1. PCA and LS Prediction

In this problem, you will be helping to analyze some clinical data for patients using PCA and least squares. Suppose the following is a data matrix consisting of scaled zero-mean patient data. The patients (i.e., data points) are the columns, while the rows from top to bottom represent scores for the height, weight, and blood pressure of the 4 patients.

\[
A = \begin{bmatrix}
-3 & 0 & 2 & 1 \\
-1 & -2 & 0 & 3 \\
-1 & -1 & 1 & 1
\end{bmatrix}
\]  

Using this information, we would like to predict a risk index for heart disease for each patient (negative if low-risk and positive if high risk). But unfortunately, there appears to be some noise mixed into the data. Let’s try to mitigate this by performing PCA to embed the training data into a lower dimension first.

(a) Let us perform PCA on \(A\) to embed the data into a single dimension. Suppose you know that \(\sigma_1 = \sqrt{8}\) and \(\vec{v}_1 = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}^T\) from the first element of \(\Sigma\) and first column of \(V\) for the SVD of \(A\). What is the first principal component \(\vec{u}_1\) of \(A\)? Show all of your work.

**Solution:** Using an intermediate step of solving for the SVD, we can find \(\vec{u}_1\) as follows:

\[
\vec{u}_1 = \frac{1}{\sigma_1} A \vec{v}_1
\]

\[
= \begin{bmatrix}
-3 & 0 & 2 & 1 \\
-1 & -2 & 0 & 3 \\
-1 & -1 & 1 & 1
\end{bmatrix}
\begin{bmatrix}
\frac{1}{2} \\
-\frac{1}{2} \\
-\frac{1}{2} \\
\frac{1}{2}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
-\frac{1}{\sqrt{2}} \\
\frac{1}{\sqrt{2}} \\
0
\end{bmatrix}
\]

(b) Regardless of your answer to the previous question, suppose \(\vec{u}_1 = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix}^T\). Project your three-dimensional data points from \(A\) into one-dimensional values. Express your answer as the vector \(\vec{z} \in \mathbb{R}^{1 \times 4}\), and show all your work.

**Solution:** Recall that we can use our principal component to project our 3D data into 1D by doing \(\vec{u}_1^T A\)

\[
\vec{u}_1^T A = \begin{bmatrix}
\frac{2}{3} & \frac{2}{3} & \frac{1}{3}
\end{bmatrix}
\begin{bmatrix}
-3 & 0 & 2 & 1 \\
-1 & -2 & 0 & 3 \\
-1 & -1 & 1 & 1
\end{bmatrix}
\]
(c) Now, suppose that for this training data we have known risk index scores already (with a scalar index for each patient):

\[ \vec{b} = \begin{bmatrix} -3 & -1 & 4 & 6 \end{bmatrix} \]  
(7)

We would like to be able to predict risk scores. Let’s try to use least squares on our scenario. In one dimension, we can set up our least squares problem using data \( \vec{d} = \vec{z}^T \) and target \( \vec{s} = \vec{b}^T \) to estimate a scalar parameter \( p_{LS} \). Perform least squares on the system \( \vec{d} = \vec{s} \) to estimate \( p_{LS} \), and show your work.

**Solution:** We perform least squares using the data points from the previous part. Thus, \( \vec{d} = \vec{z}^T = \begin{bmatrix} -3 \\ -\frac{5}{3} \\ \frac{5}{3} \\ 3 \end{bmatrix} \) is your data for least squares.

\[ p_{LS} = \left( \vec{d}^T \vec{d} \right)^{-1} \vec{d}^T \vec{s} \]  
(8)

\[ = \frac{3}{2} \]  
(9)
2. Movie Ratings and PCA

Recall from the lecture on PCA that we can think of movie ratings as a structured set of data. For every person $i$ and movie $j$, we have that person’s rating $R_{ij}$ (thought of as a real number).

Suppose that there are $m$ movies and $n$ people. Let’s think about arranging this data into a big $n \times m$ matrix $R$ with people corresponding to rows and movies corresponding to columns. So the entry in the $i$-th row and $j$-th column should be $R_{ij}$ above. Note that this is organized differently from how it was in lecture. Each row corresponds to a unique person and each column to a unique movie.

\[
R = \begin{bmatrix}
R_{11} & R_{12} & \cdots & R_{1m} \\
R_{21} & R_{22} & \cdots & R_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
R_{n1} & R_{n2} & \cdots & R_{nm}
\end{bmatrix}
\]  

(a) Suppose we believe that there is actually an underlying pattern to this rating data and that a separate study has revealed that every movie is characterized by a set of characteristics: say action and comedy. This means that every movie $j$ has a pair of numbers $a_j$ (for action) and $c_j$ (for comedy) that define it. At the same time, every person $i$ has a pair of sensitivities $f_i$ and $g_i$ that define their preferences for action vs. comedy in movies respectively. A person $i$ will rate the movie $j$ as $R_{ij} = f_i a_j + g_i c_j$.

