Final Exam Review

1. Potpourri (Spring 2021 Final)

(a)

$$\vec{v}_1 = \begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix}, \qquad \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \tag{1}$$

Run Gram-Schmidt on these vectors in this order (that is, start with $\vec{v}_1$ then $\vec{v}_2$), and extend this set to form an orthonormal basis for $\mathbb{R}^3$. Show your work.

Solution: We run the Gram-Schmidt process:

$$\vec{q}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{\vec{v}_1}{5} = \begin{bmatrix} 0 \\ 3/5 \\ 4/5 \end{bmatrix} \tag{2}$$

$$\vec{z}_2 = \vec{v}_2 - \langle \vec{v}_2, \vec{q}_1 \rangle \vec{q}_1 = \vec{v}_2 - \frac{3}{5}\vec{q}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 9/25 \\ 12/25 \end{bmatrix} = \begin{bmatrix} 0 \\ 16/25 \\ -12/25 \end{bmatrix} \tag{3}$$

$$\vec{q}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|} = \frac{\vec{v}_2}{4/5} = \begin{bmatrix} 0 \\ 4/5 \\ -3/5 \end{bmatrix} \tag{4}$$

To find a suitable final vector, we just note that $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ completes the orthonormal basis. We can also get this through Gram-Schmidt by adjoining the standard basis $\vec{v}_3 = \vec{e}_1$, $\vec{v}_4 = \vec{e}_2$, $\vec{v}_5 = \vec{e}_3$ (where $\vec{e}_i$ corresponds to each of our standard basis vectors) to our list:

$$\vec{z}_3 = \vec{v}_3 - \langle \vec{v}_3, \vec{q}_1 \rangle \vec{q}_1 - \langle \vec{v}_3, \vec{q}_2 \rangle \vec{q}_2 = \vec{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{5}$$

$$\implies \vec{q}_3 = \frac{\vec{z}_3}{\|\vec{z}_3\|} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{6}$$

$$\vec{z}_4 = \vec{v}_4 - \langle \vec{v}_4, \vec{q}_1 \rangle \vec{q}_1 - \langle \vec{v}_4, \vec{q}_2 \rangle \vec{q}_2 - \langle \vec{v}_4, \vec{q}_3 \rangle \vec{q}_3 = \vec{0} \tag{7}$$

$$\vec{z}_5 = \vec{v}_5 - \langle \vec{v}_5, \vec{q}_1 \rangle \vec{q}_1 - \langle \vec{v}_5, \vec{q}_2 \rangle \vec{q}_2 - \langle \vec{v}_5, \vec{q}_3 \rangle \vec{q}_3 = \vec{0} \tag{8}$$

2. Analyzing an LC-LC Band-Stop/Notch Filter (Spring 2021 Midterm)

In this sub-part, you will partially analyze a circuit built entirely out of L, C components as shown in Figure 1. Assume that the circuit is operating at a frequency of $\omega = \omega_s$ (i.e. $v_{\text{in}}(t) = \cos(\omega_s t)$).

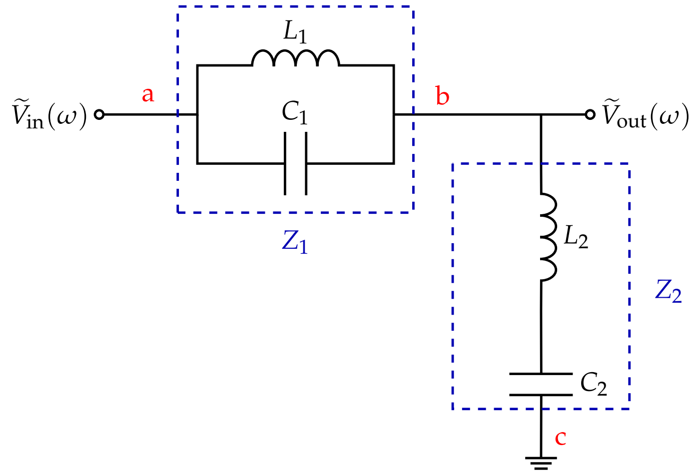
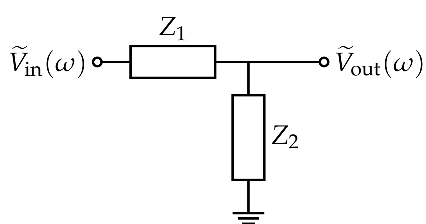


Figure 1: LC bandstop filter.

