EECS 16A Designing Information Devices and Systems I
Fall 2015 Anant Sahai, Ali Niknejad

Discussion 3B

1. **Lecture Highlights**

2. **Exploring Nullspaces**

   (a) The **column space** of a matrix is the **range** or possible outputs of a transformation/linear operation/function. It is also the **span** of the vectors that form the columns of the matrix.

   (b) The **nullspace** is the set of input vectors that output a zero vector

   For the following five matrices, answer the following questions:

   (a) What is the column span of A? What is its dimension?

   (b) What is the nullspace of A? What is its dimension?

   (c) (optional) Do the columns of A form a basis of $\mathbb{R}^2$? Why or why not?

   (a) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

   (b) $\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$

   (c) $\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$

   (d) $\begin{bmatrix} -2 & 4 \\ 3 & -6 \end{bmatrix}$

   (e) $\begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 3 \\ 0 & -1 & -2 \end{bmatrix}$

3. **Retail Store Marketing**

   Intro The retail store EehEeh Sixteen would like to design a smart register that prints a promotion to each customer when she checks out, depending on things she may be interested in. Therefore, the system can use information from the current purchase only. The problem is, the register doesn't know what the customer's interests are, and the only data the it can use about the customer are her current purchase data.

The Setting The register will have the following information at its disposal:

- A set of promotions $A_1, A_2, \ldots, A_n$. With each promotion A_i, the register also knows a vector $\vec{s_{A_i}} \in \mathbb{R}^4$ that describes the ideal (target) customer in terms of her interests in (i) party products; (ii) family products; (iii) student products and (iv) office products.
- At check out time, the register knows the spending subtotals of the customer in the following 4 categories: Food, movies, art, and books & supplies. These subtotals are denoted by T_f, T_m, T_a and T_b, respectively.

The register needs to decide based on this information which promotion is best to print on the receipt for the customer.

Your Job EehEeh Sixteen hired the same intern from the Framingham heart study to design this smart register. The intern is lost and given the awesomeness of your help last time, he needs your help again. In this problem, you will walk him through designing this smart register.

Let's break down the problem to small components, design each small component individually and then connect them all together into the complete system.

(a) First, we assume that, in addition to the promotions that are given by the store $A_1, A_2, \ldots, A_n$, an oracle provides us with the following information:

- The interests of the customer c in (i) party products; (ii) family products; (iii) student products and (iv) office products described in a vector $\vec{x_c}$
- For any given customer's interests $\vec{x_c}$ and promotion's target customer $\vec{s_A}$, a scalar value $sim(\vec{x_c}, \vec{s_A})$ that is higher the more aligned the customer c is to the promotion A (similarity score).

How can we select which promotion to print to the customer on her receipt?

(b) Since we don't really have the similarity function given to us, we will design one.

i. Would $sim_1(\vec{x_c}, \vec{s_A}) = \|\vec{x_c} - \vec{s_A}\|$ be a good similarity function? Why?

ii. What about $sim_2(\vec{x_c}, \vec{s_A}) = \frac{1}{\|\vec{x_c} - \vec{s_A}\|}$? Why?

iii. What about $sim_3(\vec{x_c}, \vec{s_A}) = \langle \vec{x_c}, \vec{s_A} \rangle$? Why?

iv. What about $sim_4(\vec{x_c}, \vec{s_A}) = \left\langle \vec{x_c}, \frac{\vec{s_A}}{\|\vec{s_A}\|} \right\rangle$? Why?

v. What about $sim_5(\vec{x_c}, \vec{s_A}) = \left\langle \frac{\vec{x_c}}{\|\vec{x_c}\|}, \frac{\vec{s_A}}{\|\vec{s_A}\|} \right\rangle$? Why?

(c) Next, we need to use the customer's spending T_f, T_m, T_a and T_b to infer her interests $\vec{x_c} = \begin{bmatrix} c_{\text{party}} \\ c_{\text{family}} \\ c_{\text{student}} \\ c_{\text{office}} \end{bmatrix}$.

For that, the intern hands you Table 1. The table describes the spending subtotals of a customers with only one interest (either party products, family products, student products or office products).

Furthermore, you can assume that a customer with more than one interest spends her money in a proportional manner to the percentages in Table 1, weighted proportionally to her different interests.