If we arrange the sensitivities into a pair of $n$-dimensional vectors $\vec{f}, \vec{g}$ for our group of $n$ people, and the movie characteristics into a pair of $m$-dimensional vectors $\vec{a}, \vec{c}$ for our group of $m$ movies, use outer products to express the rating matrix $R$ in terms of these vectors $\vec{f}, \vec{g}, \vec{a}, \vec{c}$.

Solution:

\[
R = \begin{bmatrix}
R_{11} & R_{12} & \cdots & R_{1m} \\
R_{21} & R_{22} & \cdots & R_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
R_{n1} & R_{n2} & \cdots & R_{nm}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
f_1 a_1 + g_1 c_1 & f_1 a_2 + g_1 c_2 & \cdots & f_1 a_m + g_1 c_m \\
f_2 a_1 + g_2 c_1 & f_2 a_2 + g_2 c_2 & \cdots & f_2 a_m + g_2 c_m \\
\vdots & \vdots & \ddots & \vdots \\
f_n a_1 + g_n c_1 & f_n a_2 + g_n c_2 & \cdots & f_n a_m + g_n c_m
\end{bmatrix}
\]

\[
= \begin{bmatrix}
f_1 a_1 & f_1 a_2 & \cdots & f_1 a_m \\
f_2 a_1 & f_2 a_2 & \cdots & f_2 a_m \\
\vdots & \vdots & \ddots & \vdots \\
f_n a_1 & f_n a_2 & \cdots & f_n a_m
\end{bmatrix}
+ \begin{bmatrix}
g_1 c_1 & g_1 c_2 & \cdots & g_1 c_m \\
g_2 c_1 & g_2 c_2 & \cdots & g_2 c_m \\
\vdots & \vdots & \ddots & \vdots \\
g_n c_1 & g_n c_2 & \cdots & g_n c_m
\end{bmatrix}
\]

\[
= \begin{bmatrix}
f_1 \\
f_2 \\
\vdots \\
f_n
\end{bmatrix}
\begin{bmatrix}
a_1 & a_2 & \cdots & a_n
\end{bmatrix}
+ \begin{bmatrix}
g_1 \\
g_2 \\
\vdots \\
g_n
\end{bmatrix}
\begin{bmatrix}
c_1 & c_2 & \cdots & c_n
\end{bmatrix}
\]

\[
= \vec{f} \vec{a}^\top + \vec{g} \vec{c}^\top
\]
(b) Now suppose that we want to discover the underlying nature of movies from the data itself.

Suppose for this part that you have four observed rating data vectors (corresponding to four different movies being rated by six individuals).

All of the movie data vectors just happened to be multiples of the following 6-dimensional vector

\[
\vec{w} = \begin{bmatrix}
2 \\
-2 \\
3 \\
-4 \\
4 \\
0
\end{bmatrix}.
\]

(For your convenience, note that \( \|\vec{w}\| = 7 \).)

You arrange the movie data vectors as the columns of a matrix \( R \) given by:

\[
R = \begin{bmatrix}
\vec{w} \\
-\vec{w} \\
-2\vec{w} \\
2\vec{w} \\
4\vec{w}
\end{bmatrix}
\]

You want to perform PCA (for movies) using the SVD of the matrix \( R \) to better understand the pattern in your data.

The first “principal component vector” is a unit vector that tells which direction we would want to project the columns of \( R \) onto to get the best rank-1 approximation for \( R \).

**Find this first principal component vector of the columns of \( R \) to explain the nature of your movie data points.**

(HINT: You don’t need to actually compute any SVDs in this simple case. Also, be sure to think about what size vector you want as the answer. Don’t forget that you want a unit vector!)

**Solution:** Principal component analysis is in general about understanding how best to approximate our (potentially) high-dimensional data (like recordings from a microphone, or in this case, a movie’s ratings by lots of people) with its lower-dimensional essence. The first principal component is about seeing which one-dimensional line best approximates the data points — i.e. which is the line for which projecting the data points onto it results in “estimates” that are as close as possible to the data points.

In the case of this problem, every point is explicitly given as a multiple of a single vector \( \vec{w} \) and so the data already lies on such a straight line going through the origin. So, the first principal component is just along the direction of \( \vec{w} \). Because a principal component represents a direction, it is conventional to normalize the vector to have unit length. In this case, we are told that the vector \( \vec{w} \) has length 7, and so the answer is \( \frac{\vec{w}}{7} \).

(Because the line is all that matters, you could also have used the negative of this \( -\frac{\vec{w}}{7} \).)

We can also do this using the SVD.

