Homework 3: Vectors and the Dot Product

EECS 245, Fall 2025 at the University of Michigan

due Tuesday, September 16th, 2025 at 11:59PM Ann Arbor Time

Write your solutions to the following problems by either typing them up or handwriting them on another piece of paper. Homeworks are due to Gradescope by 11:59PM on the due date. See the **syllabus** for details on the slip day policy.

Homework will be evaluated not only on the correctness of your answers, but on your ability to present your ideas clearly and logically. You should always explain and justify your conclusions, using sound reasoning. Your goal should be to convince the reader of your assertions. If a question does not require explanation, it will be explicitly stated.

Before proceeding, make sure you're familiar with the **collaboration policy**.

Total Points: 14 + 9 + 7 + 15 + 12 + 6 = 63

Problem 1: Parallelogram Law (14 pts)

a) (4 pts) Let $\vec{u} = \begin{bmatrix} 3 \\ -6 \\ 0 \\ 2 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 2 \\ 1 \\ 4 \\ -2 \end{bmatrix}$. Compute the following quantities:

(i) $\|\vec{u}\|$

(ii) $\|\vec{v}\|$

(iii) $\|\vec{u} + \vec{v}\|$

(iv) $\|\vec{u} - \vec{v}\|$

b) (3 pts) Using the same vectors as in part **a)**, compute the angle between $\vec{u}$ and $\vec{v}$. Leave your answer in terms of $\cos^{-1}$.

c) (4 pts) Now, suppose $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ are any two vectors in $\mathbb{R}^n$. Prove that:

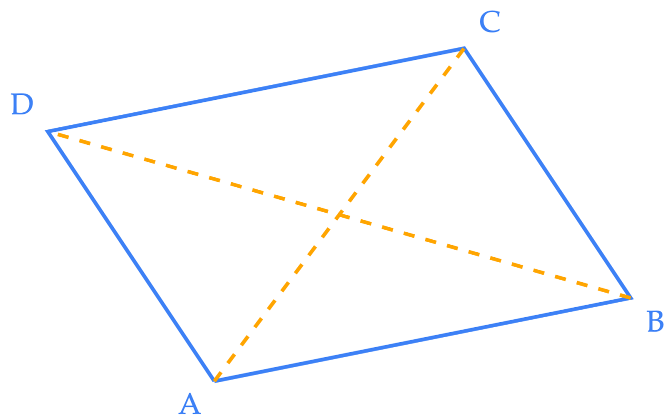
$$\|\vec{u} + \vec{v}\|^2 + \|\vec{u} - \vec{v}\|^2 = 2\|\vec{u}\|^2 + 2\|\vec{v}\|^2$$

The statement above is called the **parallelogram law** of vectors.

*Hint: The point of part **a)** was to give you a feel for which quantities are involved in this statement. Your proof should not use these values in particular. Instead, **start with the left-hand side** of the equation and use the properties of the dot product introduced in Chapter 2.2.*

d) (3 pts) Why is the equality from part **c)** called the parallelogram law? Let's explore.

Suppose points A, B, C, and D in $\mathbb{R}^n$ form a parallelogram: a polygon with four sides where opposite sides are parallel and equal in length.



Using the results of the previous part of this problem, prove that the sum of the squares of the side lengths of the parallelogram is equal to the sum of the squares of the diagonals. In other words, prove that:

$$(AB)^2 + (BC)^2 + (CD)^2 + (DA)^2 = (AC)^2 + (BD)^2$$

where AB represents the length of the segment from point A to point B, etc.

Some guidance: Define two vectors, $\vec{u}$ and $\vec{v}$, and explain why the result from the previous part of this problem implies the desired equality.

Problem 2: Linear Combinations (9 pts)

As we saw in Chapter 2.1, a **linear combination** of vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_d \in \mathbb{R}^n$ is a vector of the form

$$a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_d\vec{v}_d$$

where $a_1, a_2, \ldots, a_d$ are scalars.

Much of our study of linear algebra involves understanding **the set of possible linear combinations** of a given set of vectors. As the notes detail, our multiple linear regression problem boils down to finding the best possible linear combination of the features, so it's important that we understand how linear combinations work.

Let $\vec{v}_1 = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} 0 \\ 5 \\ 2 \end{bmatrix}$, and $\vec{x} = \begin{bmatrix} -6 \\ 1 \\ 4 \end{bmatrix}$.

You can find an interactive, three-dimensional visualization of these four vectors at this link:

https://eecs245.org/resources/homeworks/hw03/hw03-problem-2.html

We recommend you have this visual open while you work through this problem.

a) (4 pts) Find constants a, b, and c such that

$$a\vec{v}_1 + b\vec{v}_2 + c\vec{v}_3 = \vec{x}$$

In other words, write $\vec{x}$ as a linear combination of $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$.
Hint: Start by writing out the equation as a system of equations. Then, use your favorite method for solving systems of equations to find a, b, and c.

b) (2 pts) Try and find constants d and e such that

$$d\vec{v}_1 + e\vec{v}_3 = \vec{x}$$

If you are able to find constants d and e, **explain why**, even though there are two unknowns but three equations for them. If you are unable to find constants d and e, **explain why** no solution exists.

c) (3 pts) Try and find constants p and q such that

$$p\vec{v}_1 + q\vec{v}_2 = \vec{x}$$

If you are able to find constants p and q, **explain why**, even though there are two unknowns but three equations for them. If you are unable to find constants p and q, **explain why** no solution exists.

Problem 3: Correlation (7 pts)

In **Chapter 1.4**, you were told that the correlation coefficient, r, ranges between -1 and 1, where -1 implies a perfect negative linear association and 1 implies a perfect positive linear association. However, you were never given a proof of the fact that $-1 \leq r \leq 1$.

