ as the current-divider formula predicts.** Show your work.

Solution: Changing variables back to the current division problem, we have

$$\begin{bmatrix} i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} \frac{x_1}{\sqrt{R_1}} \\ \frac{x_2}{\sqrt{R_2}} \end{bmatrix} = \begin{bmatrix} \frac{R_2 i_s}{R_1+R_2} \\ \frac{R_1 i_s}{R_1+R_2} \end{bmatrix} \tag{88}$$

This matches the current division ratio from regular voltage analysis.

(c) (12 pts) We can solve any arbitrarily complicated circuit network using KCL and norm minimization, following the same technique that we used for the simple network in fig. 15. Consider a resistor network which has n resistor branches, with currents $i_1, i_2, \ldots, i_n$ across the branch resistances $R_1, R_2, \ldots, R_n$ respectively, and m total nodes each with current sources $i_{s_1}, i_{s_2}, \ldots, i_{s_m}$, which may be positive, negative or zero, as shown in fig. 16. Then the m KCL equations at the m nodes can be written as $K\vec{i} = \vec{i}_s$, where $K \in \mathbb{R}^{m\times n}$, $\vec{i} = \begin{bmatrix} i_1 \\ i_2 \\ \vdots \\ i_n \end{bmatrix} \in \mathbb{R}^n$, and $\vec{i}_s = \begin{bmatrix} i_{s_1} \\ i_{s_2} \\ \vdots \\ i_{s_m} \end{bmatrix} \in \mathbb{R}^m$. This KCL constraint $K\vec{i} = \vec{i}_s$ completely captures what is visualized in fig. 16, so you don't have to write any additional KCL. Note that fig. 15 is a simple example of fig. 16 with $n = 2$ and $m = 1$.

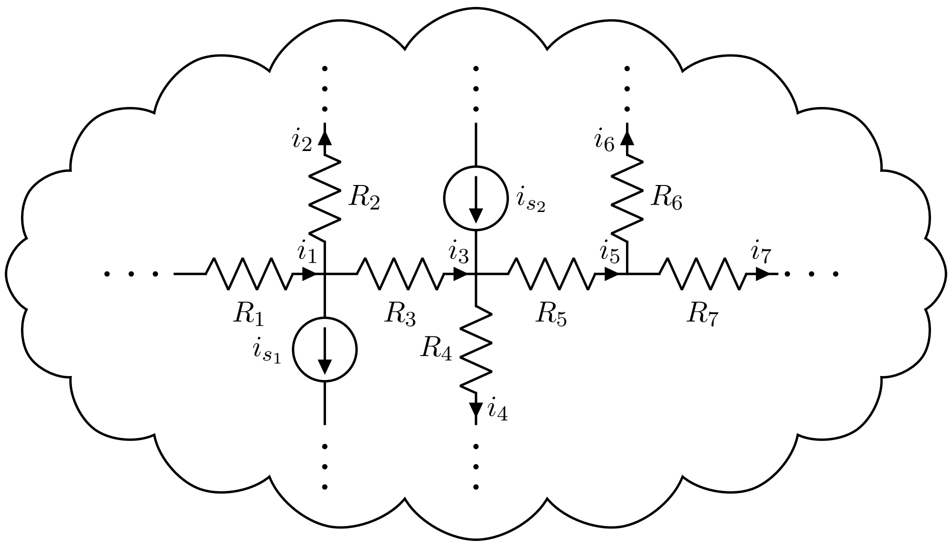


Figure 16: A section of an arbitrarily complicated network with n branches and m nodes.

i. We can change variables to $\vec{x} = D\vec{i}$ to represent the KCL constraint $K\vec{i} = \vec{i}_s$ as $A\vec{x} = \vec{i}_s$, and so the minimization of dissipated power $P = \sum_{j=1}^{n} i_j^2 R_j$ is just the minimization of $\sum_{j=1}^{n} x_j^2 = \|\vec{x}\|^2$.

Find the diagonal matrix D, and then find the matrix A in terms of D and K.