The singular value decomposition of a matrix \( R \) is a way of decomposing \( R \) into a sum of rank 1 matrices. In this sum the \( i^{th} \) rank 1 matrix is formed from taking the outer product of normalized column vectors \( \vec{u}_i \) and normalized row vectors \( \vec{w}_i^\top \), scaled by their respective singular values \( \sigma_i \).
Looking at our given $R$, we can see that the matrix itself is rank 1 as the columns are all multiples of the same vector: $\vec{w}$. Seeing this we realize we can rewrite the matrix $R$ as the following outer product:

$$R = \begin{bmatrix} 1 \\ \vec{w} \end{bmatrix} \begin{bmatrix} -1 & -2 & 2 & 4 \end{bmatrix} \quad (17)$$

However the SVD requires we normalize the vectors $\vec{u}_1$ and $\vec{v}_1^T$. In order to reconstruct $A$ properly we must scale back with the norms that we divided out to normalize.

$$\|[-1, -2, 2, 4]\| = \sqrt{25} = 5 \quad \text{and} \quad \|\vec{w}\| = 7.$$ Consequently, when we pull that out, we get $\sigma_1 = 35$ as the singular value that corresponds to the first (and only) principal component. Thus we can write the SVD of $R$ as:

$$R = \begin{bmatrix} 1 \\ \vec{w} \end{bmatrix} \begin{bmatrix} -1 \\ -2 \\ 2 \\ 4 \\ \frac{1}{7} \end{bmatrix} \quad (18)$$

Now we just have to pick which normalized vector to deem the principal component. Since our data (the movie ratings) are collected as columns we choose $\vec{w} / 7$ as the principal component.

Alternate Solution:
To find the first principal component along the columns, we can use $\Sigma$ and $U$. This is because our data is stored in the columns of $R$. We know that

$$R = U\Sigma V^T \quad (19)$$
$$\Rightarrow RR^T = U\Sigma \Sigma^T U^T \quad (20)$$

where $\Sigma \Sigma^T$ represents the square diagonal matrix with $\min(m, n)$ singular values squared on the diagonal. Plugging in for $R$ gives

$$RR^T = \begin{bmatrix} -\vec{w} & -2\vec{w} & 2\vec{w} & 4\vec{w} \end{bmatrix} \begin{bmatrix} -\vec{w} \\ -2\vec{w} \\ 2\vec{w} \\ 4\vec{w} \end{bmatrix} \quad (21)$$

$$RR^T = \vec{w}\vec{w}^T + 4\vec{w}\vec{w}^T + 4\vec{w}\vec{w}^T + 16\vec{w}\vec{w}^T \quad (22)$$

$$RR^T = 25\vec{w}\vec{w}^T \quad (23)$$

By the Spectral Theorem, we know that the eigenvectors of $RR^T$ correspond to the columns of $U$ and the eigenvalues of $RR^T$ correspond to the singular values squared. By inspection, we can see that the eigenvector is $\vec{w}$. Solving for the eigenvalue and singular value:

$$RR^T \vec{w} = 25\vec{w}\vec{w}^T \vec{w} \quad (24)$$
For PCA, we require normalized vectors, so for that reason our first principal component is \( \vec{w} = \frac{\vec{w}}{\|\vec{w}\|} \) with a corresponding singular value of \( \sigma = 35 \).

(c) Suppose that we now collected two more data points (corresponding to two more movies being rated by the same set of six people, i.e. two additional columns for our matrix) that are multiples of a different vector \( \vec{p} \) where:

\[
\vec{p} = \begin{bmatrix}
6 \\
3 \\
-2 \\
0 \\
0 \\
0
\end{bmatrix}
\]

(For your convenience, note that \( \|\vec{p}\| = 7 \) and that \( \vec{p}^\top \vec{w} = 0 \).)

We augment our ratings data matrix with these two new data points to get:

\[
R = \begin{bmatrix}
-\vec{w} & -2\vec{w} & 2\vec{w} & 4\vec{w} & -3\vec{p} & 3\vec{p}
\end{bmatrix}
\]

Find the first two principal components corresponding to the nonzero singular values of \( R \). This is what we would use to best project the movie data points onto a two-dimensional subspace.

What is the first principal component vector? What is the second principal component vector?

Justify your answer. (HINT: Think about the inner product of \( \vec{w} \) and \( \vec{p} \) and what that implies for being able to appropriately decompose \( R \). Again, very little computation is required here.)

Solution: The solution to the previous part tells you what we need to do. We need to find the best two-dimensional subspace that best represents our data.

We start by taking the SVD of \( R \).

The columns of \( R \) are all multiples of two vectors: \( \vec{w} \) and \( \vec{p} \). Each of these can be used to create a rank 1 matrix, and these can be summed together to form \( R \).

Since \( \vec{w} \) and \( \vec{p} \) are orthogonal to one another, our life is easier. This problem’s \( R \) matrix is made especially nice by seeing that a data point is either purely in the \( \vec{w} \) direction, or the \( \vec{p} \) direction.

Using this knowledge we rewrite \( R \) as:

\[
R = \begin{bmatrix}
\vec{w}
\end{bmatrix}
\begin{bmatrix}
-1 & -2 & 2 & 4 & 0 & 0
\end{bmatrix}
+ \begin{bmatrix}
\vec{p}
\end{bmatrix}
\begin{bmatrix}
0 & 0 & 0 & -3 & 3
\end{bmatrix}.
\]

The orthogonality relationships demanded by the SVD are clearly satisfied since the row-vectors involved above have disjoint support (i.e. when one is nonzero, the other is zero) and the columns are orthogonal since we’ve been told so.

