Homework 11 Solutions

CAS CS 132: Geometric Algorithms

Due: **Thursday December 7, 2023 at 11:59PM**

Submission Instructions

- Make the answer in your solution to each problem abundantly clear (e.g., put a box around your answer or used a colored font if there is a lot of text which is not part of the answer).
- Choose the correct pages corresponding to each problem in Gradescope. Note that Gradescope registers your submission as soon as you submit it, so you don't need to rush to choose corresponding pages. **For multipart questions, please make sure each part is accounted for.**

Graders have license to dock points if either of the above instructions are not properly followed.

Note. Solutions written here may be lengthy because they are expository, and may not reflect that amount of detail that you were expected to write in your own solutions.

Practice Problems

The following list of problems comes from *Linear Algebra and its Application 5th Ed* by David C. Lay, Steven R. Lay, and Judi J. McDonald. They may be useful for solidifying your understanding of the material and for studying in general. **They are optional, so please don't submit anything for them**.

- 6.1.5-8, 6.1.17-18, 6.1.19-20
- 6.2.4-6, 6.2.12-13, 6.2.23-24
- 6.3.3-5, 6.3.11-12, 6.3.21-22
- 6.5.2-4, 6.5.5, 6.5.17-18

1 Basic Analytic Geometry

$$\mathbf{v} = \begin{bmatrix} 1 \\ -1 \\ 5 \\ 7 \\ 4 \end{bmatrix} \qquad \mathbf{u} = \begin{bmatrix} 3 \\ 1 \\ 3 \\ -1 \\ 0 \end{bmatrix}$$

You must show your work for all of the following calculations.

A. (3 points) Compute the norm of $\mathbf{v}$.

B. (4 points) Find the unit vector in the direction of $\mathbf{u}$.

C. (4 points) Compute the distance between $\mathbf{u}$ and $\mathbf{v}$.

D. (4 points) Compute approximately the angle between $\mathbf{u}$ and $\mathbf{v}$ (you will need to use a calculator or Python).

E. (5 points) Using the values you have computed so far, and without computing the angle directly, determine approximately the angle between $\mathbf{v}$ and $\mathbf{v} - \mathbf{u}$. Justify your answer. (*Hint.* Think about triangle created by connecting the tips of $\mathbf{u}$, $\mathbf{v}$ by a line segment.)

Solution.

A. $\|\mathbf{v}\| = \sqrt{\mathbf{v}^T\mathbf{v}} = \sqrt{1+1+25+49+16} = \sqrt{92}$.

B. $\|\mathbf{u}\| = \sqrt{\mathbf{u}^T\mathbf{u}} = \sqrt{9+1+9+1} = \sqrt{20}$ so the unit vector is

$$\frac{\mathbf{u}}{\|\mathbf{u}\|} = \frac{1}{\sqrt{20}} \begin{bmatrix} 3 \\ 1 \\ 3 \\ -1 \\ 0 \end{bmatrix}$$

C.

$$\begin{aligned} \mathsf{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| &= \sqrt{(\mathbf{u} - \mathbf{v})^T(\mathbf{u} - \mathbf{v})} \\ &= \sqrt{2^2 + 2^2 + 2^2 + 8^2 + 4^2} \\ &= \sqrt{4+4+4+64+16} \\ &= \sqrt{92} \end{aligned}$$

D.

$$\cos\theta = \frac{1}{\|\mathbf{u}\|\|\mathbf{v}\|}\mathbf{u}^T\mathbf{v} = \frac{1}{\sqrt{20}\sqrt{92}}(3 + -1 + 15 + -7) = \frac{10}{\sqrt{20}\sqrt{92}} \approx .233$$

Then $\cos^{-1}(.223) \approx 1.335\text{rad} \approx 77.5°$.

