EECS 16A   Designing Information Devices and Systems I
Spring 2017   Babak Ayazifar, Vladimir Stojanovic   Homework 1

This homework is due January 30, 2017, at 23:59.

Self-grades are due February 2, 2017, at 23:59.

Submission Format

Your homework submission should consist of **two** files.

- `hw1.pdf`: A single pdf file that contains all your answers (any handwritten answers should be scanned) as well as your IPython notebook saved as a pdf. You can do this by printing the IPython notebook page in your browser and selecting the save to pdf option. Make sure any plots and results are showing. Also make sure you combine any separate pdfs into one file.

- `hw1.ipynb`: A single IPython notebook with all your code in it.

Submit each file to its respective assignment in Gradescope.

1. **Stojanovic's Optimal Smoothies**

Stojanovic's Optimal Smoothies has a unique way of serving its customers. To ensure the best customer experience, each customer gets a smoothie personalized to his or her tastes. Professor Stojanovic knows that a lot of customers don't know what they want, so when the customer walks up to the counter they are asked to taste four standard smoothies that cover the entire range of flavors found in smoothies.

Each smoothie is made of $\frac{1}{2}$ cup Greek yogurt, $\frac{1}{8}$ cup vanilla soy milk, $\frac{1}{2}$ cup crushed ice, and 1 cup mystery fruit. The four standard smoothies have the following recipes for the cup of mystery fruit:

Fruit(cups)	Banana Berry	Caribbean Passion	Mango-a-go-go	Strawberries Wild
Strawberries	$\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{2}{3}$
Bananas	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{5}$	$\frac{1}{3}$
Mangos	0	$\frac{1}{3}$	$\frac{3}{5}$	0
Blueberries	$\frac{1}{3}$	0	0	0

Each customer is assumed to have a score (from 0 to 10) for each fruit, and the total score for the smoothie is modeled as being computed by multiplying the score for a fruit with its proportion in the smoothie. For example, if a customer's score for strawberries is 6 and bananas is 3, then the total score for the Strawberries Wild smoothie would be $6 \times \frac{2}{3} + 3 \times \frac{1}{3} = 5$.

After a customer gives a score (from 0 to 10) for each smoothie, Professor Stojanovic then calculates (on the spot!) how much the customer likes each fruit. Then Professor Stojanovic blends up a special smoothie that will maximize the customer's score.

Professor Ayazifar was thirsty after grading midterms, so he decided to take a drink break at SOS. He walked in and gave the following ratings:

Smoothie	Score
Banana Berry	$6\frac{1}{3}$
Caribbean Passion	$6\frac{1}{3}$
Mango-a-go-go	$6\frac{4}{5}$
Strawberries Wild	$5\frac{2}{3}$

(a) What were Professor Ayazifar's ratings for each fruit? **Work this problem out by hand.**

(b) What mystery fruit combination should Professor Stojanovic put in Professor Ayazifar's personalized smoothie? What score would Professor Ayazifar give for this smoothie? There may be more than one correct answer.

2. The Framingham Risk Score

The Framingham risk score estimates the 10-year cardiovascular disease (CVD) risk of an individual. There are multiple factors (predictors) that weigh in the calculation of the Framingham score. For each individual, these factors are: gender, age, total cholesterol, level of high-density lipoprotein (HDL) cholesterol, systolic blood pressure (SBP), whether or not the individual smokes, whether or not the individual is treated for high blood pressure and whether or not the individual is diabetic. For this problem, we will focus on the algorithm that estimates the CVD risk for female individuals who smoke, are not treated for high blood pressure and are not diabetic.

To calculate the 10-year CVD risk of an individual in the group described above, a score R is first assigned based on the values of age, total cholesterol, HDL cholesterol and systolic blood pressure as follows

$$R = a \cdot \ln(\text{age (years)}) + b \cdot \ln(\text{total cholesterol (mg/dL)}) +$$
$$c \cdot \ln(\text{HDL cholesterol (mg/dL)}) + d \cdot \ln(\text{SBP (mm Hg)})$$

where a, b, c and d are constant coefficients and $\ln(\cdot)$ denotes the natural (base e) logarithm.

After the score R is calculated, it is plugged into the following formula in order to obtain the risk p (in terms of probability) of the individual suffering from a CVD in the next 10 years:

$$p = 1 - 0.95^{e^{(R-25.66)}}$$

(Note that there is a double exponent in the expression.)

When the algorithm was first devised, the only copy of the document that reported the coefficients a, b, c and d was shredded by mistake by a new intern in the hospital where the research was conducted. The intern needs to restore the values of the coefficients from existing records of hospital patients. He needs your help to do so.