However for the SVD the vectors: \( \vec{u}_1 \), \( \vec{v}_1^\top \), \( \vec{u}_2 \) and \( \vec{v}_2^\top \) must be normalized and each rank 1 matrix must be scaled by the appropriate \( \sigma \) to allow the sum to properly reconstruct \( R \). We also need...
to figure out which $\sigma_i$ is bigger so we can order them properly. In the previous part, we have already done the calculations for $\vec{w}$’s part in this story. So what remains is the $\vec{p}$ part. Clearly the norm of the relevant row is $3\sqrt{2}$ which the norm of the relevant column is 7. So the singular value in question is $21\sqrt{2}$.

Using this we can rewrite $R$ as:

$$
\begin{bmatrix}
  \vec{w} \\
  \vec{p}
\end{bmatrix}
= \begin{bmatrix}
  35 & -\frac{3}{\sqrt{2}} \\
  \frac{3}{3\sqrt{2}} & 3\sqrt{2}
\end{bmatrix}
\begin{bmatrix}
  \vec{w} \\
  \vec{p}
\end{bmatrix}
.$$  

(30)

Since $35 > 21\sqrt{2}$, we arrange them in descending order such that our singular values are $\sigma_1 = 35$ and $\sigma_2 = 21\sqrt{2}$. Thus $\vec{w}$ which corresponds to $\sigma_1$ is still the first principal component vector and $\vec{p}$ which corresponds to $\sigma_2$ is the second principal component vector.

**Alternate Solution:**

Again, to find the first principal component along the columns, we can use $\Sigma$ and $U$. This is because our data is stored in the columns of $R$.

$$
R = U \Sigma V^T
$$

(31)

$$
\Rightarrow RR^T = U \Sigma \Sigma^T U^T
$$

(32)

where $\Sigma \Sigma^T$ represents the square diagonal matrix with $\min(m, n)$ singular values squared on the diagonal. Plugging in for $R$ gives

$$
RR^T = \begin{bmatrix}
-\vec{w} & -2\vec{w} & 4\vec{w} & -3\vec{p} & 3\vec{p}
\end{bmatrix}
\begin{bmatrix}
-\vec{w} \\
-2\vec{w} \\
2\vec{w} \\
4\vec{w} \\
-3\vec{p} \\
3\vec{p}
\end{bmatrix}
$$

(33)

$$
RR^T = \vec{w}\vec{w}^T + 4\vec{w}\vec{w}^T + 4\vec{w}\vec{w}^T + 16\vec{w}\vec{w}^T + 9\vec{p}\vec{p}^T + 9\vec{p}\vec{p}^T
$$

(34)

$$
RR^T = 25\vec{w}\vec{w}^T + 18\vec{p}\vec{p}^T
$$

(35)

By the spectral decomposition, we know that the eigenvectors of $RR^T$ correspond to the columns of $U$ and the eigenvalues of $RR^T$ correspond to the singular values squared. By inspection, we can see that the first eigenvector is $\vec{w}$ and the second one is $\vec{p}$. Starting with the first:

$$
RR^T \vec{w} = 25\vec{w}\vec{w}^T \vec{w} + 18\vec{p}(\vec{p}^T \vec{w})
$$

(36)

$$
RR^T \vec{w} = (25\vec{w}^T \vec{w})\vec{w} + 0
$$

(37)

$$
\Rightarrow \lambda_1 = 25\vec{w}^T \vec{w} = 5^2 7^2
$$

(38)

$$
\Rightarrow \sigma_1 = \sqrt{\lambda_1} = 35
$$

(39)

When computing the first eigenvalue, we used the fact that $\vec{p}^T \vec{w} = 0$. We can repeat the same process for the second eigenvalue:

$$
RR^T \vec{p} = 25\vec{w}(\vec{w}^T \vec{p}) + 18\vec{p}(\vec{p}^T \vec{w})
$$

(40)
\[ RR^T \bar{p} = 0 + 18(\bar{p}^T \bar{p}) \bar{p} \]  
(41)

\[ \implies \lambda_2 = 18\bar{p}^T \bar{p} = 3^27^22 \]  
(42)

\[ \implies \sigma_2 = \sqrt{\lambda_2} = 21\sqrt{2} \]  
(43)

For PCA, we require normalized vectors, so for that reason our first principal component is \( \frac{\bar{w}}{\|\bar{w}\|} = \frac{\bar{w}}{7} \) with a corresponding singular value of \( \sigma_1 = 35 \). Our second principal component is \( \frac{\bar{p}}{\|\bar{p}\|} = \frac{\bar{p}}{7} \) with a corresponding singular value of \( \sigma_2 = 21\sqrt{2} \). As mentioned in the solution above, we choose \( \sigma_1 = 35, \sigma_2 = 21\sqrt{2} \) in this order because \( 35 > 21\sqrt{2} \).