Here, you will prove this fact, given your newfound understanding of vectors, the dot product, and angles.

a) (2 pts) Let $\vec{x}$ and $\vec{y}$ be two vectors in $\mathbb{R}^n$. We define the "mean-centered" version of $\vec{x}$ to be:

$$\vec{x}_\text{c} = \begin{bmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{bmatrix}$$

where $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ is the mean of the components of $\vec{x}$. The mean-centered version of $\vec{y}$, named $\vec{y}_\text{c}$, is defined similarly.

Express $\vec{x}_\text{c} \cdot \vec{y}_\text{c}$ using summation notation.

b) (3 pts) Prove that:

$$r = \frac{\vec{x}_\text{c} \cdot \vec{y}_\text{c}}{\|\vec{x}_\text{c}\| \|\vec{y}_\text{c}\|}$$

Do so by starting with the right-hand side of the equation, expanding it, and simplifying it until you reach the definition of r.

c) (2 pts) In 1-2 English sentences, explain why the result from part **b)** implies that $-1 \leq r \leq 1$.

Problem 4: Projections (15 pts)

In Thursday's lecture (and the soon-to-be-released Chapter 2.3), we will introduce the concept of **projecting** one vector onto one or more other vectors. In this problem, you'll see how this concept can be thought of in terms of our friend from the first two weeks of the course: calculus.

Let $\vec{x}$ and $\vec{y}$ be two vectors in $\mathbb{R}^n$. Consider the function $f : \mathbb{R} \to \mathbb{R}$, defined as:

$$f(k) = \|\vec{y} - k\vec{x}\|^2$$

By $\mathbb{R} \to \mathbb{R}$, we mean that f takes in a single real number (i.e. a scalar, **not** a vector) and outputs a single real number. This means that we can find $\frac{\mathrm{d}f}{\mathrm{d}k}$, the derivative of f with respect to k.

Note that $k\vec{x}$ is a vector that points in the same direction (or the opposite direction) as $\vec{x}$.

a) (4 pts) Rewrite $f(k)$ using the properties of the dot product. Then, show that:

$$\frac{\mathrm{d}f}{\mathrm{d}k} = -2\vec{x} \cdot \vec{y} + 2k\vec{x} \cdot \vec{x}$$

b) (2 pts) Find k^*, the value of k that minimizes $f(k)$.

c) (3 pts) Show that the vectors $k^*\vec{x}$ and $\vec{y} - k^*\vec{x}$ are orthogonal.

d) (4 pts) Now, let's study a seemingly unrelated problem.

Suppose we're given a dataset $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ and we'd like to find the optimal model parameter, w, for a simple linear model **with no intercept term**,

$$h(x_i) = wx_i$$

Find the value of w that minimizes the average loss (i.e. empirical risk) when using squared loss. (To be clear, the solution to this problem does not involve linear algebra.)

e) (2 pts) What is the connection between the answer you found in part **d)** and the answer you found in part **b)**?

Problem 5: Norms (12 pts)

In **the last section of Chapter 2.1**, we introduced the concept of vector norms *other than* the "default" Euclidean norm.

To recap, suppose $\vec{v} \in \mathbb{R}^n$.

- The L_2 norm — **which is the default norm, when no subscript is specified** — is defined as

$$\|\vec{v}\| = \|\vec{v}\|_2 = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} = \sqrt{\sum_{i=1}^{n} v_i^2}$$

- The L_1 norm is defined as

$$\|\vec{v}\|_1 = |v_1| + |v_2| + \cdots + |v_n| = \sum_{i=1}^{n} |v_i|$$

- The L_∞ norm is defined as

$$\|\vec{v}\|_\infty = \max_i |v_i|$$

Each of these norms describes a different way of measuring the length of a vector — just like how different loss functions described different ways of measuring the error of a prediction.

a) (2 pts) **In this part only**, let $\vec{v} = \begin{bmatrix} 3 \\ -6 \\ 0 \\ 2 \end{bmatrix}$. Compute $\|\vec{v}\|_2$, $\|\vec{v}\|_1$, and $\|\vec{v}\|_\infty$.

b) (3 pts) In Problem 1, we introduced the parallelogram law, which states that

$$\|\vec{u} + \vec{v}\|^2 + \|\vec{u} - \vec{v}\|^2 = 2\|\vec{u}\|^2 + 2\|\vec{v}\|^2$$

In general, the parallelogram law only holds for the L_2 norm, not necessarily other norms.

Find a counterexample involving two vectors $\vec{u}$ and $\vec{v}$ such that the parallelogram law **does not hold** for the L_1 norm.

c) (3 pts) Prove that

$$\|\vec{v}\|_2 \leq \sqrt{n}\|\vec{v}\|_\infty$$

Hint: Start by writing out the definition of the L_2 norm, and then square it to remove the square root. You will have a sum of n terms. Explain why each of those n terms is less than or equal to $\|\vec{v}\|_\infty^2$. This is most of the way to the proof, but there's still some work you'll need to do after you get to that point.

d) (4 pts) Prove that

$$\|\vec{v}\|_2 \leq \|\vec{v}\|_1$$

Hint: Start with the fact that

$$\|\vec{v}\|_1^2 = (|v_1| + |v_2| + \cdots + |v_n|)^2$$

Problem 6: Feedback (6 pts)

We'd like to get your feedback on how the course has been going so far, now that we're a few weeks in.

You can find the survey **at this link**. It is **not anonymous**, but it links to an anonymous feedback form if you'd like to provide some feedback anonymously.

Thank you for your feedback — it's helping shape our brand-new course.