(*HINT: Look at how $\vec{x}$ was defined in the previous part.*)

Solution: Define a vector $\vec{x} \in \mathbb{R}^n$ where $x_j = i_j\sqrt{R_j}\ \forall j \in [1, n]$. Then we can change variables from $\vec{i}$ to $\vec{x}$ using the relation

$$\vec{x} = \underbrace{\begin{bmatrix} \sqrt{R_1} & 0 & \ldots & 0 \\ 0 & \sqrt{R_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sqrt{R_n} \end{bmatrix}}_{D} \vec{i} \tag{89}$$

$$\implies \vec{i} = D^{-1}\vec{x} \tag{90}$$

The diagonal matrix $D \in \mathbb{R}^{n\times n}$ has non-zero diagonal elements, hence it is invertible. Hence the

KCL can be represented as

$$K\vec{i} = \vec{i}_s \tag{91}$$

$$\implies \underbrace{KD^{-1}}_{A}\vec{x} = \vec{i}_s \tag{92}$$

ii. Assume the compact SVD of A is given by $A = U\Sigma V^\top$. **Use the minimum norm solution to $A\vec{x} = \vec{i}_s$ to solve for $\vec{i}$.** Recall from the previous part that $\vec{x} = D\vec{i}$. Your final answer for $\vec{i}$ can only use U, Σ, V, D, $\vec{i}_s$ as well as standard matrix operations like inverses, etc.

Solution: We can find $\vec{i}$ by using the min norm solution for eq. (92) as follows:

$$\vec{x} = A^\dagger \vec{i}_s = V\Sigma^{-1}U^\top \vec{i}_s \tag{93}$$

$$\implies \vec{i} = D^{-1}V\Sigma^{-1}U^\top \vec{i}_s. \tag{94}$$

This connection between power-dissipation minimization and electrical circuits goes even deeper. With the tools that you learn in Math 53 and EECS 127, you can see that the very concept of voltage itself can be understood via Lagrange multipliers (dual variables) associated with the KCL constraints that are active at each node. There is a further sense in which the reason that "state" is associated with inductors and capacitors is precisely because there is energy stored in these circuit elements. This is intimately related to the concept of the Lagrangian formulation of mechanics that you might encounter in Physics and Mechanical Engineering. Different formulations of "minimalist principles" are ubiquitous in science and engineering.

11. A Proof in the Complex Case (12 points)

Suppose M is a generic $m \times n$ complex matrix with rank r and SVD $M = U\Sigma V^* = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^*$, where the matrices $U = [\vec{u}_1 \cdots \vec{u}_m]$ and $V = [\vec{v}_1 \cdots \vec{v}_n]$ have orthonormal columns according to the complex inner product (i.e. the U and V matrices are unitary) and Σ is diagonal with real non-negative diagonal entries sorted in non-ascending order.

Recall that we defined the Frobenius norm as

$$\|M\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |M_{ij}|^2} \tag{95}$$

and note that the same definition works for complex matrices.

Solution: This problem is a combination of HW 12's problem "Low Rank Approximation of a Matrix" with HW 14's problem "Adapting proofs to the complex case."

Suppose $k < r$. **Prove that the best rank-at-most-k approximation to M in the Frobenius norm is given by $\sum_{i=1}^{k} \sigma_i \vec{u}_i \vec{v}_i^*$.**

In other words, prove that no matter what the collection of vectors $\{\vec{p}_i\}$ and $\{\vec{q}_i\}$ may be,

$$\left\| M - \sum_{i=1}^{k} \sigma_i \vec{u}_i \vec{v}_i^* \right\|_F \leq \left\| M - \sum_{i=1}^{k} \vec{p}_i \vec{q}_i^* \right\|_F . \tag{96}$$

You may use the following facts without proof:

- If U and V are square unitary matrices, then $\|UA\|_F = \|A\|_F = \|AV\|_F$.
- The best rank-at-most-k approximation to a diagonal $m \times n$ real matrix Σ with the diagonal consisting of real non-negative values σ_i sorted in non-ascending order is given by $\sum_{i=1}^{k} \sigma_i \vec{e}_{m,i} \vec{e}_{n,i}^*$, where $\vec{e}_{k,i}$ is the i^{th} column of a $k \times k$ identity matrix.

  In math: No matter what the collection of vectors $\{\vec{s}_i\}$ and $\{\vec{w}_i\}$ may be,

$$\left\| \Sigma - \sum_{i=1}^{k} \sigma_i \vec{e}_{m,i} \vec{e}_{n,i}^* \right\|_F \leq \left\| \Sigma - \sum_{i=1}^{k} \vec{s}_i \vec{w}_i^* \right\|_F . \tag{97}$$

- If U and V are square unitary matrices, then the rank of a matrix A is the same as the rank of UA and the rank of AV.
- The inverse of a square unitary matrix U is given by its conjugate transpose $U^{-1} = U^*$, which is also unitary.

