1. Exploring Nullspaces

- The **column space** of a matrix is the **range** or possible outputs of a transformation/function/linear operation. It is also the **span** of the vectors that form the columns of the matrix.
- The **nullspace** is the set of input vectors that output a zero vector.

For the following five matrices, answer the following questions:

i. What is the column space of $\mathbf{A}$? What is its dimension?

ii. What is the nullspace of $\mathbf{A}$? What is its dimension?

iii. Are the column spaces of the row reduced matrix $\mathbf{A}$ and the original matrix $\mathbf{A}$ the same?

iv. Do the columns of $\mathbf{A}$ form a basis of $\mathbb{R}^2$ (or $\mathbb{R}^3$ for part (e))? 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) $\begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 3 \\ 0 & -1 & -2 \end{bmatrix}$

2. Retail Store Marketing

Intro The retail store EehEeh Sixteen would like to design a smart register that prints a promotion to each customer when they check out, depending on the things they may be interested in. The problem is that the register doesn't know what the customer's interests are, and the only data it has about the customer are their 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 their interests in (i) party products; (ii) family products; (iii) student products and (iv) office products.
- At checkout 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 to print on the receipt for the customer.

Your Job We will try do design an algorithm the register can use to print out the smart promotion. Let's break the problem down into smaller components, design each small component individually, and then combine them all together into a complete algorithm.

(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 $\vec{x_c}$
- A similarity function $sim(\vec{x_c}, \vec{s_A})$ that returns a scalar comparing a given customer's interests $\vec{x_c}$ and a promotion's target customer $\vec{s_A}$. The scalar value returned by $sim(\vec{x_c}, \vec{s_A})$ is higher the more aligned the customer c is with the promotion A (similarity score).

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

(b) Since the similarity function is not given to us, we will design one ourselves.

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 subtotals T_f, T_m, T_a and T_b to infer their interests $\vec{x_c} = \begin{bmatrix} c_{\text{party}} \\ c_{\text{family}} \\ c_{\text{student}} \\ c_{\text{office}} \end{bmatrix}$. Table **??** describes the spending subtotals of a customer with only one interest (either party products, family products, student products, or office products).

	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 spending habits of customers in each category

Furthermore, you can assume that a customer with more than one interest spends their money proportionally to the percentages given in Table **??**.

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 their spending habits in food, movies, art and books & supplies.

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

ii. Now assume we have a customer who told us that they are 90% interested in student products and 10% interested in office products and that they are willing to spend T dollars on merchandise. What would be their spending subtotals on food, movies, art and books & supplies?

iii. Now, we go back to our original setting: we only know the values T_f, T_m, T_a and T_b of the customer – the 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. Represent this system using a matrix vector product.

(d) Will there ever be a customer for which the system devised in part **??** yields no solutions or infinite solutions?

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