E. Since $\mathbf{v}$ and $\mathbf{v}-\mathbf{u}$ have the same length, the triangle described in the hint is isosceles. Therefore, the angle between $\mathbf{v}$ and $\mathbf{v}-\mathbf{u}$ can be computed as $\approx 180° - 2(77.5°) = 27°$

2 The Gram-Schmidt Process

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 2 \\ -2 \end{bmatrix} \qquad \mathbf{u} = \begin{bmatrix} 2 \\ -2 \\ 1 \\ -5 \end{bmatrix}$$

You must show your work for all of the following calculations.[1]

A. (4 points) Find the component of $\mathbf{v}_2$ orthogonal to $\mathbf{v}_1$ (that is, find the vector $\mathbf{z}$ such that $\mathbf{v}_2 = \hat{\mathbf{v}}_2 + \mathbf{z}$ where $\hat{\mathbf{v}}_2$ is the orthogonal projection of $\mathbf{v}_2$ onto $\mathbf{v}_1$). We will refer to this as $\mathbf{v}_2'$ below.

B. (4 points) Find the component of $\mathbf{v}_3$ orthogonal to $\mathbf{v}_1$. We will refer to this as $\mathbf{v}_3'$ below.

C. (4 points) Find the component of $\mathbf{v}_3'$ orthogonal to $\mathbf{v}_2'$. We will refer to this as $\mathbf{v}_3''$ below.

D. (3 points) Compute the inner product $\mathbf{v}_3''$ and $\mathbf{v}_1$.

E. (5 points) Find $[\mathbf{u}]_{\mathcal{B}}$ where $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3''\}$

Solution.

A.

$$\hat{\mathbf{v}}_2 = \frac{\mathbf{v}_2^T \mathbf{v}_1}{\mathbf{v}_1^T \mathbf{v}_1} \mathbf{v}_1 = \mathbf{v}_1$$

$$\mathbf{v}_2' = \mathbf{v}_2 - \hat{\mathbf{v}}_2 = \mathbf{v}_2 - \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

B.

$$\hat{\mathbf{v}}_3 = \frac{\mathbf{v}_3^T \mathbf{v}_1}{\mathbf{v}_1^T \mathbf{v}_1} \mathbf{v}_1 = \frac{\mathbf{v}_1}{3}$$

$$\mathbf{v}_3' = \mathbf{v}_3 - \hat{\mathbf{v}}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_1}{3} = \begin{bmatrix} 2/3 \\ 0 \\ 5/3 \\ -7/3 \end{bmatrix}$$

[1] The Gram-Schmidt process is an algorithm for converting a basis into an orthogonal basis. We won't have time to cover Gram-Schmidt in full detail in this course, but this problem details how the procedure works for three vectors.

C.

$$\hat{\mathbf{v}_3'} = \frac{\mathbf{v}_3'^T \mathbf{v}_2'}{(\mathbf{v}_2')^T \mathbf{v}_2'} \mathbf{v}_2' = -\mathbf{v}_2'$$

$$\mathbf{v}_3'' = \mathbf{v}_3' + \mathbf{v}_2' = \begin{bmatrix} -1/3 \\ 1 \\ 5/3 \\ -4/3 \end{bmatrix}$$

D. $\langle \mathbf{v}_3'', \mathbf{v}_1 \rangle = 1(1) + 0(0) + 1(2) + 1(-2) = 0$

E. Since $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3''\}$ is an orthogonal set, we can use the equation for finding a solution to the equation

$$\mathbf{u} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2' + x_3 \mathbf{v}_3''$$

So we have

$$[\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} \frac{\mathbf{u}^T \mathbf{v}_1}{\mathbf{v}_1^T \mathbf{v}_1} \\ \frac{\mathbf{u}^T \mathbf{v}_2'}{(\mathbf{v}_2')^T \mathbf{v}_2'} \\ \frac{\mathbf{u}^T \mathbf{v}_3''}{(\mathbf{v}_3'')^T \mathbf{v}_3''} \end{bmatrix} = \begin{bmatrix} -2/3 \\ -3 \\ 1 \end{bmatrix}$$