Solution:

Write $M = U\Sigma V^*$ and fix $\{\vec{s}_i\}$, $\{\vec{w}_i\}$. We use the fact that $\|UA\|_F = \|A\|_F$ and $\|AV^*\|_F = \|A\|_F$ to get

$$\left\| \Sigma - \sum_{i=1}^{k} \sigma_i \vec{e}_{m,i} \vec{e}_{n,i}^* \right\|_F \leq \left\| \Sigma - \sum_{i=1}^{k} \vec{s}_i \vec{w}_i^* \right\|_F \tag{98}$$

$$\left\| U\left(\Sigma - \sum_{i=1}^{k} \sigma_i \vec{e}_{m,i}\vec{e}_{n,i}^{*}\right)\right\|_F \leq \left\| U\left(\Sigma - \sum_{i=1}^{k} \vec{s}_i\vec{w}_i^{*}\right)\right\|_F \tag{99}$$

$$\left\| U\left(\Sigma - \sum_{i=1}^{k} \sigma_i \vec{e}_{m,i}\vec{e}_{n,i}^{*}\right)V^*\right\|_F \leq \left\| U\left(\Sigma - \sum_{i=1}^{k} \vec{s}_i\vec{w}_i^{*}\right)V^*\right\|_F \tag{100}$$

$$\left\| U\Sigma V^* - \sum_{i=1}^{k} \sigma_i U\vec{e}_{m,i}\vec{e}_{n,i}^{*}V^*\right\|_F \leq \left\| U\Sigma V^* - \sum_{i=1}^{k} U\vec{s}_i\vec{w}_i^{*}V^*\right\|_F \tag{101}$$

$$\left\| U\Sigma V^* - \sum_{i=1}^{k} \sigma_i (U\vec{e}_{m,i})(V\vec{e}_{n,i})^{*}\right\|_F \leq \left\| U\Sigma V^* - \sum_{i=1}^{k} (U\vec{s}_i)(V\vec{w}_i)^{*}\right\|_F \tag{102}$$

$$\left\| M - \sum_{i=1}^{k} \sigma_i \vec{u}_i\vec{v}_i^{*}\right\|_F \leq \left\| M - \sum_{i=1}^{k} \vec{p}_i\vec{q}_i^{*}\right\|_F . \tag{103}$$

Here we have $\vec{p}_i = \sigma_i U\vec{s}_i$ and $\vec{q}_i = V\vec{w}_i$. Since U, V are orthonormal, and hence invertible, picking $\vec{p}_i = U\vec{s}_i$ is equivalent to picking $\vec{s}_i = U^{-1}\vec{p}_i = U^*\vec{p}_i$, and similarly picking $\vec{q}_i = V\vec{w}_i$ is equivalent to picking $\vec{w}_i = V^{-1}\vec{q}_i = V^*\vec{q}_i$. So we have shown the statement for all $\vec{p}_i$ and $\vec{q}_i$ as desired.

12. System ID for Continuous Systems (16 points)

So far we have seen system identification only done for discrete-time systems. But what if we really want to identify some underlying continuous-time model instead? We will explore how to do so in this problem.

(a) (8pts) Suppose we believed that our system was of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = \lambda x(t) + bu(t) \tag{104}$$

where $x(t) \in \mathbb{R}$ and $u(t) \in \mathbb{R}$ is a scalar input.

Given an initial condition $x(t_0)$, and that $u(t)$ is some constant input $\bar{u}$ over the interval $[t_0, t_f)$, then for all $t \in [t_0, t_f)$, we know that this differential equation eq. (104) has the unique solution

$$x(t) = x(t_0)e^{\lambda(t-t_0)} + \frac{e^{\lambda(t-t_0)} - 1}{\lambda}b\bar{u}. \tag{105}$$

Assume that we record the state at known times $\tau_0, \tau_1, \ldots, \tau_n$ as having corresponding state values $x_0 = x(\tau_0), x_1 = x(\tau_1), \ldots, x_n = x(\tau_n)$. The continuous-time input is known to be piecewise constant $u(t) = u_i$ for $t \in [\tau_i, \tau_{i+1})$, where we know the sequence of inputs $u_0, u_1, \ldots, u_{n-1}$.