(d) In the previous part, you had

\[
R = \begin{bmatrix}
\bar{w} & -2\bar{w} & 2\bar{w} & 4\bar{w} & -3\bar{p} & 3\bar{p}
\end{bmatrix}
\]

with \( \|\bar{w}\| = 7 \) and \( \|\bar{p}\| = 7 \), satisfying \( \bar{p}^T \bar{w} = 0 \).

Plot the movie data points \( \vec{r}_i \) (for all \( i \), where \( \vec{r}_i \) denotes the \( i \)-th column of \( R \)) projected onto the first and second principal component vectors of the columns of \( R \). The horizontal axis should reflect the coordinate along the first principal component, and the vertical axis should reflect the coordinate along the second principal component. **Label each point, and the axes. Remember that principal component vectors are normalized.**
Solution: Once we know what the principal components are, we know that the first four data points are just multiples of the first principal component and the last two data points are just multiples of the second principal component. What multiples? For the first four, the multiples are clearly $-7, -14, 14, 28$ since the norm of $\vec{w}$ is 7. For the final two, the multiples are clearly $-21, +21$ since the norm of $\vec{p}$ is also 7. Plotting:
3. PCA Plots

In each plot below, some $d$-dimensional data is projected onto two unit vectors. The $x$-coordinate is the projection onto the first vector (written as $\vec{v}_1$), and the $y$-coordinate is the projection onto the second vector (written as $\vec{v}_2$). Mathematically, we can say that we have some data matrix $D = \begin{bmatrix} \vec{d}_1 & \vec{d}_2 & \ldots & \vec{d}_n \end{bmatrix}$ where each $\vec{d}_i \in \mathbb{R}^d$. We then project each $\vec{d}_i$ onto the column space of the matrix $V = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix}$ and plot the projection coefficients below. We say that a plot is “valid” if $\vec{v}_1$ could be the first principal component.

(HINT: Note that the mean or “center of mass” of the data points is the origin, $(0, 0)$, in all of the plots. The procedure for this problem is very similar to the procedure in Discussion 14B, where you were similarly asked to judge the “validity” of scatterplots.)

(a) Is the following plot valid?

![Diagram of PCA plot]

**Solution:**

\[\bigotimes\text{Valid} \quad \bigotimes\text{Invalid}\]

In this case, the coordinates on $\vec{v}_2$ axis seems to have a larger sum of squares than the $\vec{v}_1$ axis, since the coordinates on the $\vec{v}_2$ axis range from -2 to 2.
(b) Is the following plot valid?

Solution:

In this case, the coordinates on the $\vec{v}_1$ axis seem to maximize the sum of squares (i.e., there is no other axis that we can construct which will have a larger sum of squares of coordinates). Another way to approach this problem would be to notice that the data is centered around the origin, so the first principal component will be the direction of greatest “spread” in the data, which is the $\vec{v}_1$ direction.
(c) Is the following plot valid?

Solution:

Since our data appears to be centered (the data is symmetric about the $x$- and $y$-axes), we can find the direction that maximizes the spread in our data. In this case, the direction appears to be given by the axis $\frac{1}{\sqrt{2}} \vec{v}_1 + \frac{1}{\sqrt{2}} \vec{v}_2$. 
(d) Is the following plot valid? Assume that $k \to \infty$.

![Plot Image]

**Solution:**

Any axis we choose can be represented as $\alpha \vec{v}_1 + \beta \vec{v}_2$, a linear combination of the axes shown in the plot. We require $\alpha^2 + \beta^2 = 1$ since the axes have to be normal, so as long as $\alpha \neq 0$, we will be maximizing the sum of squares of coordinates. Specifically, if we choose an axis with $\alpha \neq 0$, then the sum of squares of coordinates will be $\infty$ in the limit, and $\vec{v}_1$ is an axis with $\alpha \neq 0$ so the plot is valid.
4. (OPTIONAL) Linearization of a Scalar System

In this question, we linearize the scalar differential equation

\[
\frac{d}{dt} x(t) = \sin(x(t)) + u(t)
\]  

(44)

around equilibria, discretize it, and apply feedback control to stabilize the resulting system.

(a) The first step is to find the equilibria that we will linearize around. Recall that equilibria are the values of \((x, u)\) such that \(\frac{d}{dt} x(t) = 0\). Suppose we want to linearize around equilibria \((x^*, u^*)\) where \(u^* = 0\). Sketch \(\sin(x)\) for \(-4\pi \leq x \leq 4\pi\) and intersect it with a horizontal line at 0. Then, argue why \(x^*_m = m\pi\) and \(u^* = 0\) are equilibria of the system (44).

Solution:

![Figure 1: Plot of \(\sin(x)\)](image)

We can see that all the multiples of \(\pi\) are where the line intersects the sine wave. At \((x^*_m, u^*)\) we have

\[
\frac{d}{dt} x(t) = \sin(x^*_m) + u^*
\]  

(45)

\[
= \sin(m\pi) + 0
\]  

(46)

\[
= 0
\]  

(47)

\[
= 0
\]  

(48)

so \((x^*_m, u^*)\) are equilibria of the system.

