Homework 13

This homework is due on Saturday, April 29, 2023 at 11:59PM. Self-grades and HW Resubmissions are due the following Saturday, May 6, 2023 at 11:59PM.

1. **Study Groups**

   If you are a student who participated in the study group survey we gave in the early weeks of the semester, we would really appreciate your feedback on the group you were matched with. If you did not participate in a study group, we would appreciate your input on what factors went into this decision. Please fill out this form to provide any feedback. For this question, write that you completed the survey.

2. Practical SVD System ID

Please answer all of the questions in the Jupyter notebook associated with this homework.

3. PCA Introduction

Let $X \in \mathbb{R}^{m\times n}$ be defined as $X := \begin{bmatrix} \vec{x}_1 & \cdots & \vec{x}_n \end{bmatrix}$ where each $\vec{x}_i \in \mathbb{R}^m$. Let X have an SVD $X = U\Sigma V^\top$. Now, let $U_\ell := \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_\ell \end{bmatrix}$ where $\vec{u}_i$ is the ith column of U. In other words, U_ℓ is the first ℓ columns of U. In this problem, we will go about showing that

$$U_\ell \in \underset{W\in\mathbb{R}^{m\times \ell}}{\operatorname{argmin}} \sum_{i=1}^{n} \left\| \vec{x}_i - WW^\top \vec{x}_i \right\|^2 \tag{1}$$

where $W^\top W = I_\ell$ (i.e., it is a matrix with orthonormal columns). This is an important result for deriving PCA, as you will see in lecture.

(a) **First, show that**

$$\left\| \vec{x}_i - WW^\top \vec{x}_i \right\|^2 = \|\vec{x}_i\|^2 - \left\| W^\top \vec{x}_i \right\|^2 \tag{2}$$

(HINT: Expand the left hand side of the equation above using transposes. That is, use the fact that $\|\vec{v}\|^2 = \vec{v}^\top \vec{v}$.)

(b) Using the result from the previous part, we can simplify the original optimization problem to say

$$\underset{W\in\mathbb{R}^{m\times \ell}}{\operatorname{argmin}} \sum_{i=1}^{n} \left\| \vec{x}_i - WW^\top \vec{x}_i \right\|^2 = \underset{W\in\mathbb{R}^{m\times \ell}}{\operatorname{argmin}} \sum_{i=1}^{n} \left(\|\vec{x}_i\|^2 - \left\| W^\top \vec{x}_i \right\|^2 \right) \tag{3}$$

$$\iff \underset{W\in\mathbb{R}^{m\times \ell}}{\operatorname{argmin}} \sum_{i=1}^{n} \left(- \left\| W^\top \vec{x}_i \right\|^2 \right) \tag{4}$$

$$\iff \underset{W\in\mathbb{R}^{m\times \ell}}{\operatorname{argmax}} \sum_{i=1}^{n} \left\| W^\top \vec{x}_i \right\|^2 \tag{5}$$

where we get the second line from noticing that we cannot change $\vec{x}_i$, so we remove it from the optimization problem. Then, we pull out the negative to turn the minimization problem into a maximization problem. Now, let $W := \begin{bmatrix} \vec{w}_1 & \cdots & \vec{w}_\ell \end{bmatrix}$. **Show that**

$$\sum_{i=1}^{n} \left\| W^\top \vec{x}_i \right\|^2 = \sum_{k=1}^{\ell} \vec{w}_k^\top \left(XX^\top \right) \vec{w}_k \tag{6}$$

You may use the fact that $\sum_{i=1}^{n} \vec{x}_i \vec{x}_i^\top = XX^\top$. *(HINT: Start by expanding out the norm squared expression as the sum of squares of the individual entries of $W^\top \vec{x}_i$.)*

(c) **Use the result of the previous part to show that**

$$\sum_{i=1}^{n} \left\| W^\top \vec{x}_i \right\|^2 = \sum_{k=1}^{\ell} \vec{\tilde{w}}_k^\top \Sigma\Sigma^\top \vec{\tilde{w}}_k \tag{7}$$

where $\vec{\tilde{w}}_k = U^\top \vec{w}_k$. Then, argue that $\Sigma\Sigma^\top$ can be written as

$$\Sigma\Sigma^\top = \begin{bmatrix} \sigma_1^2 & & & & & \\ & \ddots & & & & \\ & & \sigma_r^2 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} \tag{8}$$

where $r = \text{rank}(X)$ *(HINT: Use the SVD of X to simplify the $XX^\top$ term from the previous part.)*

(d) From the previous part, we have the following expression:

$$\sum_{i=1}^{n} \left\|W^\top \vec{x}_i\right\|^2 = \sum_{k=1}^{\ell} \vec{\tilde{w}}_k^\top \begin{bmatrix} \sigma_1^2 & & & & & \\ & \ddots & & & & \\ & & \sigma_r^2 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} \vec{\tilde{w}}_k \tag{9}$$

One may show (via Cauchy-Schwarz) that

$$\sum_{k=1}^{\ell} \vec{\tilde{w}}_k^\top \begin{bmatrix} \sigma_1^2 & & & & & \\ & \ddots & & & & \\ & & \sigma_r^2 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} \vec{\tilde{w}}_k \leq \sum_{k=1}^{\ell} \sigma_k^2 \tag{10}$$

if $\vec{\tilde{w}}_k$ are required to be orthonormal (you are not required to show this). **Using this fact, find some specific values of $\vec{\tilde{w}}_i$ such that we attain eq. (10) with equality. Then, use this to show that U_ℓ maximizes $\sum_{i=1}^{n} \left\|W^\top \vec{x}_i\right\|^2$ and hence is a solution to the original optimization problem.**

4. Linearizing for understanding amplification

Linearization isn't just something that is important for control, robotics, machine learning, and optimization — it is one of the standard tools used across different areas, 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.