Solution: This question part builds on Discussion 6A most closely, together with some of the framing of Discussion 6B. It tests whether students understand where the equations come from in system identification.

We now want to formulate this as a system ID question by relating the unknown parameters λ, b to the data we have. However, the relationship between the parameters and the data we collected is now non-linear. **For the data point x_{i+1}, use eq. (105) to write out how x_{i+1} should be related to λ and b in the form**

$$x_{i+1} = f(\lambda, x_i, \tau_i, \tau_{i+1}) + bg(\lambda, x_i, u_i, \tau_i, \tau_{i+1}). \tag{106}$$

What are the functions f and g?

Solution: We directly use (105) by plugging in $t = \tau_{i+1}$ and $t_0 = \tau_i$. We can use this formula as the input $u(t)$ is a constant u_i throughout this interval.

$$f(\lambda, x_i, \tau_i, \tau_{i+1}) = x_i e^{\lambda(\tau_{i+1}-\tau_i)} \tag{107}$$

$$g(\lambda, x_i, u_i, \tau_i, \tau_{i+1}) = \frac{u_i}{\lambda}(e^{\lambda(\tau_{i+1}-\tau_i)} - 1) \tag{108}$$

(b) (8pts) The previous part gave rise to a sequence of n equations of the form eq. (106). Because of observation noise and imperfection in our model, we are going to assume that these equations hold only approximately and hope to find values for the two parameters λ, b that minimize the cost function:

$$c(\lambda, b) = \sum_{i=0}^{n-1} \ell\left(x_{i+1}, f(\lambda, x_i, \tau_i, \tau_{i+1}) + bg(\lambda, x_i, u_i, \tau_i, \tau_{i+1})\right) \tag{109}$$

where $f(\lambda, x, \sigma, \tau)$ and $g(\lambda, x, u, \sigma, \tau)$ are given nonlinear scalar functions, and $\ell(s, p)$ is a loss function that penalizes how far the prediction p is from the measured state s. For example, you could use squared loss $\ell(s, p) = (s - p)^2$.

To find the best possible λ_*, b_*, you observe that you want to solve the nonlinear system of equations:

$$\frac{\partial}{\partial \lambda} c(\lambda, b)\Big|_{\lambda_*, b_*} = 0 \tag{110}$$

$$\frac{\partial}{\partial b} c(\lambda, b)\Big|_{\lambda_*, b_*} = 0 \tag{111}$$

and decide to do so using Newton's method starting with an initial guess λ_0, b_0 and linearizing the system of equations eq. (110) and eq. (111) to get a system of linear equations to solve at each step.

The system of linear equations at each iteration $j + 1$ can be expressed in vector form as:

$$A \begin{bmatrix} \lambda - \lambda_j \\ b - b_j \end{bmatrix} = \vec{y}. \tag{112}$$

Solution: This problem engages with the ideas in HW13's problem entitled "Linearization to help classification: discovering logistic regression and how to solve it" and even more so, what you saw in Discussion 13A. The only difference here is the larger context: the exact same techniques work to do system identification as to do classification.

What are the entries of the matrix A and the vector $\vec{y}$ in terms of the appropriate partial derivatives of $c(\lambda, b)$ evaluated at the appropriate arguments?

Assume that you can use PyTorch to compute whatever derivatives of $c(\lambda, b)$ that you want — all given functions are sufficiently differentiable. You don't have to take the derivatives by hand. You just need to tell us what derivatives and what arguments to evaluate them at.

Solution: We want to solve a nonlinear vector equation $\vec{f}(\vec{x}) = \vec{0}$. Linearizing this equation gives us

$$\vec{f}(\vec{x}) \approx \vec{f}(\vec{x}^*) + \frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}^*)(\vec{x} - \vec{x}^*) = \vec{0} \tag{113}$$

We want to solve for $\vec{x}$ as a function of our current guess $\vec{x}^*$, which means we want to solve

$$\frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}^*)(\vec{x} - \vec{x}^*) = -\vec{f}(\vec{x}^*) \tag{114}$$

For our objective, we want

$$\vec{f}(\vec{x}) = \begin{bmatrix} \frac{\partial c(\lambda_j, b_j)}{\partial \lambda} \\ \frac{\partial c(\lambda_j, b_j)}{\partial b} \end{bmatrix} = \vec{0} \tag{115}$$