3 Orthogonal Projection Matrices

$$\mathbf{v} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} \qquad A = \begin{bmatrix} 1 & 2 & -5 \\ 1 & 1 & 3 \\ 8 & 14 & -6 \\ 1 & 2 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & -4 & -2 & 9 & 0 \\ 0 & 2 & 2 & -6 & 1 \\ 3 & -5 & 1 & 9 & 5 \\ 0 & 1 & 1 & -2 & 1 \end{bmatrix}$$

You must show your work for all of the following calculations.

A. (2 points) Find the matrix which implements orthogonal projection onto span$\{\mathbf{v}\}$.

B. (3 points) For an arbitrary vector $\mathbf{x}$ in $\mathbb{R}^n$, find an expression for the matrix which implements orthogonal projection onto span$\{\mathbf{x}\}$.

C. (3 points) Find approximately the matrix which implements orthogonal projection onto $\operatorname{Col} A$. You should use Python for this. Write down the NumPy expressions you used for your calculation.

D. (6 points) Find approximately the matrix which implements orthogonal projection onto $\operatorname{Col} B$. You should use Python for this. (*Hint.* First determine the rank of B.) Write down the NumPy expressions you used for your calculation.

E. (3 points) What is the relationship between $\operatorname{Col} A$ and $\operatorname{Col} B$? Justify your answer.

F. (3 points) Suppose that C is an arbitrary $m \times n$ *orthonormal* matrix. Find an expression for the matrix which implements orthogonal projection onto $\operatorname{Col} C$. Your expression should be as simplified as possible.

Solution.

A. If we want to determine the matrix which implements orthogonal projection onto span$\{\mathbf{v}\}$, we need to determine how the transformation affects the standard basis:

$$\frac{\mathbf{v}^T\mathbf{e}_1}{\mathbf{v}^T\mathbf{v}}\mathbf{v} = \frac{2\mathbf{v}}{14}$$
$$\frac{\mathbf{v}^T\mathbf{e}_2}{\mathbf{v}^T\mathbf{v}}\mathbf{v} = \frac{3\mathbf{v}}{14}$$
$$\frac{\mathbf{v}^T\mathbf{e}_3}{\mathbf{v}^T\mathbf{v}}\mathbf{v} = \frac{\mathbf{v}}{14}$$

Therefore, the matrix which implements this transformation is

$$\frac{1}{14}[2\mathbf{v} \;\; 3\mathbf{v} \;\; \mathbf{v}] = \frac{1}{14}\begin{bmatrix} 4 & 6 & 2 \\ 6 & 9 & 3 \\ 2 & 3 & 1 \end{bmatrix}$$

B. We can either recognize the pattern from the previous problem, or we can use the equation

$$\hat{\mathbf{b}} = A(A^TA)^{-1}A^T\mathbf{b}$$

In this case, $A = \mathbf{v}$ so the 'A^TA' part is $\mathbf{v}^T\mathbf{v}$, which is just a number, so we can move it to the front of the equation to get

$$\hat{\mathbf{b}} = \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}\mathbf{b}$$

So the final expression is $\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}$.

C. Since A has full rank, its columns form a basis for $\operatorname{Col} A$. Therefore, we can use the above equation directly. In python, this would be

```
a @ np.linalg.inv(a.T @ a) @ a.T
```

which gives the matrix

```
[[ 0.60869565 -0.26086957  0.13043478 -0.39130435]
 [-0.26086957  0.82608696  0.08695652 -0.26086957]
 [ 0.13043478  0.08695652  0.95652174  0.13043478]
 [-0.39130435 -0.26086957  0.13043478  0.60869565]]
```