In this part, we will use this information to devise a system of linear equations, which we can solve to infer any customer's interests $\vec{x_c}$ given her spending on food, movies, art and books & supplies.

i. To get started, let's first assume we have a customer who told us she is purely interested in student products, and she told us he is willing to spend T dollars on merchandise. What would be her spending subtotals on food, movies, art and books & supplies?

	Spending Category			
Interest Category	**Food**	**Movies**	**Art**	**Books & Supplies**
Party	40%	33%	22%	5%
Family	70%	10%	10%	10%
Student	20%	10%	15%	55%
Office	5%	2%	20%	73%

Table 1: The distribution of spending of people in each category.

ii. Now assume we have a customer who told us she is purely interested in office products, and we know she is willing to spend T dollars on merchandise. What would be her spending subtotals on food, movies, art and books & supplies?

iii. Now assume we have a customer who told us she is 90% interested in student products and 10% interested in office products, and we know she is willing to spend T dollars on merchandise. What would be her spending subtotals on food, movies, art and books & supplies?

iv. Now, we go back to our original setting, what we really know about the customer are the values of T_f, T_m, T_a and T_b – spending subtotals on food, movies, art and books & supplies, respectively (and subsequently the total spending T). We would like to solve for the customer's interests c_p, c_f, c_s and c_o – interests in products for party, family, students and office, respectively. Write a system of linear equations, whose solutions will be the customer's interests.

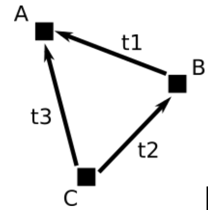
(d) Combine the different parts into a complete algorithm.

(e) Run the algorithm on a customer, Jane Doe, that spent \$6 on food, \$4 on movies, \$1 on art and \$5 on books. With promotions A_1, A_2, A_3 and A_4 targeted at customers with preferences $\vec{s_{A_1}} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$, $\vec{s_{A_2}} = \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{3} \end{bmatrix}$, $\vec{s_{A_3}} = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{5}{2} \\ -\frac{1}{2} \end{bmatrix}$ and $\vec{s_{A_4}} = \begin{bmatrix} 0 \\ \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix}$. Which promotion will be printed on the receipt?

(f) Will there ever be a customer for which the system devised in part (c) will yield no solutions or infinite solutions?

4. Traffic Flows Let's go through the first few parts of HW3 traffic flows problem. Suppose your goal is to measure flow rates of vehicles along roads in Berkeley. However, the city's limited budget prohibits you from placing a traffic sensor along every road. Fortunately, you realize that there is a specific constraint that traffic must obey: the number of cars entering an intersection must equal the number of cars exiting the intersection. We'll see how this constraint helps us determine how many sensors you need for a given set of roads, and where we should place them.

(a) Begin with a loop of road with three intersections, A, B, and C. t_1 cars flow from B to A per hour. t_2 cars flow from C to B per hour. And t_3 cars flow from C to A per hour.



[H!]

Because we have determined that the number of cars in the network is conserved, the total number of cars per hour flowing into each node is zero. For example, at node B, $t_2 - t_1 = 0$. Let's write this constraint as a system of linear equations. We can represent the flows on each road as a vector $\vec{t} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$.

Find the matrix G such that the equation

$$\begin{bmatrix} & & \\ & G & \\ & & \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

represents the constraint that the sum of flows into each node is zero. This matrix is called the *incidence matrix*. What does each row of this matrix represent? What does each column of this matrix represent?

(b) We can place sensors on a road to measure the flow through it. But, as we mentioned earlier, the budget is limited. Our goal is to figure out the minimum number of sensors needed to measure flow along every road.

Suppose for the network above we have one sensor reading, $t_1 = 10$. Is it possible to calculate the flow rates t_2 and t_3?

5. Span

If you have extra time, determine whether b lies in the span of A.

For the following problems determine whether b lies in the span of A.

(a) $A = \begin{bmatrix} 8 & 12 & 16 \\ 12 & 5 & 7 \\ 16 & 8 & 9 \end{bmatrix}, b = \begin{bmatrix} 4 \\ 8 \\ 9 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}, b = \begin{bmatrix} 8 \\ 7 \\ 6 \end{bmatrix}$

(c) $A = \begin{bmatrix} 1 & 1 & -6 \\ 1 & -2 & 3 \\ 6 & -3 & -9 \end{bmatrix}, b = \begin{bmatrix} -3 \\ 0 \\ 3 \end{bmatrix}$