(b) Linearize system (44) around the equilibrium \((x^*_0, u^*) = (0, 0)\). What is the resulting linearized scalar differential equation for \(\delta x(t) = x(t) - x^*_0 = x(t) - 0\), involving \(\delta u(t) = u(t) - u^* = u(t) - 0\)?

Solution: We have

\[
\frac{dx}{dt} = f(x(t), u(t)) = \sin(x(t)) + u(t)
\]  

(49)
\[
\frac{d}{dt}\delta x(t) \approx \frac{\partial f}{\partial x}(x^*, u^*)\delta x(t) + \frac{\partial f}{\partial u}(x^*, u^*)\delta u(t) \\
\quad = \cos(0)\delta x(t) + (1)\delta u(t) \\
\quad = \delta x(t) + \delta u(t). \tag{50}
\]

(c) Given an arbitrary, continuous linear system as in
\[
\frac{dx(t)}{dt} = \lambda x(t) + bu(t) \tag{53}
\]
discretizing it into intervals of \(\Delta\) gives the discrete-time system
\[
x[i+1] = e^{\lambda \Delta}x[i] + \frac{b(e^{\lambda \Delta} - 1)}{\lambda}u[i] \tag{54}
\]
Using this result, discretize the approximate linear system. Is the (approximate) discrete-time system stable?

**Solution:** By pattern matching from eq. (52) to eq. (53), we have that \(\lambda = 1\) and \(b = 1\). Plugging into eq. (54), we get
\[
\delta x[i+1] = e^{\lambda \Delta} \delta x[i] + \delta u[i](e^\lambda - 1) \tag{55}
\]
For a linear scalar discrete time recurrence relation, stability is determined by the coefficient of the system’s variable, in this case \(\delta x\). We know that if the magnitude of this coefficient is between -1 and 1, our system is stable. But \(e^\lambda > 1\) for all positive \(\Delta\) (and \(\Delta\) by definition has to be positive). Hence, our system is not stable.

(d) Suppose for the linearized discrete-time system that you found in the previous part, we apply the feedback law
\[
\delta u[i] = -k(\delta x[i] - x^*). \tag{56}
\]
For what range of \(k\) values would the resulting linearized discrete-time system be stable? Your answer will depend on \(\Delta\).

**Solution:** Based on our definition of \(\delta x\), we have, \(\delta u[n] = -k\delta x[n]\). Substituting and grouping the terms, we get
\[
\delta x[i+1] = \delta x[i]\left( e^\Delta - k(e^\Delta - 1) \right) \tag{56}
\]
Hence, we want the above coefficient to between -1 and 1.
\[
-1 < e^\Delta - k(e^\Delta - 1) < 1 \tag{57}
\]
\[
-(1 + e^\Delta) < -k(e^\Delta - 1) < 1 - e^\Delta \tag{58}
\]
\[
\Rightarrow 1 < k < \frac{e^\Delta + 1}{e^\Delta - 1} \tag{59}
\]
5. **(OPTIONAL) Linearization to Understand Amplification**

Linearization isn’t just something that is important for control, robotics, machine learning, and optimization — it is also one of the standard tools used across different hardware disciplines, including circuits.

The circuit below is a voltage amplifier, where the element inside the box is a bipolar junction transistor (BJT). You do not need to know what a BJT is to do this question.

The BJT in the circuit can be modeled quite accurately as a nonlinear, voltage-controlled current source, where the collector current $I_C$ is given by:

$$I_C(V_{\text{in}}) = I_S \cdot e^{\frac{V_{\text{in}}}{V_{TH}}},$$

where $V_{TH}$ is the thermal voltage. We can assume that $V_{TH} = 26\, \text{mV}$ at room temperature. $I_S$ is a constant we are not providing because we want you to find ways of eliminating it in favor of other quantities whenever possible.

The goal of this circuit is to pick a particular point $(V_{\text{in}}^*, V_{\text{out}}^*)$ so that any small variation $\delta V_{\text{in}}$ in the input voltage $V_{\text{in}}$ can be amplified to a relatively larger variation $\delta V_{\text{out}}$ in the output voltage $V_{\text{out}}$. In other words, if $V_{\text{in}} = V_{\text{in}}^* + \delta V_{\text{in}}$ and $V_{\text{out}} = V_{\text{out}}^* + \delta V_{\text{out}}$, then we want the magnitude of the ‘amplification gain’ given by $\left|\frac{\delta V_{\text{out}}}{\delta V_{\text{in}}}\right|$ to be large. We’re going to investigate this amplification using linearization.

**(NOTE:** In this problem, $\delta V$ is single variable indicating a small variation in $V$, not $\delta \times V$.)

**(a) Write a symbolic expression for $V_{\text{out}}$ as a function of $I_C$, $V_{DD}$ and $R$ in Fig 2.**

**Solution:**

$$V_{\text{out}} = V_{DD} - R I_C$$

since we have a voltage drop of $I_C R$ across the resistor and the top voltage is $V_{DD}$. 

