EECS 16A Designing Information Devices and Systems I
Spring 2020 Homework 4

This homework is due Friday February 21, 2020, at 23:59.
Self-grades are due Monday February 24, 2020, at 23:59.

Submission Format
Your homework submission should consist of **one** file.

- `hw4.pdf`: A single PDF file that contains all of your answers (any handwritten answers should be scanned) as well as your IPython notebook saved as a PDF.

  If you do not attach a PDF "printout" of your IPython notebook, you will not receive credit for problems that involve coding. Make sure that your results and your plots are visible. Assign the IPython printout to the correct problem(s) on Gradescope.

- Practice problems are not graded, but are to help with learning.

Submit the file to the appropriate assignment on Gradescope.

1. Finding Null Spaces and Column Spaces

Learning Objectives*: Null spaces and column spaces are two fundamental vector spaces associated with matrices and they describe important attributes of the transformations that these matrices represent. This problem explores how to find and express these spaces.*

Definition (Null space): The null space of a matrix, $A \in \mathbb{R}^{m \times n}$, is the set of all vectors $\vec{x} \in \mathbb{R}^n$ such that $A\vec{x} = \vec{0}$. The null space is notated as $N(A)$ and the definition can be written in set notation as:

$$N(A) = \{\vec{x} \mid A\vec{x} = \vec{0}, \vec{x} \in \mathbb{R}^n\}$$

Definition (Column space): The column space of a matrix, $A \in \mathbb{R}^{m \times n}$, is the set of all vectors $A\vec{x} \in \mathbb{R}^m$ for all choices of $\vec{x} \in \mathbb{R}^n$. Equivalently, it is also the span of the set of A's columns. The column space can be notated as $\mathrm{C}(A)$ or $\mathrm{range}(A)$ and the definition can be written in set notation as:

$$\mathrm{C}(A) = \{A\vec{x} \mid \vec{x} \in \mathbb{R}^n\}$$

(a) Consider matrices in $\mathbb{R}^{3 \times 5}$. What is the maximum possible number of linearly independent column vectors?

(b) You are given the following matrix $\mathbf{A}$.

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & -2 & 3 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Find a set of vectors that span the column space of $\mathbf{A}$. What is the minimum number of vectors required to span the column space of $\mathbf{A}$? (This is the dimension of the column space of $\mathbf{A}$.)

(c) The dimension of the null space is the minimum number of vectors needed to span it. Find a set of vectors that span the null space of $\mathbf{A}$ (the matrix from part (b)). What is the dimension of the null space of $\mathbf{A}$?

(d) What do you notice about the sum of the dimensions of the null space and the column space in relation to the dimensions of $\mathbf{A}$?

(e) (**Practice**) Now consider the new matrix, $\mathbf{B} = \mathbf{A}^T$,

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & -1 & 0 \\ 3 & 1 & 0 \end{bmatrix}$$

Find a set of vectors that span the column space of $\mathbf{B}$. What is the minimum number of vectors required to span the column space of $\mathbf{B}$?

(f) **(Practice)** Find a set of vectors that spans the null space of the following matrix. This problem requires systematic calculations, but is helpful if you want more practice.

$$\mathbf{C} = \begin{bmatrix} 2 & -4 & 4 & 8 \\ 1 & -2 & 3 & 6 \\ 2 & -4 & 5 & 10 \\ 3 & -6 & 7 & 14 \end{bmatrix}$$

(g) **(Practice)** Find the column space and its dimension, and the nullspace and its dimension of the following matrix.

$$\mathbf{D} = \begin{bmatrix} 1 & -1 & -3 & 4 \\ 3 & -3 & -5 & 8 \\ 1 & -1 & -1 & 2 \end{bmatrix}$$

2. Cubic Polynomials

Learning Goal: *This problem shows us that we can treat fixed-degree polynomials as a vector space. Furthermore, many operations on polynomials are linear operations in this vector space and can be represented by matrices.*

(a) Show that the set of all cubic polynomials

$$p(t) = p_0 + p_1 t + p_2 t^2 + p_3 t^3,$$

where $t \in [a,b]$ and the coefficients p_k are real scalars, forms a vector space. Call this vector space V.