Then during each Newton's iteration, we want to solve the linear system of equations

$$\begin{bmatrix} \frac{\partial^2 c}{\partial\lambda\partial\lambda}(\lambda_j, b_j) & \frac{\partial^2 c}{\partial\lambda\partial b}(\lambda_j, b_j) \\ \frac{\partial^2 c}{\partial b\partial\lambda}(\lambda_j, b_j) & \frac{\partial^2 c}{\partial b\partial b}(\lambda_j, b_j) \end{bmatrix} \begin{bmatrix} \lambda - \lambda_j \\ b - b_j \end{bmatrix} = \begin{bmatrix} -\frac{\partial c}{\partial\lambda}(\lambda_j, b_j) \\ -\frac{\partial c}{\partial b}(\lambda_j, b_j) \end{bmatrix} \tag{116}$$

$$\frac{\partial^2 c}{\partial\vec{x}\partial\vec{x}}(\lambda_j, b_j) \begin{bmatrix} \lambda - \lambda_j \\ b - b_j \end{bmatrix} = -\frac{\partial c}{\partial\vec{x}}^\top(\lambda_j, b_j) \tag{117}$$

Thus the matrix A is the Hessian of $c(\lambda, b)$ evaluated at λ_j, b_j, and the vector $\vec{y}$ is the transpose of the derivative of $c(\lambda, b)$ evaluated at λ_j, b_j. You don't need to see this connection and full credit is given if you just write out the partial derivative entries of A and $\vec{y}$.

The fact that you now have the tools to be able to do this problem is a testament to the awesome power of what you have been taught in 16AB. Even to do system-identification in continuous-time for a linear system, we need to be able to deal with fitting nonlinear equations. But once we have the ability to do this kind of fitting, we can actually fit even more sophisticated nonlinear differential equations. The advent of convenient tools for automatic differentiation like PyTorch have made it practically true in many cases that if we can simulate a model, we can also fit that model to data.

13. Movie Ratings with Missing Entries (23 points)

In a matrix R, you have users' movie ratings. However, not all users watched all the movies.

$$R = \left[\begin{array}{cccc|cc} 0.50 & 0.00 & 0.50 & 0.50 & 0.20 & 1.0 \\ 0.60 & 0.20 & 0.40 & 0.50 & ? & ? \\ 0.50 & 0.50 & 0.00 & 0.25 & 0.60 & 1.0 \\ 0.60 & 0.10 & 0.50 & 0.55 & ? & ? \\ \hline 1.00 & 0.40 & 0.60 & ? & ? & ? \end{array}\right] \quad (118)$$

where the element at the ith row and jth column indicates the rating of movie i by user j. A "?" means that there's no rating for that movie.

Our goal is to predict ratings for the missing entries, so we can recommend movies to users. In order to do this, you want to find the hidden goodness vectors for the movies, and the hidden sensitivity vectors of the users. However, due to missing entries, it is not possible to run an SVD on the entire rating matrix R.

It turns out that we have a submatrix R' in R that does not have any missing entries.

$$R' = \begin{bmatrix} 0.50 & 0.00 & 0.50 & 0.50 \\ 0.60 & 0.20 & 0.40 & 0.50 \\ 0.50 & 0.50 & 0.00 & 0.25 \\ 0.60 & 0.10 & 0.50 & 0.55 \end{bmatrix} \quad (119)$$

We provide a decomposition of this matrix:

$$R' = \underbrace{\begin{bmatrix} 0.5 & 0.0 \\ 0.4 & 0.2 \\ 0.0 & 0.5 \\ 0.5 & 0.1 \end{bmatrix}}_{G} \underbrace{\begin{bmatrix} 4.0 & 0.0 \\ 0.0 & 2.0 \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} 0.25 & 0.0 & 0.25 & 0.25 \\ 0.5 & 0.5 & 0.0 & 0.25 \end{bmatrix}}_{S} \quad (120)$$

where the ith row of the matrix G represents the goodness row vector $\vec{g}_i^\top$ of the movie i, the jth column of the matrix S represents the sensitivity vector $\vec{s}_j$ of user j, and each diagonal entry of the matrix D shows the weight each attribute has in determining the rating of a movie by a user.

NOTE: This decomposition in eq. (120) is not an SVD; G and S do not have orthonormal vectors.