(a) **Find $\widetilde{V}_{\text{out}}(\omega)$ in terms of Z_1, Z_2, $\widetilde{V}_{\text{in}}(\omega)$.** *You do not need to compute $\widetilde{V}_{\text{in}}(\omega)$ for this part.* Show your work.

Solution: The equivalent circuit we get is as follows:



Since we are in the phasor domain, these impedances can be treated as resistors, and our output voltage phasor can be found as a function of the input voltage phasor by applying the voltage-divider equation. Doing so, we find that:

$$\widetilde{V}_{\text{out}}(\omega) = \frac{Z_2}{Z_1 + Z_2} \widetilde{V}_{\text{in}}(\omega) \tag{9}$$

(b) **Find Z_1, the equivalent impedance between terminals a and b, in terms of L_1, C_1, and ω_s.** Leave your answer in the form $j\frac{M}{N}$, where M and N are real.

What is the impedance Z_1 at $\omega_s = \frac{1}{\sqrt{L_1 C_1}}$

Solution: The impedances are in parallel. Using that $Z_L = \mathrm{j}\omega_s L$ and $Z_C = \frac{1}{\mathrm{j}\omega_s C}$ (at the frequency ω_s of the input phasor), we find:

$$Z_1 = Z_{L_1} \parallel Z_{C_1} \tag{10}$$

$$= \frac{Z_{L_1} Z_{C_1}}{Z_{L_1} + Z_{C_1}} \tag{11}$$

$$= \frac{j\omega_s L_1 \frac{1}{j\omega_s C_1}}{j\omega_s L_1 + \frac{1}{j\omega_s C_1}} \quad (12)$$

$$= \frac{j\omega_s L_1}{\left(j\omega_s L_1 \frac{1}{j\omega_s C_1}\right) + 1} \quad (13)$$

$$= \frac{j\omega_s L_1}{1 - \omega_s^2 L_1 C_1} \quad (14)$$

At $\omega_s = \frac{1}{\sqrt{L_1 C_1}}$:

$$Z_1 = \frac{j\omega_s L_1}{1 - \omega_s^2 L_1 C_1} \quad (15)$$

$$Z_1 = \frac{j\omega_s L_1}{0} \quad (16)$$

$$Z_1 = \infty \quad (17)$$

(c) **Find Z_2, the equivalent impedance between terminals b and c, in terms of L_2, C_2, and ω_s.** Leave your answer in the form $j\frac{M}{N}$, where M and N are real.

What is the impedance Z_2 at $\omega_s = \frac{1}{\sqrt{L_2 C_2}}$?

Solution: Simplifying the series LC combination, we find (since $\frac{1}{j} = -j$):

$$Z_2 = Z_{L_2} + Z_{C_2} \quad (18)$$

$$= j\omega_s L_2 + \frac{1}{j\omega_s C_2} \quad (19)$$

$$= j\left(\omega_s L_2 - \frac{1}{\omega_s C_2}\right) \quad (20)$$

$$= j\left(\frac{\omega_s^2 L_2 C_2 - 1}{\omega_s C_2}\right) \quad (21)$$

At $\omega_s = \frac{1}{\sqrt{L_2 C_2}}$:

$$Z_2 = j\left(\frac{L_2}{\sqrt{L_2 C_2}} - \frac{\sqrt{L_2 C_2}}{C_2}\right) \quad (22)$$

$$= j\left(\sqrt{\frac{L_2}{C_2}} - \sqrt{\frac{L_2}{C_2}}\right) \quad (23)$$

$$= 0 \quad (24)$$

3. Control (Fall 2019 Midterm 2)

Suppose that we have a two-dimensional discrete-time system governed by:

$$\vec{x}(t+1) = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{7}{6} \end{bmatrix} \vec{x}(t) + \vec{w}(t) \tag{25}$$

(a) **Is the system stable? Why or why not?**

Here, we give you that

$$A = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{7}{6} \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} -\frac{4}{3} & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \tag{26}$$

and that the characteristic polynomial $\det(\lambda I - A) = \lambda^2 + \frac{11}{6}\lambda + \frac{2}{3}$.