(b) Consider the set of real-valued monomials given below:

$$\varphi_0(t) = 1, \quad \varphi_1(t) = t, \quad \varphi_2(t) = t^2, \quad \varphi_3(t) = t^3,$$

where $t \in \mathbb{R}$.
Show that every real-valued cubic polynomial

$$p(t) = p_0 + p_1 t + p_2 t^2 + p_3 t^3$$

defined over the interval $[a,b]$ can be written as a linear combination of the monomials $\varphi_0(t)$, $\varphi_1(t)$, $\varphi_2(t)$, and $\varphi_3(t)$. In particular, show that

$$p(t) = \vec{c}^T \vec{\varphi}(t),$$

where

$$\vec{c}^T = \begin{bmatrix} c_0 & c_1 & c_2 & c_3 \end{bmatrix}$$

is a vector of appropriately chosen coefficients and

$$\vec{\varphi}(t) = \begin{bmatrix} \varphi_0(t) \\ \varphi_1(t) \\ \varphi_2(t) \\ \varphi_3(t) \end{bmatrix} = \begin{bmatrix} 1 \\ t \\ t^2 \\ t^3 \end{bmatrix}.$$

(c) The monomials $\varphi_k(t) = t^k$, for $k = 0,1,2,3$, constitute a basis for the vector space of real-valued cubic polynomials defined over the interval $[a,b]$. Justify why this is true. What is the dimension of V?

(d) Express the derivatives of the basis polynomials $\varphi_i(t)$ for $i = 0,1,2,3$ in terms of the $\varphi_i(t)$ for $i = 0,1,2,3$.

(e) Let **D** be a 4x4 matrix. Use the previous part to help you find the entries of **D**, such that for any polynomial

$$p(t) = \vec{c}^T \vec{\varphi}(t),$$

its derivative can be expressed as

$$\frac{\mathrm{d}}{\mathrm{d}t} p(t) = (D\vec{c})^T \vec{\varphi}(t).$$

Hint: What are the dimensions of $(\mathbf{D}\vec{c})^T$? (Reminder: The dimensions of a matrix or vector is a different concept than the dimensions of a vector space).

(f) **(Practice):** A curve is a continuous mapping from the real line to $\mathbb{R}^N$. A cubic Bézier curve—used extensively in computer graphics—is a type of curve that uses as a basis the following special subset of what are more broadly known as *Bernstein polynomials*:

$$\beta_0(t) = (1-t)^3, \quad \beta_1(t) = 3t(1-t)^2, \quad \beta_2(t) = 3t^2(1-t), \text{ and } \beta_3(t) = t^3.$$

Prove that the Bernstein polynomials $\beta_k(t)$ defined above form a basis for the space of cubic polynomials. To do this, show that any real-valued polynomial

$$p(t) = p_0 + p_1 t + p_2 t^2 + p_3 t^3$$

can be expressed as a linear combination of the polynomials $\beta_k(t)$ and determine the coefficients in that linear combination. In particular, determine the coefficients in the expansion

$$p(t) = \widehat{p}_0 \beta_0(t) + \widehat{p}_1 \beta_1(t) + \widehat{p}_2 \beta_2(t) + \widehat{p}_3 \beta_3(t).$$

Hint: Determine a matrix **R**, such that

$$\vec{\beta}(t) = \mathbf{R}\vec{\varphi}(t),$$

where

$$\vec{\beta}(t) = \begin{bmatrix} \beta_0(t) \\ \beta_1(t) \\ \beta_2(t) \\ \beta_3(t) \end{bmatrix} = \begin{bmatrix} (1-t)^3 \\ 3t(1-t)^2 \\ 3t^2(1-t) \\ t^3 \end{bmatrix}.$$

Without solving for the inverse, show that **R** is invertible. Determine its inverse $\mathbf{R}^{-1}$, from which you can determine the coefficients $\widehat{p}_k$. You may use IPython to find $\mathbf{R}^{-1}$.

3. Introduction to Eigenvalues and Eigenvectors

Learning Goal: *Practice algorithmic computation of eigenvalues and eigenvectors. The importance of eigenvalues and eigenvectors will become clear in the following problems.*

For each of the following matrices, find their eigenvalues and the corresponding eigenvectors. For simple matrices, you may do this my inspection if you prefer.

(a) $\begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}$

(b) $\begin{bmatrix} 22 & 6 \\ 6 & 13 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

(d) Let $A \in \mathbb{R}^{n \times n}$ be a general square matrix. Show that the set of eigenvectors corresponding to a particular eigenvalue of this matrix

$$\{\vec{x} \in \mathbb{R}^n : A\vec{x} = \lambda \vec{x}, \lambda \in \mathbb{R}\}$$

is a subspace.

4. Operations on Subspaces

Let $\mathbb{V}$ be a vector space with subspaces $\mathbb{S}$ and $\mathbb{T}$.

(a) Consider the following set of vectors: $\mathbb{S} + \mathbb{T} := \{\vec{s} + \vec{t} \mid \vec{s} \in \mathbb{S}, \vec{t} \in \mathbb{T}\}$ Show that $\mathbb{S} + \mathbb{T}$ is a subspace of $\mathbb{V}$.

(b) Consider the intersection of $\mathbb{S}$ and $\mathbb{T}$: $\mathbb{S} \cap \mathbb{T} = \{\vec{v} \in \mathbb{V} \mid \vec{v} \in \mathbb{S}, \vec{v} \in \mathbb{T}\}$. Show that this is a subspace.