Solution: This problem largely follows the modeling spirit of HW 11's problem "Movie Ratings and PCA." The major shift is that instead of using the SVD to get the low-rank approximations, we are simply providing those to you differently so that the orthogonality-properties of the SVD bases cannot be used. In that sense, we are connecting to ideas you saw in Discussion 9B.

(a) (4pts) Suppose $\vec{s}_6$ (the sensitivity vector of user 6) is:

$$\vec{s}_6 = \begin{bmatrix} 0.5 \\ 1.0 \end{bmatrix} \quad (121)$$

Use this to **estimate the rating of movie 2 as rated by user 6.** (Show work that uses $\vec{s}_6$. Unjustified answers will get no credit.)

Solution: This rating is in R_{26}. From the structure of R', if we generalize the sensitivity model, then this can be calculated with $\vec{g}_2^\top D\vec{s}_6$.

$$R_{26} = \vec{g}_2^\top D\vec{s}_6 = d_1 s_{61} g_{21} + d_2 s_{62} g_{22} \quad (122)$$

$$= 4 \cdot 0.5 \cdot 0.4 + 2 \cdot 1.0 \cdot 0.2 \quad (123)$$

$$= 1.2 \quad (124)$$

where d_i is ith diagonal element of matrix D.

(b) (6pts) For the 5th movie, we have three ratings and want to find two parameters of goodness. **Formulate a least squares problem $A\vec{g}_5 \approx \vec{b}$ to estimate $\vec{g}_5$ (goodness vector of movie 5).** You need to tell us A explicitly as a matrix with numerical entries. We give you that $\vec{b} = \begin{bmatrix} 1.00 \\ 0.40 \\ 0.60 \end{bmatrix}$ since those are the three ratings we know for this movie.

Solution: Considering the $R_{51}, \ldots, R_{54}$ in the 5th row of the ratings matrix R, we have

$$\begin{bmatrix} g_{51} & g_{52} \end{bmatrix} \underbrace{\begin{bmatrix} 4.0 & 0.0 \\ 0.0 & 2.0 \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} 0.25 & 0.0 & 0.25 & 0.25 \\ 0.5 & 0.5 & 0.0 & 0.25 \end{bmatrix}}_{S} = \begin{bmatrix} 1.00 & 0.40 & 0.60 & ? \end{bmatrix} \tag{125}$$

Removing the 4th column of S (since there is no corresponding equation) and transposing each side,

$$\underbrace{\begin{bmatrix} 0.25 & 0.5 \\ 0.0 & 0.5 \\ 0.25 & 0.0 \end{bmatrix}}_{S_{1:3}^\top} \underbrace{\begin{bmatrix} 4.0 & 0.0 \\ 0.0 & 2.0 \end{bmatrix}}_{D} \begin{bmatrix} g_{51} \\ g_{52} \end{bmatrix} = \begin{bmatrix} 1.00 \\ 0.40 \\ 0.60 \end{bmatrix} \tag{126}$$

Setting $A = S_{1:3}^\top D$ and $\vec{b}$ to be the ratings vector, we have a least squares problem

$$\min_{\vec{x}} \|A\vec{x} - \vec{b}\|^2 \tag{127}$$

where

$$A = \begin{bmatrix} 1.0 & 1.0 \\ 0.0 & 1.0 \\ 1.0 & 0.0 \end{bmatrix}, \vec{x} = \begin{bmatrix} g_{51} \\ g_{52} \end{bmatrix}, \vec{b} = \begin{bmatrix} 1.00 \\ 0.40 \\ 0.60 \end{bmatrix} \tag{128}$$

(c) (3pts) We now consider a ratings matrix R without missing entries (that is different from the previous

R) where the matrix is partitioned into four blocks $R_{11}, R_{12}, R_{21}, R_{22}$ as below.

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \tag{129}$$

In order to find the optimal number of principal components, we compute a PCA model from the SVD of R_{11} with k principal components, with $k = 2, 3, 4, 5$. We then use the chosen components and the singular values of R_{11} together with the information in R_{12} and R_{21} to create an estimate $\widehat{R}_{22}$ for the held-out ratings in R_{22}. We can also use the first k terms of the SVD of R_{11} to reconstruct $\widehat{R}_{11}$ as the best rank-k approximation to R_{11}.