Solution: The eigenvalues are $\lambda_1 = -\frac{4}{3}$ and $\lambda_2 = -\frac{1}{2}$. The system is unstable because we have an eigenvalue with absolute value greater than 1.

(b) Suppose that there is no disturbance and we can now influence the system using a scalar input $u(t)$ to our system:

$$\vec{x}(t+1) = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{7}{6} \end{bmatrix} \vec{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t) \tag{27}$$

Is the system controllable?

Solution: Yes, and we can see this by writing out the controllability matrix

$$\mathcal{C} = \begin{bmatrix} \vec{b} & A\vec{b} \end{bmatrix} = \begin{bmatrix} 1 & -\frac{2}{3} \\ 0 & \frac{1}{3} \end{bmatrix} \tag{28}$$

The columns are linearly independent, so $\mathcal{C}$ is full rank and the system is controllable.

Note: the solutions for this and subsequent problems define $\mathcal{C} = \begin{bmatrix} \vec{b} & A\vec{b} \end{bmatrix}$ whereas, in this iteration of the course, we have defined it as $\mathcal{C} = \begin{bmatrix} A\vec{b} & \vec{b} \end{bmatrix}$. You can choose either way to write the controllability matrix, as long as you maintain consistency, especially in the context of deriving CCF transforms. Furthermore, note that permuting the columns of a matrix won't change its rank.

(c) We want to set the closed-loop eigenvalues of the system

$$\vec{x}(t+1) = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{7}{6} \end{bmatrix} \vec{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t) \tag{29}$$

to be $\lambda_1 = -\frac{5}{6}, \lambda_2 = \frac{5}{6}$ using state feedback

$$u(t) = -\begin{bmatrix} k_1 & k_2 \end{bmatrix} \vec{x}(t) \tag{30}$$

What specific numeric values of k_1 and k_2 should we use?

Solution: The new system matrix is

$$\widetilde{A} = \begin{bmatrix} -\frac{2}{3} - k_1 & \frac{1}{3} - k_2 \\ \frac{1}{3} & -\frac{7}{6} \end{bmatrix} \tag{31}$$

and we find the characteristic polynomial

$$\left(\lambda + \frac{7}{6}\right) \cdot \left(\lambda + \frac{2}{3} + k_1\right) - \frac{1}{9} + \frac{1}{3}k_2 = \lambda^2 + \left(\frac{11}{6} + k_1\right)\lambda + \frac{2}{3} + \frac{7}{6}k_1 + \frac{1}{3}k_2 \tag{32}$$

Since we want $\lambda_1 = -\frac{5}{6}, \lambda_2 = \frac{5}{6}$, we would want the characteristic polynomial to be of the form

$$\lambda^2 - \frac{25}{36} = 0 \tag{33}$$

Comparing the terms, we want

$$\frac{2}{3} + \frac{7}{6}k_1 + \frac{1}{3}k_2 = -\frac{25}{36} \tag{34}$$

$$\frac{11}{6} + k_1 = 0 \tag{35}$$

and we solve

$$k_1 = -\frac{11}{6} \tag{36}$$

$$k_2 = \frac{7}{3} \tag{37}$$

4. Controllable Canonical Form (Fall 2019 Final)

Suppose that we have a two-dimensional continuous time system governed by:

$$\frac{d}{dt}\vec{x}(t) = \begin{bmatrix} 1 & -1 \\ 0 & -4 \end{bmatrix} \vec{x}(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(t) \tag{38}$$

We would like to put this system into Controllable Canonical Form (CCF) to use state feedback to place the eigenvalues at desired locations. For your convenience, the characteristic polynomial is

$$\det\left(\lambda I - \begin{bmatrix} 1 & -1 \\ 0 & -4 \end{bmatrix}\right) = (\lambda - 1)(\lambda + 4) = \lambda^2 + 3\lambda - 4 \tag{39}$$

(a) **Is the system stable? Why or why not?**

Solution: The system is not stable because it has an eigenvalue $\lambda = +1$ that has a non-negative real part. The system will tend to explode along that direction.

The $\lambda = -4$ eigenvalue is perfectly stable by contrast.

