Homework 01

1. Course Setup

Please complete the following steps to get access to all course resources.

(a) Visit the course website at `http://eecs127.github.io/` and familiarize yourself with the syllabus.

(b) Verify that you can access the class Ed site at `https://edstem.org/us/courses/93760/discussion`.

(c) Verify that you can access the class Gradescope site at `https://www.gradescope.com/courses/1228786`.

2. What Prerequisites Have You Taken?

The prerequisites for this course are

- MATH 54 (Linear Algebra & Differential Equations),
- CS 70 (Discrete Mathematics & Probability Theory), and
- MATH 53 (Multivariable Calculus).

Please fill out the following Google form: `https://forms.gle/gLgvN769ZxoDNzn59` to tell us which of these courses, or their equivalents, you have taken. If you are unsure of course material overlap, please refer to the MATH 54, CS 70 websites (`http://www.sp22.eecs70.org/`, `https://lin-lin.github.io/MATH54/`), and the MATH 53 textbook (*Multivariable Calculus* by James Stewart). **For the response to this question, write the secret word revealed at the end of the form.**

The course material this semester will rely on knowledge from these prerequisite courses. If you feel shaky on this material, please use the first week to reacquaint yourself with it. We expect you to handle this review on your own; we will not prioritize questions about prerequisite material in office hours.

Solution: The secret word is: `PinkpantheR`

3. Orthogonality

Let $\vec{x}, \vec{y} \in \mathbb{R}^n$ be two linearly independent unit-norm vectors; that is, $\|\vec{x}\|_2 = \|\vec{y}\|_2 = 1$.

(a) Show that the vectors $\vec{u} = \vec{x} - \vec{y}$ and $\vec{v} = \vec{x} + \vec{y}$ are orthogonal.

Solution: Orthogonal means dot product is 0. When x, y are both unit-norm, we have

$$(\vec{x} - \vec{y})^\top (\vec{x} + \vec{y}) = \vec{x}^\top \vec{x} - \vec{y}^\top \vec{y} - \vec{y}^\top \vec{x} + \vec{x}^\top \vec{y} = \vec{x}^\top \vec{x} - \vec{y}^\top \vec{y} = \|\vec{x}\|_2^2 - \|\vec{y}\|_2^2 = 1 - 1 = 0. \quad (1)$$

(b) Find an orthonormal basis for $\operatorname{span}(\vec{x}, \vec{y})$, the subspace spanned by $\vec{x}$ and $\vec{y}$.

Solution:

We seek an orthonormal basis for $S = \operatorname{span}(\vec{x}, \vec{y})$. Since $\vec{x}, \vec{y}$ are linearly independent, we know $\dim(S) = 2$. We present two main approaches.

Approach 1: Symmetric Basis (using sum and difference) From Part (a), we know that $\vec{u} = \vec{x} - \vec{y}$ and $\vec{v} = \vec{x} + \vec{y}$ are orthogonal. To show they form a basis for S, we must verify that $\operatorname{span}(\vec{u}, \vec{v}) = \operatorname{span}(\vec{x}, \vec{y})$. We can justify this in three ways:

- **Justification A:** Let $\vec{z} \in \operatorname{span}(\vec{x}, \vec{y})$, so $\vec{z} = \lambda\vec{x} + \mu\vec{y}$. We can rewrite this in terms of $\vec{u}$ and $\vec{v}$ as $\vec{z} = \alpha\vec{u} + \beta\vec{v}$ by solving for α, β:

$$\alpha = \frac{\lambda - \mu}{2}, \quad \beta = \frac{\lambda + \mu}{2}. \quad (2)$$

Thus, any vector in $\operatorname{span}(\vec{x}, \vec{y})$ is in $\operatorname{span}(\vec{u}, \vec{v})$.

- **Justification B:** We can express $\vec{x}$ and $\vec{y}$ directly in terms of $\vec{u}$ and $\vec{v}$:

$$\vec{x} = \frac{\vec{u} + \vec{v}}{2}, \quad \vec{y} = \frac{\vec{v} - \vec{u}}{2}. \quad (3)$$

Since $\vec{x}, \vec{y}$ are linear combinations of $\vec{u}, \vec{v}$ (and vice versa), their spans are identical.

- **Justification C:** Since $\vec{u}$ and $\vec{v}$ are orthogonal (and non-zero), they are linearly independent. Thus, $\operatorname{span}(\vec{u}, \vec{v})$ is a 2-dimensional subspace. Since $\vec{u}, \vec{v}$ consist of linear combinations of $\vec{x}, \vec{y}$, we know $\operatorname{span}(\vec{u}, \vec{v}) \subseteq \operatorname{span}(\vec{x}, \vec{y})$. Since $\operatorname{span}(\vec{x}, \vec{y})$ is also 2-dimensional, the subspace must be equal to itself.