Solution: This part of the problem builds on the ideas you saw in both lecture and exercised for yourself in HW 10's problem "Orthonormalization for Speeding Up Model Order Selection." Only the context has changed, but the spirit of doing model order selection by minimizing the validation error is the same.

The training errors $\|R_{11} - \widehat{R}_{11}\|_F^2$ and validation errors $\|R_{22} - \widehat{R}_{22}\|_F^2$ for each candidate choice for k are given in the table below.

Select	k	Training error	Validation error
◯	1	1.428	3.104
◯	2	0.414	2.494
◯	3	0.093	0.462
◯	4	0.017	0.090
◯	5	0.011	0.132
◯	6	0.006	0.161

Find the optimal number of principal components k we should use and fill in the appropriate bubble. (No need to give any justification.)

Solution: As we increase the number of principal components k, we will observe training error decreasing, while validation error first decreases but increases at some point. The optimal k can be found by selecting k which gives minimum validation error, which is $k = 4$.

(d) (10pts) Suppose that we want to approximate R with a rank-k matrix $\widehat{R} = CXL^\top$ where C is known (e.g. it has a specific k columns selected from R), and so is $L^\top$ (e.g. it has a specific k rows selected from R). The only freedom is in choosing the k by k matrix X.

We want to find X that minimizes the Frobenius norm error between R and $\widehat{R}$:

$$\underset{X}{\operatorname{argmin}} \|R - \underbrace{CXL^\top}_{\widehat{R}}\|_F^2 \tag{130}$$

This is a least-squares problem since $CXL^\top$ is linear in the entries of the matrix X and minimizing the Frobenius norm squared is just minimizing a sum of squared errors. Suppose we further know that C has linearly independent columns and that $L^\top$ has linearly independent rows. It turns out that the optimal $X = \left((C^\top C)^{-1}C^\top\right) R \left(L(L^\top L)^{-1}\right)$.

Solution: This part of this problem has a direct spiritual lineage to HW 14's problem "Minimum Norm Variants" in its emphasis of understanding an optimization problem by looking at things in the appropriate SVD bases. The underlying ideas on the nature of projections have been seen in multiple discussions including 9B.

Suppose that we know the full SVDs of C and $L^\top$:

$$C = \begin{bmatrix} U_C & U_{C,n} \end{bmatrix} \begin{bmatrix} \Sigma_C \\ 0 \end{bmatrix} V_C^\top, \qquad L^\top = U_L \begin{bmatrix} \Sigma_L & 0 \end{bmatrix} \begin{bmatrix} V_L^\top \\ V_{L,n}^\top \end{bmatrix}. \tag{131}$$

Using these SVDs and remembering how they simplify projections, we notice that:

$$\widehat{R} = CXL^\top = \big(C(C^\top C)^{-1}C^\top\big)R\big(L(L^\top L)^{-1}L^\top\big) = U_C U_C^\top R V_L V_L^\top. \tag{132}$$

This suggests that the orthonormal bases $U = [U_C \quad U_{C,n}]$ and $V = [V_L \quad V_{L,n}]$ are interesting to consider, so we notice that

$$R = UU^\top RVV^\top \tag{133}$$

$$= \begin{bmatrix} U_C & U_{C,n} \end{bmatrix} \begin{bmatrix} U_C^\top \\ U_{C,n}^\top \end{bmatrix} R \begin{bmatrix} V_L & V_{L,n} \end{bmatrix} \begin{bmatrix} V_L^\top \\ V_{L,n}^\top \end{bmatrix} \tag{134}$$

$$= (U_C U_C^\top + U_{C,n}U_{C,n}^\top)R(V_L V_L^\top + V_{L,n}V_{L,n}^\top) \tag{135}$$

$$= \underbrace{U_C U_C^\top R V_L V_L^\top}_{\widehat{R}} + U_C U_C^\top R V_{L,n} V_{L,n}^\top + U_{C,n}U_{C,n}^\top R V_L V_L^\top + U_{C,n}U_{C,n}^\top R V_{L,n} V_{L,n}^\top. \tag{136}$$

Use eq. (136) together with eq. (132) to establish that this X satisfies the key condition of least-squares optimality: **show that the residual $R - \widehat{R}$ is orthogonal to the estimate $\widehat{R}$** when you use the inner product corresponding to the Frobenius norm — which basically treats a matrix as a big vector. $\langle A, B\rangle_F = \text{trace}(A^\top B) = \text{trace}(B^\top A) = \text{trace}(AB^\top) = \text{trace}(BA^\top)$.