(c) As before, $\mathbb{S}$ and $\mathbb{T}$ are subspaces of the vector space $\mathbb{V}$. Now, assume $\mathbb{S} \cap \mathbb{T} = \{\vec{0}\}$. Let $B_S = \{\vec{s}_1, \vec{s}_2, \ldots, \vec{s}_k\}$ be a basis for $\mathbb{S}$, and $B_T = \{\vec{t}_1, \vec{t}_2, \ldots, \vec{t}_n\}$ be a basis for $\mathbb{T}$.
Show that $B_S \cup B_T = \{\vec{s}_1, \vec{s}_2, \ldots, \vec{s}_k, \vec{t}_1, \vec{t}_2, \ldots, \vec{t}_n\}$ is a basis for $\mathbb{S} + \mathbb{T}$.

(d) If $\mathbb{S} \cap \mathbb{T}$ is $\{\vec{0}\}$, what is the dimension of $\mathbb{S} + \mathbb{T}$ in terms of the dimensions of $\mathbb{S}$ and $\mathbb{T}$?

(e) (**Practice**) For arbitrary subspaces $\mathbb{S}$ and $\mathbb{T}$ of vector space $\mathbb{V}$, show that $\dim(\mathbb{S} + \mathbb{T}) = \dim(\mathbb{S}) + \dim(\mathbb{T}) - \dim(\mathbb{S} \cap \mathbb{T})$.

5. Image Compression

In this question, we explore how eigenvalues and eigenvectors can be used for image compression. A grayscale image can be represented as a data grid. Say a symmetric, square image is represented by a symmetric matrix $\mathbf{A}$, such that $\mathbf{A}^T = \mathbf{A}$. We can transform the images to vectors to make it easier to process them as data, but here, we will understand them as 2D data. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $\mathbf{A}$ with corresponding eigenvectors $\vec{v}_1, \ldots, \vec{v}_n$. Also, let these eigenvectors be normalized (unit norm). Then, the matrix can be represented as the expansion

$$\mathbf{A} = \lambda_1 \vec{v}_1 \vec{v}_1^T + \lambda_2 \vec{v}_2 \vec{v}_2^T + \cdots + \lambda_n \vec{v}_n \vec{v}_n^T.$$

This is known as the *spectral decomposition* of $\mathbf{A}$. Note that the eigenvectors must be normalized for this expansion to be valid because we know that if $\vec{v}_i$ is an eigenvector, then any scalar multiple $\alpha \vec{v}_i$ is also an eigenvector. If we scaled every eigenvector on the right hand side of the equation by α, then the left hand side would change from $\mathbf{A}$ to $\alpha^2 \mathbf{A}$.

The previous expansion shows that the matrix $\mathbf{A}$ can be synthesized by its n eigenvalues and eigenvectors. However, $\mathbf{A}$ can also be *approximated* with the k largest eigenvalues and the corresponding eigenvectors. That is,

$$\mathbf{A} \approx \lambda_1 \vec{v}_1 \vec{v}_1^T + \lambda_2 \vec{v}_2 \vec{v}_2^T + \cdots + \lambda_k \vec{v}_k \vec{v}_k^T.$$

(a) Use the IPython notebook `prob4.ipynb` and the image file `pattern.npy`. Run the associated code block, which computes the eigenvalues and eigenvectors of the image `pattern.npy` and then sorts them in descending order. Note that `numpy.linalg.eig` returns normalized eigenvectors by default. Mathematically, how many eigenvectors are required to fully capture the information within the image?

(b) In the IPython notebook, find an approximation for the image using the 100 largest eigenvalues and eigenvectors. Does the approximate image capture most of the features of the original image?

(c) Repeat part (b) with $k = 50$. By further experimenting with the code, what seems to be the lowest value of k that retains most of the salient features of the given image?

6. Traffic Flows

Learning Objective: *The learning objective of this problem is to see how the concept of nullspaces can be applied to flow problems.*

Your goal is to measure the flow rates of vehicles along roads in a town. It is prohibitively (too) expensive to place a traffic sensor along every road. You realize, however, that the number of cars flowing into an intersection must equal the number of cars flowing out. You can use this "flow conservation" to determine the traffic along all roads in a network by measuring the flow along only some roads. In this problem, we will explore this concept.

(a) Let's begin with a network with three intersections, A, B and C. Define the flow t_1 as the rate of cars (cars/hour) on the road between B and A, flow t_2 as the rate on the road between C and B, and flow t_3 as the rate on the road between C and A.

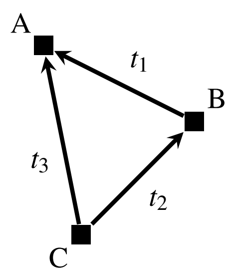

Figure 1: A simple road network.