(b) Recall that our original system is:

$$\frac{d}{dt}\vec{x}(t) = \begin{bmatrix} 1 & -1 \\ 0 & -4 \end{bmatrix} \vec{x}(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(t) \tag{40}$$

We would like to change the coordinates to bring the system into CCF:

$$\frac{d}{dt}\vec{x}(t) = \begin{bmatrix} 0 & 1 \\ a_0 & a_1 \end{bmatrix} \vec{x}(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) \tag{41}$$

Compute the T basis such that $\vec{z}(t) = T^{-1}\vec{x}(t)$ or equivalently (if you want), give us T^{-1} directly. To help you along, here are some calculations already done for you:

$$\begin{bmatrix} 1 & 0 \\ 1 & -4 \end{bmatrix}^{-1} = \frac{1}{4}\begin{bmatrix} 4 & 0 \\ 1 & -1 \end{bmatrix} \tag{42}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & a \end{bmatrix}^{-1} = \begin{bmatrix} -a & 1 \\ 1 & 0 \end{bmatrix} \tag{43}$$

What is a_0? What is a_1?

Solution: The characteristic polynomial of $\begin{bmatrix} 0 & 1 \\ a_0 & a_1 \end{bmatrix}$ is $\lambda^2 - a_1\lambda - a_0$ and since changing coordinates doesn't change the characteristic polynomial, we know that $a_0 = 4$ and $a_1 = -3$.

Now, we can compute the controllability matrix for the original system $C = \begin{bmatrix} 1 & 0 \\ 1 & -4 \end{bmatrix}$ as well as for the system that is already in CCF: $C_z = \begin{bmatrix} 0 & 1 \\ 1 & -3 \end{bmatrix}$.

At this point, we know that the original system's A matrix in the C basis looks like the transpose of A_z. Since we also know that C_z has the same basic relationship, we know that $T = CC_z^{-1}$ is the basis in which the original system's matrix will look like A_z. Computing this:

$$T = CC_z^{-1} \tag{44}$$

$$= \begin{bmatrix} 1 & 0 \\ 1 & -4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -3 \end{bmatrix}^{-1} \tag{45}$$

$$= \begin{bmatrix} 1 & 0 \\ 1 & -4 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix} \tag{46}$$

$$= \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix} \tag{47}$$

We could also compute T^{-1} if we wanted by the same essential reasoning:

$$T^{-1} = C_z C^{-1} \tag{48}$$

$$= \begin{bmatrix} 0 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & -4 \end{bmatrix}^{-1} \tag{49}$$

$$= \begin{bmatrix} 0 & 1 \\ 1 & -3 \end{bmatrix} \cdot \frac{1}{4} \begin{bmatrix} 4 & 0 \\ 1 & -1 \end{bmatrix} \tag{50}$$

$$= \frac{1}{4} \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix} \tag{51}$$

If we wanted, we could check our work by computing

$$T^{-1} \begin{bmatrix} 1 & -1 \\ 0 & -4 \end{bmatrix} T = \frac{1}{4} \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & -4 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix} \tag{52}$$

$$= \frac{1}{4} \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 4 & -4 \end{bmatrix} \tag{53}$$

$$= \begin{bmatrix} 0 & 1 \\ 4 & -3 \end{bmatrix} \tag{54}$$

And so, this checks out as far as the system matrix goes. We can also see what happens with $T^{-1}\vec{b} = \frac{1}{4} \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ which also checks out.

Finally, it is important to note that we can also compute the a_0, a_1 by computing

$$C^{-1}A^2\vec{b} = \frac{1}{4} \begin{bmatrix} 4 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ 16 \end{bmatrix} \tag{55}$$

$$= \begin{bmatrix} 4 \\ -3 \end{bmatrix} \tag{56}$$

That also works.

(c) Using state feedback

$$u(t) = \begin{bmatrix} \widetilde{k}_0 & \widetilde{k}_1 \end{bmatrix} \vec{z}(t) \tag{57}$$

place the closed-loop eigenvalues at $\lambda_1 = -1, \lambda_2 = -2$. **What is $\widetilde{k}_0$? What is $\widetilde{k}_1$?**

(Notice that we are asking for the feedback gains in terms of $\vec{z}(t)$ not the original $\vec{x}(t)$. If you have time, feel free to check your work by the original $\vec{x}(t)$.)