Conclusion: The vectors $\vec{u}, \vec{v}$ form an orthogonal basis. To get an *orthonormal* basis, we normalize them:

$$\left(\frac{\vec{x} - \vec{y}}{\|\vec{x} - \vec{y}\|_2}, \frac{\vec{x} + \vec{y}}{\|\vec{x} + \vec{y}\|_2} \right). \quad (4)$$

Approach 2: Gram-Schmidt Process Alternatively, we can construct an orthonormal basis from $(\vec{x}, \vec{y})$ directly using the Gram-Schmidt algorithm.

i. Set the first basis vector as $\vec{z}_1 = \vec{x}$ (since $\|\vec{x}\|_2 = 1$).

ii. Construct the second vector $\vec{z}_2$ by projecting $\vec{y}$ onto $\vec{x}$, removing that component, and normalizing:

$$\vec{z}_2 = \frac{\vec{y} - (\vec{y}^\top \vec{x})\vec{x}}{\|\vec{y} - (\vec{y}^\top \vec{x})\vec{x}\|_2}. \quad (5)$$

The set $(\vec{z}_1, \vec{z}_2)$ forms a valid orthonormal basis.

4. Least Squares

The Michaelis-Menten model for enzyme kinetics relates the rate y of an enzymatic reaction to the concentration x of a substrate, as follows:

$$y = \frac{\beta_1 x}{\beta_2 + x}, \tag{6}$$

for constants $\beta_1, \beta_2 > 0$.

(a) Show that the model can be expressed as a linear relation between the values $1/y = y^{-1}$ and $1/x = x^{-1}$. Specifically, give an equation of the form $y^{-1} = w_1 + w_2 x^{-1}$, specifying the values of w_1 and w_2 in terms of β_1 and β_2.

Solution: Inverting each side of the equation, we have

$$\begin{aligned} y^{-1} &= \left(\frac{\beta_1 x}{\beta_2 + x}\right)^{-1} && (7)\\ &= \frac{\beta_2 + x}{\beta_1 x} && (8)\\ &= \frac{\beta_2}{\beta_1 x} + \frac{x}{\beta_1 x} && (9)\\ &= \frac{\beta_2}{\beta_1} x^{-1} + \frac{1}{\beta_1} && (10)\\ &= \frac{1}{\beta_1} + \frac{\beta_2}{\beta_1} x^{-1}. && (11)\\ & && (12) \end{aligned}$$

The above equation has exactly the desired form $y^{-1} = w_1 + w_2 x^{-1}$ for $w_1 = \frac{1}{\beta_1}$ and $w_2 = \frac{\beta_2}{\beta_1}$.

(b) In general, reaction parameters β_1 and β_2 (and, thus, w_1 and w_2) are not known a priori and must be fit from data — for example, using least squares. Suppose you collect m measurements (x_i, y_i), $i = 1, \ldots, m$ over the course of a reaction. Formulate the least squares problem

$$\vec{w}^\star = \underset{\vec{w}}{\operatorname{argmin}} \, \|X\vec{w} - \vec{y}\|_2^2\,, \tag{13}$$

where $\vec{w}^\star = \begin{bmatrix} w_1^\star & w_2^\star \end{bmatrix}^\top$, and you must specify $X \in \mathbb{R}^{m \times 2}$ and $\vec{y} \in \mathbb{R}^m$. Specifically, your solution should include explicit expressions for X and $\vec{y}$ as a function of (x_i, y_i) values and a final expression for $\vec{w}^\star$ in terms of X and $\vec{y}$, which should contain only matrix multiplications, transposes, and inverses. Assume without loss of generality that $x_1 \neq x_2$.

Solution: To formulate the least squares problem as stated, X and $\vec{y}$ values should be set to

$$X = \begin{bmatrix} 1 & \cdots & 1 \\ x_1^{-1} & \cdots & x_m^{-1} \end{bmatrix}^\top, \quad \vec{y} = \begin{bmatrix} y_1^{-1} & \cdots & y_m^{-1} \end{bmatrix}^\top. \tag{14}$$

To solve this least squares problem, we note that the optimal residual vector $X\vec{w}^\star - \vec{y}$ must be orthogonal to $\mathcal{R}(X)$ by the orthogonality principle, and we have

$$X^\top (X\vec{w}^\star - \vec{y}) = 0. \tag{15}$$

Note that $X^T X$ is invertible because $\forall i \neq j, i, j \in \{1, \ldots, m\}, x_i \neq x_j$. Therefore, rearranging we get,

$$\vec{w}^\star = (X^\top X)^{-1} X^\top \vec{y}. \tag{16}$$

(c) Assume that we have used the above procedure to calculate values for $w_1^\star$ and $w_2^\star$, and we now want to estimate $\widehat{\vec{\beta}} = \begin{bmatrix} \widehat{\beta}_1 & \widehat{\beta}_2 \end{bmatrix}^\top$. Write an expression for $\widehat{\vec{\beta}}$ in terms of $w_1^\star$ and $w_2^\star$.