(Note: The directions of the arrows in the figure are the way that we define positive flow by convention. For example, if there were 100 cars per hour traveling from A to C, then $t_3 = -100$. The flows now are not fractions of water in reservoirs as in the pumps setting, but numbers of cars.)

We assume the "flow conservation" constraints: the net number of cars per hour flowing into each intersection is zero. For example at intersection B, we have the constraint $t_2 - t_1 = 0$. The full set of constraints (one per intersection) is:

$$\begin{cases} t_1 + t_3 = 0 \\ t_2 - t_1 = 0 \\ -t_3 - t_2 = 0 \end{cases}$$

As mentioned earlier, we can place sensors on a road to measure the flow through it, but we have a limited budget, and we would like to determine all of the flows with the smallest possible number of sensors.

Suppose for the network above we have one sensor reading, $t_1 = 10$. Can we figure out the flows along the other roads? (That is, the values of t_2 and t_3). If we can, find the values of t_2 and t_3.

(b) Now suppose we have a larger network, as shown in Figure 2.

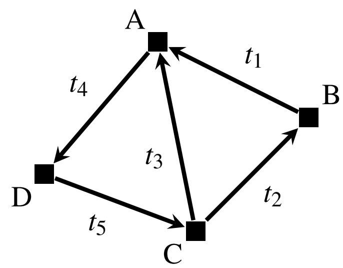

Figure 2: A larger road network.

We would again like to determine the traffic flows on all roads, using measurements from some sensors. A Berkeley student claims that we need two sensors placed on the roads CA (measuring t_3) and DC (measuring t_5). A Stanford student claims that we need two sensors placed on the roads CB (measuring t_2) and BA (measuring t_1). Write out the system of linear equations that represents this flow graph. Is it possible to determine all traffic flows, $\left[t_1, t_2, t_3, t_4, t_5\right]^T$, with the Berkeley student's suggestion? How about the Stanford student's suggestion?

(c) We would like a more general way of determining the possible traffic flows in a network. Suppose we write the traffic flow on all roads as a vector $\vec{t} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \end{bmatrix}$. As a first step, let us try to write all the flow conservation constraints (one per intersection) as a matrix equation.

Construct a 4×5 matrix $\mathbf{B}$ such that the equation $\mathbf{B}\vec{t} = \vec{0}$:

$$\begin{bmatrix} & & \\ & \mathbf{B} & \\ & & \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

represents the flow conservation constraints for the network in Figure 2.

Hint: Each row is the constraint of an intersection. You can construct $\mathbf{B}$ using only $0, 1$, and -1 entries. This matrix is called the **incidence matrix**. What constraint does each column of $\mathbf{B}$ represent?

(d) Again, suppose we write the traffic flow on all roads as a vector $\vec{t} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \end{bmatrix}$. Then, determine the subspace of all valid traffic flows for the network of Figure 2. Notice that the set of all vectors $\vec{t}$ that satisfy $\mathbf{B}\vec{t} = \vec{0}$

is exactly the null space of the matrix $\mathbf{B}$. That is, we can find all valid traffic flows by computing the null space of $\mathbf{B}$. What is the dimension of the nullspace?

(e) Notice that we can represent the Berkeley student's measurement as $\mathbf{M}_B\vec{t}$, where:

$$\mathbf{M}_B\vec{t} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \vec{t} = \begin{bmatrix} t_3 \\ t_5 \end{bmatrix}$$

Write a matrix $\mathbf{M}_S$ that can be used to represent the Stanford student's measurement.

(f) Now let us analyze more general road networks. Say there is a road network graph G, with incidence matrix $\mathbf{B}_G$. If $\mathbf{B}_G$ has a k-dimensional null space, does this mean measuring the flows along ***any*** k **roads** is always sufficient to recover all of the true flows? Prove or give a counterexample.

Hint: Consider the Stanford student from part (b).

(g) (**Practice**) Assume that $\vec{u}$ and $\vec{t}$ are distinct valid flows, that is $\mathbf{B}_G\vec{u} = \mathbf{B}_G\vec{t} = \vec{0}$. Can you recover all of the network's true flows if $(\vec{u} - \vec{t})$ belongs to the nullspace of $\mathbf{M}_S$?

Clarification: A âĂIJvalidâĂİ flow is one that is possible without violating the constraints on the nodes (so flow in must equal to flow out). There may be many valid flows, but only one âĂIJtrueâĂİ flow.

(h) (**Challenge: Practice**) If the incidence matrix $\mathbf{B}_G$ has a k-dimensional null space, does this mean we can **always pick a set of** k **roads** such that measuring the flows along these roads is sufficient to recover the exact flows? Prove or give a counterexample.

7. Homework Process and Study Group

Who else did you work with on this homework? List names and student ID's. (In case of homework party, you can also just describe the group.) How did you work on this homework?