Solution: The key here is to realize what we want the new characteristic polynomial to be $(\lambda - \lambda_1)(\lambda - \lambda_2) = (\lambda + 1)(\lambda + 2) = \lambda^2 + 3\lambda + 2$. In CCF form, the coefficients of the characteristic polynomial are visible on the bottom row. So we need $4 + \widetilde{k}_0 = -2$ which means that $\widetilde{k}_0 = -6$. We also need $-3 + \widetilde{k}_1 = -3$ which means that $\widetilde{k}_1 = 0$.

To check our work, we realize that $\begin{bmatrix} -6 & 0 \end{bmatrix} T^{-1} = \begin{bmatrix} -6 & 0 \end{bmatrix} \frac{1}{4} \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} & \frac{3}{2} \end{bmatrix}$ are the claimed gains in the original coordinates. This makes the original coordinate closed-loop matrix be

$$\begin{aligned} \begin{bmatrix} 1 & -1 \\ 0 & -4 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} -\frac{3}{2} & \frac{3}{2} \end{bmatrix} &= \begin{bmatrix} 1 - \frac{3}{2} & -1 + \frac{3}{2} \\ 0 - \frac{3}{2} & -4 + \frac{3}{2} \end{bmatrix} && (58) \\ &= \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ -\frac{3}{2} & -\frac{5}{2} \end{bmatrix} && (59) \end{aligned}$$

which has eigenvalues at -2 and -1 as desired since $(\lambda + \frac{1}{2})(\lambda + \frac{5}{2}) + \frac{3}{4} = \lambda^2 + 3\lambda + \frac{5}{4} + \frac{3}{4} = \lambda^2 + 3\lambda + 2$.

5. I bet Cal will win this year (Spring 2021 Final)

As huge fans of the Big Game, you and your friend want to bet on whether Cal or Stanford will win this year. You want to predict this year's result by analyzing historical records. Therefore, you decide to model this as a binary classification problem and do PCA for dimension reduction on the data you collected. The "+1" class represents victories of Cal and "−1" represents victories of Stanford.

After some research, you obtained a data matrix $A \in \mathbb{R}^{n \times d}$,

$$A = \begin{bmatrix} - & \vec{x}_1^\top & - \\ - & \vec{x}_2^\top & - \\ & \vdots & \\ - & \vec{x}_n^\top & - \end{bmatrix} \tag{60}$$

where each of the n rows $\vec{x}_i^\top$ denotes a game and each of the d columns of A contains information of a possibly relevant factor of the games (weather, location, date, air quality, etc).

(a) Let the full SVD of $A = U\Sigma V^\top$, where A is given in eq. (60).

You project your data along $\vec{v}_1$ and $\vec{v}_2$ (the first two principal components along the rows), and for comparison you also project your data along two randomly chosen directions $\vec{w}_1$ and $\vec{w}_2$ as well. You get the two pictures in Figure 2, but you forgot to label the axes. Of the two figures below, which one is the projection onto the principal components and which one is the projection onto the random directions? **Match axes (i), (ii), (iii), (iv) to $\vec{w}_1, \vec{w}_2, \vec{v}_1$, and $\vec{v}_2$, and justify your answer.**

Note that there may be multiple correct matchings; you only need to find and justify one of them.