Throughout the problem, you can approximate any numbers up to the fourth decimal. For example, you can approximate 0.23456789 by 0.2346 and 0.24296 by 0.2430.

(a) The intern dug up some of the records for patients in the study group who fit the criteria of the formula in question. The records are summarized in the table 1. Use these records to devise a system of linear equations where a, b, c and d are the unknowns.

(b) Solve the system of linear equations that you devised in question (a) of this problem. For this question, you can use IPython.

Note: The natural logarithm of a number a (i.e. $\ln(a)$) can be obtained as `np.log(a)` in IPython.

Patient ID	Age (years)	Total Chol. (mg/dL)	HDL Chol. (mg/dL)	SBP (mm Hg)	Risk (p)
1	66	198	55	132	0.1550
2	61	180	47	124	0.1108
3	60	180	50	120	0.0940
4	23	132	45	132	0.0105

Table 1: Patient records for 10-year CVD risk assessment

Note: Some of the values in the algorithm were modified from the original study values.

3. **Filtering out the troll**

You attended a very important public speech and recorded it using a recording device that consists of two directional microphones. However, there was this particular person in the audience who was trolling around. When you went back home to listen to the recording you realized that the two recordings are dominated by the troll and you cannot hear the speech. Fortunately since you had two microphones, you realize that there is a way to combine the two recordings such that trolling is removed. Recollecting the scene, the locations of the speaker and the troll are shown in Figure 1.
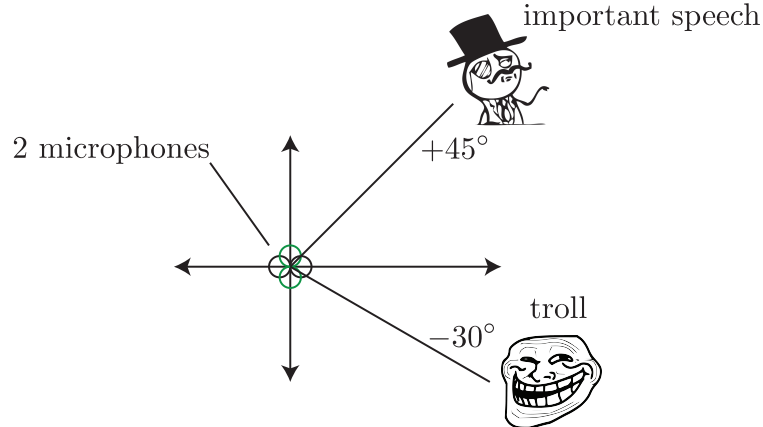

Figure 1:

The way your recording device works is that each microphone weighs the audio signal depending on the angle of the audio source, relative to the *x* axis, hence the name *directional microphones*. More specifically, if the audio source is located at an angle of θ, the first microphone will record the audio signal with weight $f_1(\theta) = \cos(\theta)$, and the second microphone will record the audio signal with weight $f_2(\theta) = \sin(\theta)$. For example, an audio source that lies on the *x* axis will be recorded with the first microphone with weight equal to 1 (since $\cos(0) = 1$), but will not be picked by microphone two (since $\sin(0) = 0$). Note that the weights can also be negative.

Graphically, the directional characteristics of the microphones are given in Figures 2 and 3 (colors red and blue denote positive and negative values of the weight respectively). Putting all of this together, assume that there are two speakers, *A* and *B*, at angles θ and ψ, respectively. Assume that speaker *A* produces an audio signal represented by the vector $\vec{a} \in \mathbb{R}^n$. That is, the *i*-th component of $\vec{a}$ is the signal at the *i*-th time step. Similarly, assume speaker *B* produces an audio signal $\vec{b}$,

Then the first microphone will record the signal

$$\vec{m}_1 = \cos(\theta) \cdot \vec{a} + \cos(\psi) \cdot \vec{b},$$

and the second microphone will record the signal

$$\vec{m}_2 = \sin(\theta) \cdot \vec{a} + \sin(\psi) \cdot \vec{b},$$
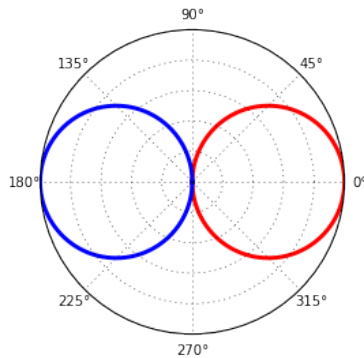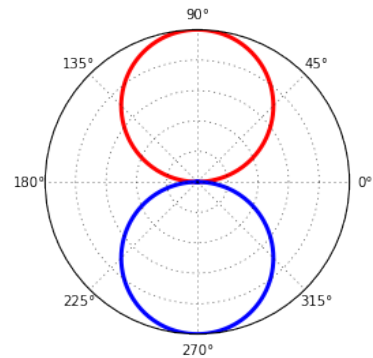