Solution: To calculate $\widehat{\vec{\beta}}$, we can simply reverse the calculations from part 4(a):

$$w_1 = \frac{1}{\beta_1} \implies \beta_1 = \frac{1}{w_1}, \tag{17}$$

$$w_2 = \frac{\beta_2}{\beta_1} \implies \beta_2 = \beta_1 w_2 = \frac{w_2}{w_1}. \tag{18}$$

Thus, $\widehat{\vec{\beta}} = \begin{bmatrix} \frac{1}{w_1^\star} & \frac{w_2^\star}{w_1^\star} \end{bmatrix}^\top$.

NOTE: This problem was taken (with some edits) from the textbook *Optimization Models* by Calafiore and El Ghaoui.

5. Subspaces and Dimensions

Consider the set $\mathcal{S}$ of points $(x_1, x_2, x_3) \in \mathbb{R}^3$ such that

$$x_1 + 2x_2 + 3x_3 = 0, \quad 3x_1 + 2x_2 + x_3 = 0. \tag{19}$$

(a) Find a 2×3 matrix A for which $\mathcal{S}$ is exactly the null space of A.

Solution: Recall the definition of the null space of a matrix A as the set of all vectors $\vec{x}$ such that $A\vec{x} = \vec{0}$. The equations

$$x_1 + 2x_2 + 3x_3 = 0 \tag{20}$$

$$3x_1 + 2x_2 + x_3 = 0 \tag{21}$$

can be written in matrix-vector form as

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{22}$$

The set of $\vec{x} = \begin{bmatrix} x_1, x_2, x_3 \end{bmatrix}^\top$ which satisfy this equation form the null space of $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$. This is the matrix we are looking for.

(b) Determine the dimension of $\mathcal{S}$ and find a basis for it.

Solution: Recall the definitions of a basis and the dimension of a subspace, which are related. A basis for a space is a set of linearly independent vectors that span the space. The dimension of this space is then the number of vectors in the basis.

To find the dimension, we solve the equation and find that any solution to the equations is of the form $x_1 = x_3$, $x_2 = -2x_3$, where x_3 is free. Thus, the solutions are of the form $\begin{bmatrix} 1, -2, 1 \end{bmatrix}^\top u$ for $u \in \mathbb{R}$, and so $\mathcal{S} = \operatorname{span}\left(\begin{bmatrix} 1, -2, 1 \end{bmatrix}^\top\right)$. Hence, the dimension of $\mathcal{S}$ is 1, and a basis for $\mathcal{S}$ is the vector $\begin{bmatrix} 1, -2, 1 \end{bmatrix}^\top$.

6. Vector Spaces and Rank

The *rank* of a $m \times n$ matrix A, $\operatorname{rank}(A)$, is the dimension of its *range*, also called span, and denoted $\mathcal{R}(A) := \{A\vec{x} \, : \, \vec{x} \in \mathbb{R}^n\}$.

(a) Assume that $A \in \mathbb{R}^{m \times n}$ takes the form $A = \vec{u}\vec{v}^\top$, with $\vec{u} \in \mathbb{R}^m$, $\vec{v} \in \mathbb{R}^n$, and $\vec{u}, \vec{v} \neq \vec{0}$. (Note that a matrix of this form is known as a *dyad*.) Find the rank of A.

HINT: Consider the quantity $A\vec{x}$ for arbitrary $\vec{x}$, i.e., what happens when you multiply any vector by matrix A.

Solution: For any $\vec{x} \in \mathbb{R}^n$, we have that $A\vec{x} = \vec{u}\vec{v}^\top\vec{x} = \vec{u}(\vec{v}^\top\vec{x}) = (\vec{v}^\top\vec{x})\vec{u}$. Note that $\vec{v}^\top\vec{x}$ is a scalar that can take on any value depending on choice of $\vec{x}$. Since the range of A is the subspace reachable through any choice of $\vec{x}$, $\mathcal{R}(A)$ is simply the 1-dimensional subspace spanned by $\vec{u}$ (i.e., the line pointing along $\vec{u}$). Since a single vector (namely, $\vec{u}$) spans $\mathcal{R}(A)$, the rank of A is 1.