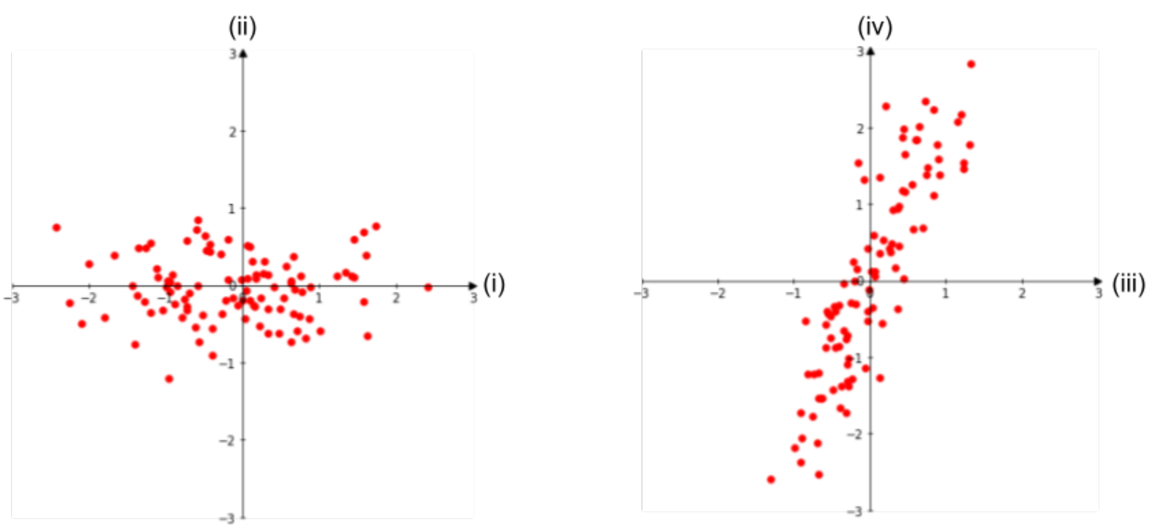


Figure 2: Projected datasets.

Solution: You may recall from the notebook on neuron classification — when we project data along principal components the data is aligned to orthogonal axes where as when we projected it along random components it was not aligned to any axes. So we can deduce from this that (i) and (ii) must correspond in some order to the principal components and (iii) and (iv) must correspond to the random directions. Further observation leads us to see that axis (ii) has more spread than axis (i) and therefore must correspond to the larger singular value, i.e. to the first principal component $\vec{v}_1$.

i. $\vec{v}_1$ – This is the axis with the maximal spread of the data and therefore must correspond to the largest singular value. the single axis, across both plots, across which there is maximal spread of the data.

ii. $\vec{v}_2$ – the corresponding axis to $\vec{v}_1$ and also an axis for which the spread of the data is axis-aligned.

iii. $\vec{w}_1$ or $\vec{w}_2$ – seemingly a random projection. We don't know which one is $\vec{w}_1$ and $\vec{w}_2$ since they are random unit vectors and as such are independent of the data, so we can't tell from the plot.

iv. $\vec{w}_2$ or $\vec{w}_1$ – seemingly a random projection. We don't know which one is $\vec{w}_1$ and $\vec{w}_2$ since they are random unit vectors and as such are independent of the data, so we can't tell from the plot.

(b) In order to reduce the dimension of the data, we would like to project the data onto the first k principal components along the rows of A, where k is less than the original data dimension d. **Show how to find the new coordinates $\vec{z}_i$ of the data point $\vec{x}_i$ after this projection.** You may use the SVD of A.

Solution: Let

$$V_k = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_k \end{bmatrix} \tag{61}$$

Since we are projecting onto the columns of V_k, the new coordinate of $\vec{x}_i$ after dimension reduction is $\vec{z}_i = (V_k^\top V_k)^{-1} V_k^\top \vec{x}_i = V_k^\top \vec{x}_i$. Note that $V_k^\top V_k = I_k$ since V_k has orthonormal columns. Note also that this is equivalent to saying $\vec{z}_i$ is obtained by taking the first k entries of $V^\top \vec{x}_i$.

(c) Using the data you have, you trained a classifier $\vec{w}_\star$. For any new data point after dimension reduction $\vec{z}_{\text{new}}$, the value of $\text{sign}(\vec{w}_\star^\top \vec{z}_{\text{new}})$ tells you whether the data point belongs to the "+1" class or to the "−1" class. Now suppose you have obtained two new data points, $\vec{z}_a$ and $\vec{z}_b$. Based on Figure 3 showing $\vec{w}_\star$, $\vec{z}_a$ and $\vec{z}_b$, **predict the class of $\vec{z}_a$ and $\vec{z}_b$ using $\vec{w}_\star$, and justify your answer.**

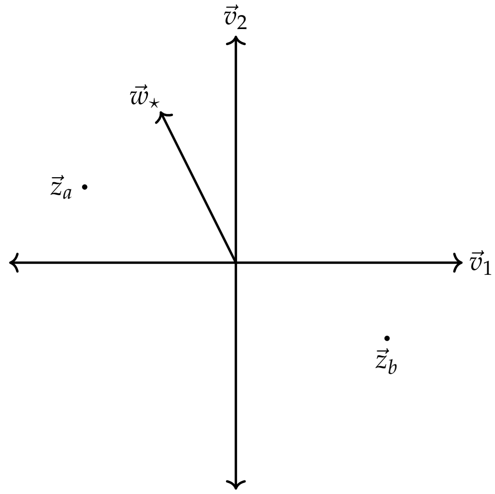


Figure 3: Dataset projected onto $\vec{v}_1$ and $\vec{v}_2$ with $\vec{w}_\star$

Solution: Since we are classifying based on $\text{sign}(\vec{w}_\star^\top \vec{z}_{\text{new}})$, from the graph we can see that $\vec{w}_\star^\top \vec{z}_a > 0$, thus $\vec{z}_a$ is predicted to be in class "+1". Similarly, $\vec{w}_\star^\top \vec{z}_b < 0$, and thus $\vec{z}_b$ is predicted to be in class "-1".