---

© UCB EECS 16B, Spring 2024. All Rights Reserved. This may not be publicly shared without explicit permission.
(b) Now let’s linearize $I_C$ in the neighborhood of an input voltage $V_{in}^\ast$ and a specific $I_C^\ast$. Assume that you have a found a particular pair of input voltage $V_{in}^\ast$ and current $I_C^\ast$ that satisfy the current equation (60).

We can look at nearby input voltages and see how much the current changes. We can write the linearized expression for the collector current around this point as:

$$I_C(V_{in}) = I_C(V_{in}^\ast) + g_m(V_{in} - V_{in}^\ast) = I_C^\ast + g_m \delta V_{in} \quad (62)$$

where $\delta V_{in} = V_{in} - V_{in}^\ast$ is the change in input voltage, and $g_m$ is the slope of the local linearization around $(V_{in}^\ast, I_C^\ast)$. What is $g_m$ here as a function of $I_C^\ast$ and $V_{TH}$?

(HINT: Find $g_m$ by taking the appropriate derivative around the equilibrium point. You should recognize a part of your equation is equal to the current equilibrium point $I_C^\ast = I_C(V_{in}^\ast)$, so your final form should not depend on $I_S$.)

**Solution:** We start out by writing out the linearization form that we are looking for:

$$I_C(V_{in}) = I_C^\ast + g_m \delta V_{in} \quad (63)$$

Now, taking the first derivative of $I_C$ around $V_{in}^\ast$:

$$g_m = \frac{dI_C(V_{in}^\ast)}{dV_{in}} \quad (64)$$

$$= \frac{1}{V_{TH}} I_S e^{V_{in}^\ast/V_{TH}} \quad (65)$$

$$= \frac{I_C^\ast}{V_{TH}} \quad (66)$$

where in the last line, we recognize that $I_C^\ast = I_S e^{V_{in}^\ast/V_{TH}}$, and therefore knowledge of $I_S$ is not required to determine $g_m$ if $I_C^\ast$ and $V_{TH}$ are known.
For understanding this linearization graphically, we can choose to plot $I_C$ vs. $V_{in}$ and look at how the slope (i.e. $g_m$) changes with different values of $(V_{in}^*, I_C^*)$. In Fig. 3, we see the linearizations around $V_{in}^* = 0.65$ V ($I_C^* = 1$ mA) and $V_{in}^* = 0.7$ V ($I_C^* = 9$ mA) given in part (d) below.

(c) We now have a linear relationship between small changes in current and voltage, $\delta I_C = g_m \delta V_{in}$ around a known solution $(V_{in}^*, I_C^*)$.

As a reminder, the goal of this problem is to pick a particular point $(V_{in}^*, V_{out}^*)$ so that any small variation $\delta V_{in}$ in the input voltage $V_{in}$ can be amplified to a relatively larger variation $\delta V_{out}$ in the output voltage $V_{out}$. In other words, if $V_{in} = V_{in}^* + \delta V_{in}$ and $V_{out} = V_{out}^* + \delta V_{out}$, then we want the magnitude of the “amplification gain” given by $\left| \frac{\delta V_{out}}{\delta V_{in}} \right|$ to be large.

Plug your linearized equation for $I_C$ into the answer from part (a). It may help to define the output voltage equilibrium point as $V_{out}^*$, where

$$V_{out}^* = V_{DD} - R I_C^*$$

so that we can view $V_{out} = V_{out}^* + \delta V_{out}$ when we have $V_{in} = V_{in}^* + \delta V_{in}$.

Find the linearized relationship between $\delta V_{out}$ and $\delta V_{in}$. The ratio $\frac{\delta V_{out}}{\delta V_{in}}$ is called the “small-signal voltage gain” of this amplifier around this equilibrium point.

Solution: We have two equations for $V_{out}$:

$$V_{out} = V_{out}^* + \delta V_{out}$$

and

$$V_{out} = V_{DD} - R I_C$$

We know from equation (62) that $I_C = I_C^* + g_m \delta V_{in}$, so we can re-write the above two equations as:

$$V_{out}^* + \delta V_{out} = V_{DD} - R (I_C^* + g_m \delta V_{in})$$
We also know the output voltage equilibrium point $V_{\text{out}}^*$ is related to the current equilibrium point $I_C^*$ as $V_{\text{out}}^* = V_{DD} - RI_C^*$, hence:

$$V_{DD} - RI_C^* + \delta V_{\text{out}} = V_{DD} - R(I_C^* + g_m \delta V_{\text{in}})$$  \hspace{1cm} (71)

$$\Rightarrow \delta V_{\text{out}} = -RG_m \delta V_{\text{in}}$$  \hspace{1cm} (72)

We re-arrange and solve for the small-signal voltage gain:

$$\frac{\delta V_{\text{out}}}{\delta V_{\text{in}}} = -R \frac{I_C^*}{V_{TH}}$$  \hspace{1cm} (73)