(HINT 1: Given that the Frobenius inner-product between two matrices of the same size can be interpreted as either the sum of the inner-products of all the rows in A with their B counterparts (trace$(AB^\top)$) or all the columns in A with their B counterparts (trace$(A^\top B)$), why do you think that the first term in eq. (136) must be orthogonal to each of the final three terms in eq. (136)?)

(HINT 2: What subspace do the rows of the second term in eq. (136) live in? What subspace do the columns of the third term in eq. (136) live in? .... What subspaces do the rows and columns of $\widehat{R}$ live in according to eq. (132)?)

Solution:
We prove the orthogonality of residual $R - CXL^\top$ and the estimate $CXL^\top$.
Expanding each term, we have:

$$CXL^\top = U_C U_C^\top R V_L V_L^\top \tag{137}$$
$$R - CXL^\top = U_C U_C^\top R V_{L,n} V_{L,n}^\top + U_{C,n} U_{C,n}^\top R V_L V_L^\top + U_{C,n} U_{C,n}^\top R V_{L,n} V_{L,n}^\top \tag{138}$$

From $\text{Col}(U_C) \perp \text{Col}\big(U_{C,n}\big)$ and $\text{Row}\left(V_L^\top\right) \perp \text{Row}\left(V_{L,n}^\top\right)$, and the definition of matrix multiplication (i.e. AB has columns that are linear combinations of the columns of A and has rows that are linear combinations of the rows of B), we find

$$\text{Row}\Big((U_C U_C^\top R V_L) V_L^\top\Big) \perp \text{Row}\Big((U_C U_C^\top R V_{L,n}) V_{L,n}^\top\Big) \tag{139}$$
$$\text{Col}\Big(U_C (U_C^\top R V_L V_L^\top)\Big) \perp \text{Col}\Big(U_{C,n} (U_{C,n}^\top R V_L V_L^\top)\Big) \tag{140}$$
$$\text{Row}\Big((U_C U_C^\top R V_L) V_L^\top\Big) \perp \text{Row}\Big((U_{C,n} U_{C,n}^\top R V_{L,n}) V_{L,n}^\top\Big) \tag{141}$$

For any two matrix A, B which $\text{Row}(A) \perp \text{Row}(B)$, the inner product $\langle A, B\rangle_F = \text{trace}(AB^\top) = 0$ since A's rows are orthogonal to $B^\top$'s columns. Because of this orthogonality, the entire matrix $AB^\top$ is filled with zeros, so clearly the trace is also zero.
Similarly, for any two matrix A, B which $\text{Col}(A) \perp \text{Col}(B)$, the inner product $\langle A, B\rangle_F = \text{trace}(A^\top B) = 0$ since $A^\top$'s rows are orthogonal to B's columns.
Using this property, we find:

$$\langle U_C U_C^\top R V_L V_L^\top, U_C U_C^\top R V_{L,n} V_{L,n}^\top \rangle_F = 0 \tag{142}$$
$$\langle U_C U_C^\top R V_L V_L^\top, U_{C,n} U_{C,n}^\top R V_L V_L^\top \rangle_F = 0 \tag{143}$$
$$\langle U_C U_C^\top R V_L V_L^\top, U_{C,n} U_{C,n}^\top R V_{L,n} V_{L,n}^\top \rangle_F = 0 \tag{144}$$

Summing all three inner products yields $\langle CXL^\top, R - CXL^\top\rangle_F = 0$.
This concludes that the estimate $CXL^\top$ and the residual $R - CXL^\top$ is orthogonal, proving the optimitzer X of $\min_X \|R - CXL^\top\|_F^2$ is indeed

$$X = \left((C^\top C)^{-1} C^\top\right) R \left(L (L^\top L)^{-1}\right). \tag{145}$$

[Doodle page! Draw us something if you want or give us suggestions or complaints. You can also use this page to report anything suspicious that you might have noticed.

If needed, you can also use this space to work on problems. But if you want the work on this page to be graded, make sure you tell us on the problem's main page.]

Contributors:

- Michael Danielczuk.
- Kumar Krishna Agrawal.
- Ashwin Vangipuram.
- Ayan Biswas.
- Vladimir Stojanovic.
- Moses Won.
- Anant Sahai.
- Druv Pai.
- Jichan Chung.