(b) Show that for arbitrary $A, B \in \mathbb{R}^{m \times n}$,

$$\operatorname{rank}(A + B) \leq \operatorname{rank}(A) + \operatorname{rank}(B), \tag{23}$$

i.e., the rank of the sum of two matrices is less than or equal to the sum of their ranks.

HINT: First, show that $\mathcal{R}(A + B) \subseteq \mathcal{R}(A) + \mathcal{R}(B)$, meaning that any vector in the range of $A + B$ can be expressed as the sum of two vectors, each in the range of A and B, respectively. Remember that for any matrix A, $\mathcal{R}(A)$ is a subspace, and for any two subspaces S_1 and S_2, $\dim(S_1 + S_2) \leq \dim(S_1) + \dim(S_2)$.[1] (Note that the sum of vector spaces $S_1 + S_2$ is not equivalent to $S_1 \cup S_2$, but is defined as $S_1 + S_2 := \{\vec{s_1} + \vec{s_2} | \vec{s_1} \in S_1, \vec{s_2} \in S_2\}$.)

Solution: Given any vector $\vec{v} \in \mathcal{R}(A + B)$, there must by definition exist $\vec{x} \in \mathbb{R}^n$ such that $\vec{v} = (A+B)\vec{x}$. Thus, $\vec{v} = (A + B)\vec{x} = \underbrace{A\vec{x}}_{\in \mathcal{R}(A)} + \underbrace{B\vec{x}}_{\in \mathcal{R}(B)}$, so $\mathcal{R}(A + B) \subseteq \mathcal{R}(A) + \mathcal{R}(B)$, as hinted.

Computing the dimension of each side of the subset relationship, it follows that

$$\dim(\mathcal{R}(A + B)) \leq \dim(\mathcal{R}(A) + \mathcal{R}(B)). \tag{24}$$

Using the second part of the hint, we have that

$$\dim(\mathcal{R}(A) + \mathcal{R}(B)) \leq \dim(\mathcal{R}(A)) + \dim(\mathcal{R}(B)). \tag{25}$$

Combining the previous two equations,

$$\dim(\mathcal{R}(A + B)) \leq \dim(\mathcal{R}(A)) + \dim(\mathcal{R}(B)), \tag{26}$$

i.e., by definition,

$$\operatorname{rank}(A + B) \leq \operatorname{rank}(A) + \operatorname{rank}(B) \tag{27}$$

as desired.

(c) Consider an $m \times n$ matrix A that takes the form $A = UV^\top$, with $U \in \mathbb{R}^{m \times k}$, $V \in \mathbb{R}^{n \times k}$. Show that the rank of A is less than or equal to k. *HINT: Use parts 6(a) and 6(b), and remember that this decomposition*

[1] This fact can be proved by taking a basis of S_1 and extending it to a basis of S_2 (during which we can only add *at most* $\dim(S_2)$ basis vectors). This extended basis must now also be a basis of $S_1 + S_2$. Thus, $\dim(S_1 + S_2) \leq \dim(S_1) + \dim(S_2)$.

can also be written as the dyadic expansion

$$A = UV^\top = \begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_k \end{bmatrix} \begin{bmatrix} \vec{v}_1^\top \\ \vdots \\ \vec{v}_k^\top \end{bmatrix} = \sum_{i=1}^{k} \vec{u}_i \vec{v}_i^\top, \tag{28}$$

for $U = \begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_k \end{bmatrix}$ *and* $V = \begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_k \end{bmatrix}$.

Solution: Starting with the dyadic expansion above, iteratively pulling out terms from this summation, and using the result from 6(a) that the rank of a dyadic matrix is 1 (or 0, if any $\vec{v}_i = \vec{0}$), we know by the rank relation from 6(b) that

$$\operatorname{rank}(A) = \operatorname{rank}\left(\sum_{i=1}^{k} \vec{u}_i \vec{v}_i^\top\right) \leq \operatorname{rank}\left(\sum_{i=1}^{k-1} \vec{u}_i \vec{v}_i^\top\right) + \underbrace{\operatorname{rank}\left(\vec{u}_k \vec{v}_k^\top\right)}_{\text{0 or 1}} \leq \dots \leq k, \tag{29}$$

as desired.

7. Homework Process

With whom did you work on this homework? List the names and SIDs of your group members.

NOTE: If you didn't work with anyone, you can put "none" as your answer.