You are not required to simplify it beyond this point. However, recognize that $I_C^* R = V_{DD} - V_{\text{out}}^*$, so we can relate the small-signal voltage gain directly to the output voltage equilibrium point:

$$\frac{\delta V_{\text{out}}}{\delta V_{\text{in}}} = -\frac{V_{DD} - V_{\text{out}}^*}{V_{TH}}$$  \hspace{1cm} (74)

This suggests that we want the voltage “gap” between the supply voltage $V_{DD}$ and the output voltage bias (i.e. DC) point $V_{\text{out}}^*$ to be large if we want a large voltage gain. For a fixed $V_{DD}$, this means a lower $V_{\text{out}}^*$. However, when you learn about BJT devices properly, you will see the output bias voltage can only go so low before our models fails. We also notice from equations (73) and (74) that to get a higher voltage gain, we need a larger bias (DC) current $I_C^*$ (to get a lower $V_{\text{out}}^*$ means we need a larger $I_C^*$ through the resistor). In other words, to get higher voltage gain, we need to burn more power.

(d) Assuming that $V_{DD} = 10$ V, $R = 1$ kΩ, and $I_C^* = 1$ mA when $V_{\text{in}}^* = 0.65$ V, verify that the magnitude of the small-signal voltage gain $\left| \frac{\delta V_{\text{out}}}{\delta V_{\text{in}}} \right|$ is approximately 38.

Next, if $I_C^* = 9$ mA when $V_{\text{in}}^* = 0.7$ V with all other parameters remaining fixed, verify that the magnitude of the small-signal voltage gain $\left| \frac{\delta V_{\text{out}}}{\delta V_{\text{in}}} \right|$ between the input and the output around this equilibrium point is approximately 346.

(HINT: Remember $V_{TH} = 26$ mV.)

**Solution:** Just plugging in to equation (73):

$$\left| \frac{\delta V_{\text{out}}}{\delta V_{\text{in}}} \right| = \frac{I_C^* R}{V_{TH}} = \left( \frac{1 \text{ mA} \times 1 \text{ kΩ}}{26 \text{ mV}} \right) = \frac{1 \text{ V}}{26 \text{ mV}} \approx 38$$  \hspace{1cm} (75)

Now, if $I_C^* = 9$ mA when $V_{\text{in}}^* = 0.7$ V, we have

$$\left| \frac{\delta V_{\text{out}}}{\delta V_{\text{in}}} \right| = \frac{I_C^* R}{V_{TH}} = \left( \frac{9 \text{ mA} \times 1 \text{ kΩ}}{26 \text{ mV}} \right) = \frac{9 \text{ V}}{26 \text{ mV}} \approx 346$$  \hspace{1cm} (76)

As an aside, notice here that $V_{\text{out}}^*$ has already been pulled down to around 1 V (= $V_{DD} - RI_C^* = 10$ V − 1 kΩ × 9 mA). Realistically, this is close to as low as $V_{\text{out}}^*$ can get for this device; the small-signal voltage gain is close to its upper limit. When you first saw the BJT circuit, it may not have been obvious that $V_{DD}$ and $V_{TH}$ provide the fundamental limit on the small-signal gain for such circuits - you may have been tempted to say the upper bound was related to the resistor value. But the simple linearization analysis in part (c) reveals $V_{DD}$ and $V_{TH}$ are setting the true limit. Courses like EE105 and EE140 further develop these insights in circuit design along with feedback control in interesting and very practical ways.
(e) If you wished to make an amplifier with as large of a small signal gain as possible, **which operating (bias) point would you choose among** $V_{\text{in}}^* = 0.65 \text{ V}$ and $V_{\text{in}}^* = 0.7 \text{ V}$?

**Solution:** We would choose $V_{\text{in}}^* = 0.7 \text{ V}$ since the magnitude of the small signal gain in this equilibrium point is much higher than that at $V_{\text{in}}^* = 0.65 \text{ V}$.

Note that since $I_C^*$ is related to $V_{\text{in}}^*$ by (60), and $V_{\text{out}}^*$ is related to $I_C^*$ by (61), just choosing $V_{\text{in}}^*$ fixes the small signal gain of the circuit in Fig. 2.

This shows you that by appropriately biasing (choosing an equilibrium point), we can adjust our gain for small signals. While we just wanted to show you a simple application of linearization here, these ideas are developed a lot further in EE105, EE140, and other courses to create things like op-amps and other analog information-processing systems. Simple voltage amplifier circuits like these are used in everyday circuits like the sensors in your smartwatch, wireless transceivers in your phone, and communication circuits in CPUs and GPUs.

**Contributors:**
- Nikhil Shinde.
- Nathan Lambert.
- Anant Sahai.
- Daniel Abraham.
- Wahid Rahman.
- Sally Hui.
- Anish Muthali.
- Kris Pister.
- Alex Devonport.
- Regina Eckert.