(d) Assume $d = 6$, $k = 4$, and $\vec{w}_\star = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}^\top$. Let $A = U\Sigma V^\top$ for A defined in eq. (60), and you find that V is given by the identity matrix, i.e. $V = I_d$. Now suppose the data point for this year's big game $\vec{x}_{2021} = \begin{bmatrix} 3 & 6 & 4 & 1 & 9 & 6 \end{bmatrix}^\top$. **Would you bet on Cal or Stanford to win? Justify your answer.** A quick reminder that "+1" denotes victories of Cal and "-1" denotes victories of Stanford.

(HINT: Don't forget to project your data onto the principal components.)

Solution: First, we need to preprocess this data point and project it onto the k-dimensional subspace just like what we did to the training points

$$\vec{z}_{2021} = V_k^\top \vec{x}_{2021} \tag{62}$$

$$= V_k^\top \begin{bmatrix} 3 \\ 6 \\ 4 \\ 1 \\ 9 \\ 6 \end{bmatrix} \tag{63}$$

$$= \begin{bmatrix} 3 \\ 6 \\ 4 \\ 1 \end{bmatrix} \tag{64}$$

Then, we compute the classifier's predicted value

$$p_{2021} = \vec{w}_\star^\top \vec{z}_{2021} \tag{65}$$

$$= \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 6 \\ 4 \\ 1 \end{bmatrix} \tag{66}$$

$$= 6 \tag{67}$$

Therefore, the classifier predicts the label for this data point to be "+1", thus you should bet for Cal to win this year!

6. Optimization and Singular Values (Spring 2021 Final)

We are going to focus on a special optimization problem that is related to the underlying structure of the SVD. More specifically, we want to solve for s in the following maximization problem

$$s = \max_{\|\vec{x}\| \neq 0} \frac{\|A\vec{x}\|^2}{\|\vec{x}\|^2} \tag{68}$$

Here, we have $A \in \mathbb{R}^{m \times n}$. Let $m > n$ so that A is a tall matrix and $\text{rank}(A) = n$. Let the full SVD be given by $A = U\Sigma V^\top$. Define $\vec{x}^\star \in \mathbb{R}^n$ to be the optimal vector that achieves the maximum in equation (68). That is,

$$\vec{x}^\star = \operatorname*{argmax}_{\|\vec{x}\| \neq 0} \frac{\|A\vec{x}\|^2}{\|\vec{x}\|^2} \tag{69}$$

$$s = \frac{\left\|A\vec{x}^\star\right\|^2}{\left\|\vec{x}^\star\right\|^2} \tag{70}$$

(a) We start by attempting to simplify the optimization problem. **Prove that for any $\vec{x}$, we have $\|A\vec{x}\| = \left\|\Sigma V^\top \vec{x}\right\|$.** Note that you must justify and explain every step for full credit, just equations without an explanation may not be awarded full credit.

Solution: We can directly plug in the SVD of A into the left hand side:

$$\|A\vec{x}\| = \left\|U\Sigma V^\top \vec{x}\right\| \tag{71}$$

$$= \left\|U(\Sigma V^\top \vec{x})\right\| \tag{72}$$

$$= \left\|\Sigma V^\top \vec{x}\right\| \tag{73}$$

We used the fact that U is an orthonormal matrix and therefore preserves the norm of any vector.