Figure 2: Directional characteristics of mic. 1

Figure 3: Directional characteristics of mic. 2

(a) Using the notation above, let the important speaker be speaker A (with signal $\vec{a}$) and let the person trolling be "speaker" B (with signal $\vec{b}$). Express the recordings of the two microphones $\vec{m}_1$ and $\vec{m}_2$ (i.e. the signals recorded by the first and the second microphones, respectively) as a linear combination of $\vec{a}$ and $\vec{b}$.

(b) Recover the important speech $\vec{a}$, as a weighted combination of $\vec{m}_1$ and $\vec{m}_2$. In other words, write $\vec{a} = u \cdot \vec{m}_1 + v \cdot \vec{m}_2$ (where u and v are scalars). What are the values of u and v?

(c) Partial IPython code can be found in `prob1.ipynb`. Complete the code to get a clean signal of the important speech. What does the speaker say? (Optional: Where is the speech taken from?)

Note: You may have noticed that the recordings of the two microphones sounded remarkably similar. This means that you could recover the real speech from two "trolled" recordings that sounded almost identical! Leave out the fact that the recordings are actually different, and have some fun with your friends who aren't lucky enough to be taking EE16A.

4. **Homework process and study group**

Who else did you work with on this homework? List names and student ID's. (In case of hw party, you can also just describe the group.) How did you work on this homework?
Working in groups of 3-5 will earn credit for your participation grade.

5. **(PRACTICE) Finding charges from potential measurements**

Note: As a practice problem, this problem is optional and will not be worth any points. You are expected to be familiar with the concepts in this problem. Consider saving this problem for studying for exams.

We have three point charges Q_1, Q_2, and Q_3 whose positions are known, and we want to determine their charges. In order to do that, we take three potential measurements U_1, U_2 and U_3 at three different locations. The positions of the charges and potentials are shown in Figure 4.

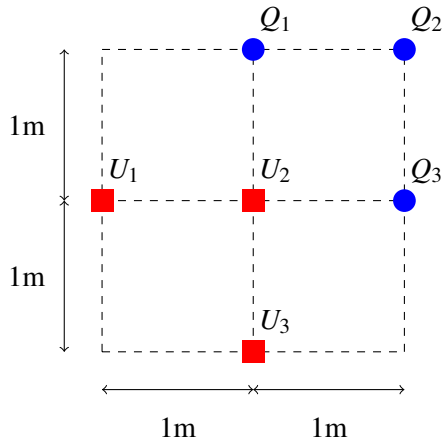For the purpose of this problem, the following equation is true:

$$U = k\frac{Q}{r}$$

Figure 4:

at a point r meters away (for some fixed physical constant k; this problem does not require its numerical value).

Furthermore, the potential contribution from different point charges add up linearly. For example, in the setup of Figure 4, the potential measured at point U_2 is

$$U_2 = k\frac{Q_1}{1} + k\frac{Q_2}{\sqrt{2}} + k\frac{Q_3}{1}$$

Given that the actual potential measurements in the setup of Figure 4 are

$$U_1 = k\frac{4+3\sqrt{5}+\sqrt{10}}{2\sqrt{5}},$$
$$U_2 = k\frac{2+4\sqrt{2}}{\sqrt{2}},$$
$$U_3 = k\frac{4+\sqrt{5}+3\sqrt{10}}{2\sqrt{5}}$$

write the system of linear equations relating the potentials to charges. Solve the system to find the charges Q_1, Q_2, Q_3. You may use your IPython notebook to solve the system.

IPython hint: For constants a_i, b_i, c_i, y_i, you can solve the system of linear equations

$$a_1x_1 + a_2x_2 + a_3x_3 = y_1,$$
$$b_1x_1 + b_2x_2 + b_3x_3 = y_2,$$
$$c_1x_1 + c_2x_2 + c_3x_3 = y_3,$$

in IPython with the following code

```
import numpy as np
a = np.array([
    [a1, a2, a3],
    [b1, b2, b3],
    [c1, c2, c3]
```

```
])
b = np.array([y1, y2, y3])
x = np.linalg.solve(a, b)
```

The square root of a number *a* can be obtained as `np.sqrt(a)` in IPython.