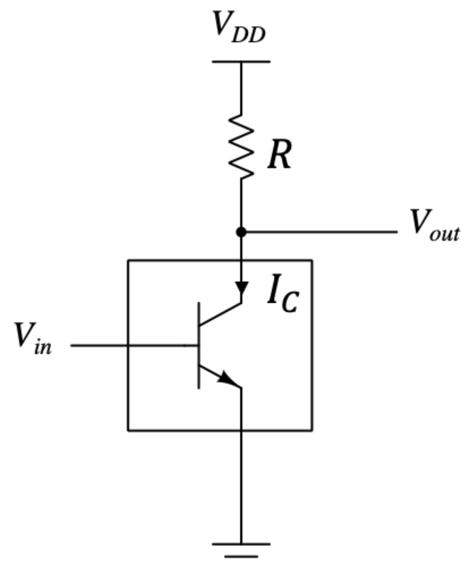

Figure 1: Voltage amplifier circuit using a BJT

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:

$$I_C(V_{\text{in}}) = I_S \cdot e^{\frac{V_{\text{in}}}{V_{\text{TH}}}}, \tag{11}$$

where V_{TH} is the thermal voltage. We can assume $V_{\text{TH}} = 26\,\text{mV}$ at room temperature. I_S is a constant whose exact value we are not giving you 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 $(V_{\text{in}}^\star, V_{\text{out}}^\star)$ so that any small variation δV_{in} in the input voltage V_{in} can be amplified to a relatively larger variation δV_{out} in the output voltage V_{out}. In other words, if $V_{\text{in}} = V_{\text{in}}^\star + \delta V_{\text{in}}$ and $V_{\text{out}} = V_{\text{out}}^\star + \delta V_{\text{out}}$, 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, δV is single variable indicating a small variation in V, not $\delta \times V$.)

(a) **Write a symbolic expression for V_{out} as a function of I_C, V_{DD} and R in Fig 1.**

(b) Now let's linearize I_C in the neighborhood of an input voltage $V_{\text{in}}^\star$ and a specific $I_C^\star$. Assume that you have a found a particular pair of input voltage $V_{\text{in}}^\star$ and current $I_C^\star$ that satisfy the current equation (11).

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:

$$I_C(V_{\text{in}}) = I_C(V_{\text{in}}^\star) + g_m(V_{\text{in}} - V_{\text{in}}^\star) = I_C^\star + g_m\,\delta V_{\text{in}} \tag{12}$$

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

(HINT: Find g_m by taking the appropriate derivative around the operating point. You should recognize a part of your equation is equal to the current operating point $I_C^\star = I_C(V_{\text{in}}^\star)$, so your final form should not depend on I_S. Also, note that in circuits terminology, "operating point" is defined to be the point around which we linearize input-output relationship.)

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

As a reminder, the goal of this problem is to pick a particular point $(V_{\text{in}}^\star, V_{\text{out}}^\star)$ so that any small variation δV_{in} in the input voltage V_{in} can be amplified to a relatively larger variation δV_{out} in the output voltage V_{out}. In other words, if $V_{\text{in}} = V_{\text{in}}^\star + \delta V_{\text{in}}$ and $V_{\text{out}} = V_{\text{out}}^\star + \delta V_{\text{out}}$, 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.

Plug in your linearized equation for I_C in the answer from part (a). It may help to define the output voltage operating point as $V_{\text{out}}^\star$, where

$$V_{\text{out}}^\star = V_{\text{DD}} - RI_C^\star \tag{13}$$

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

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

(d) Assuming that $V_{\text{DD}} = 10\,\text{V}$, $R = 1\,\text{k}\Omega$, and $I_C^\star = 1\,\text{mA}$ when $V_{\text{in}}^\star = 0.65\,\text{V}$, **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 $I_C^\star = 9\,\text{mA}$ when $V_{\text{in}}^\star = 0.7\,\text{V}$ 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 operating point is approximately 346.**

(HINT: Remember $V_{\text{TH}} = 26\,\text{mV}$.

)

(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 $V_{\text{in}}^\star = 0.65\,\text{V}$ and $V_{\text{in}}^\star = 0.7\,\text{V}$?**

This shows you that by appropriately biasing (choosing an operating point), we can adjust what our gain is 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:

- Anish Muthali.
- Kris Pister.
- Alex Devonport.
- Anant Sahai.
- Regina Eckert.
- Wahid Rahman.