Alternative Solution:

Alternatively, we can use the inner product definition of the norm in order to simplify the expression. Recall that $\|\vec{x}\| = \sqrt{\langle \vec{x}, \vec{x} \rangle} = \sqrt{\vec{x}^\top \vec{x}}$. Using this fact, we can write out

$$\|A\vec{x}\| = \sqrt{(A\vec{x})^\top (A\vec{x})} \tag{74}$$

$$= \sqrt{(U\Sigma V^\top \vec{x})^\top (U\Sigma V^\top \vec{x})} \tag{75}$$

$$= \sqrt{\vec{x}^\top V \Sigma^\top (U^\top U) \Sigma V^\top \vec{x}} \tag{76}$$

We use the fact that U is an orthonormal matrix, implying that $U^\top U = I$. Substituting this back in gives us

$$\|A\vec{x}\| = \sqrt{\vec{x}^\top V \Sigma^\top \Sigma V^\top \vec{x}} \tag{77}$$

$$= \sqrt{(\Sigma V^\top \vec{x})^\top (\Sigma V^\top \vec{x})} \tag{78}$$

$$= \left\|\Sigma V^\top \vec{x}\right\| \tag{79}$$

(b) Using a change of variables, we can in fact turn our original maximization problem into

$$s = \max_{\|\vec{w}\| \neq 0} \frac{\|\Sigma \vec{w}\|^2}{\|\vec{w}\|^2}. \tag{80}$$

Find the correct change of variables that relates $\vec{x}$ and $\vec{w}$ and show that optimization problems (68) and (80) are equivalent.

(HINT: The change of variables you are looking for can also be thought of as a change of basis.)

Solution: Plugging in the result from part (a) into the original maximization problem yields

$$s = \max_{\|\vec{x}\| \neq \vec{0}} \frac{\|\Sigma V^\top \vec{x}\|^2}{\|\vec{x}\|^2} \tag{81}$$

Seeing what the maximization problem transforms into, we see we want to choose $V^\top \vec{x} = \vec{w}$. This is since multiplication by $V^\top$ won't change the norm of a vector due to it being orthonormal. So, the problem becomes

$$s = \max_{\|V\vec{w}\| \neq \vec{0}} \frac{\|\Sigma \vec{w}\|^2}{\|V\vec{w}\|^2} \tag{82}$$

$$= \max_{\|\vec{w}\| \neq \vec{0}} \frac{\|\Sigma \vec{w}\|^2}{\|\vec{w}\|^2} \tag{83}$$

(c) **Let σ_1 be the largest singular value of matrix A. Find a $\vec{w}^\star$, such that $\|\Sigma \vec{w}^\star\|^2 = \sigma_1^2 \|\vec{w}^\star\|^2$. Justify your answer.**

Solution: We first define $\vec{e}_i$ (the i^{th} standard basis vector) as the vector with a 1 in the i^{th} entry and zero everywhere else (i.e. the i^{th} column of the identity matrix). If $\vec{w}^\star = c\vec{e}_1 = \begin{bmatrix} c & 0 & \dots & 0 \end{bmatrix}^\top$ for some constant c, then

$$\|\Sigma \vec{w}^\star\|^2 = \|\Sigma c \vec{e}_1\|^2 = \sigma_1^2 c^2 = \sigma_1^2 \|\vec{w}^\star\|^2 \tag{84}$$

Part (d) will tell us that for all $\vec{w}$, $\frac{\|\Sigma \vec{w}\|}{\|\vec{w}\|}^2 \leq \sigma_1^2$. This means that we have found the $\vec{w^\star}$ that achieves the upper bound. So, the value of s from (68) must be equal to σ_1^2.

(d) **Prove that for all $\vec{w}$ we have $\|\Sigma \vec{w}\|^2 \leq \sigma_1^2 \|\vec{w}\|^2$. Show your work.**

Solution: We first note that σ_1 is the largest singular value, so it is greater than or equal to all σ_i. We start by rewriting $\|\Sigma \vec{w}\|^2$ as a summation,

$$\|\Sigma \vec{w}\|^2 = \sum_{i=1}^{n} \sigma_i^2 w_i^2. \tag{85}$$

Next, we use the fact that σ_1 is greater than or equal to all of the σ_i and invoke the inequality,

$$\sum_{i=1}^{n} \sigma_i^2 w_i^2 \leq \sum_{i=1}^{n} \sigma_1^2 w_i^2. \tag{86}$$

Finally, we can pull out the common σ_1^2 from the summation and substitute $\sum_{i=1}^{n} w_i^2$ as the norm-squared of $\vec{w}$,

$$\sum_{i=1}^{n} \sigma_1^2 w_i^2 = \sigma_1^2 \sum_{i=1}^{n} w_i^2 = \sigma_1^2 \|\vec{w}\|^2, \tag{87}$$

as desired.