D. We cannot use the equation above directly because B the rank of B is 3, and so it does not have full rank. So we need to determine a set of three columns which are linearly independent. We can do this in a number of ways (I'll leave this to you) but we can find that the first, second and fourth columns are linearly independent. So with `b0 = b[:,[0, 1, 3]]` we can use

```
b0 @ np.linalg.inv(b0.T @ b0) @ b0.T
```

which gives the matrix

```
[[ 0.60869565 -0.26086957  0.13043478 -0.39130435]
 [-0.26086957  0.82608696  0.08695652 -0.26086957]
 [ 0.13043478  0.08695652  0.95652174  0.13043478]
 [-0.39130435 -0.26086957  0.13043478  0.60869565]]
```

E. We see that the orthogonal projections onto $\operatorname{Col} A$ and $\operatorname{Col} B$ are identical. This means that in fact the column spaces are identical as well.

F. In the case that C is orthonormal, we can use the fact that $C^TC = I$ to simplify the above equation:

$$C(C^TC)^{-1}C^T = CI^{-1}C^T = CC^T$$

4 Least Squares

$$A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ 1 & -1 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix} \qquad U = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \qquad \mathbf{c} = \begin{bmatrix} 6 \\ 5 \\ 4 \\ 7 \\ 2 \\ 3 \end{bmatrix}$$

A. (3 points) Find the normal equations for $A\mathbf{x} = \mathbf{b}$.

B. (2 points) Find a least squares solution for $A\mathbf{x} = \mathbf{b}$. Then use `np.linalg.lstsq` to verify your solution. You should read the NumPy documentation to verify how this function works. Include in your solution what Python code your wrote to verify your solution.

C. (1 points) Is the least squares solution you found in the previous part unique? Justify your answer.

D. (4 points) Find the normal equations for $U\mathbf{x} = \mathbf{c}$.

E. (4 points) Find two linearly independent least squares solutions for $U\mathbf{x} = \mathbf{c}$.

F. (6 points) Let C be an arbitrary $m \times n$ matrix of rank k and let $\mathbf{b}$ be an arbitrary vector in $\mathbb{R}^m$. Find an expression for the maximum size of a linearly independent set of least squares solutions for $C\mathbf{x} = \mathbf{b}$ where $\mathbf{b} \neq \mathbf{0}$. You do not need to prove that your expression is correct, but you should justify your answer. (*Hint.* Think about how to generate multiple least squares solutions in the case that there is more than one. Your expression should be in terms of m, n, and k, thought it may not use every one of these values.)

Solution.

A.

$$A^T A = \begin{bmatrix} 2 & 1 \\ 1 & 6 \end{bmatrix}$$
$$A^T \mathbf{b} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

so the normal equations are given by

$$\begin{bmatrix} 2 & 1 \\ 1 & 6 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

B. Since the columns of A^TA are linearly independent, this matrix is invertible and

$$(A^TA)^{-1} = \frac{1}{11}\begin{bmatrix} 6 & -1 \\ -1 & 2 \end{bmatrix}$$

Then a least squares solution is given by

$$(A^TA)^{-1}A^T\mathbf{b} = \frac{1}{11}\begin{bmatrix} 6 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 18/11 \\ 8/11 \end{bmatrix}$$

C. This solution is unique because the columns of A are linearly independent (and because A^TA is invertible.

D.

$$U^TU = \begin{bmatrix} 6 & 3 & 3 \\ 3 & 3 & 0 \\ 3 & 0 & 3 \end{bmatrix}$$

$$U^T\mathbf{c} = \begin{bmatrix} 27 \\ 12 \\ 15 \end{bmatrix}$$

so the normal equations are given by

$$\begin{bmatrix} 6 & 3 & 3 \\ 3 & 3 & 0 \\ 3 & 0 & 3 \end{bmatrix}\mathbf{x} = \begin{bmatrix} 27 \\ 12 \\ 15 \end{bmatrix}$$

E. The reduced echelon form of the above system is

$$\begin{bmatrix} 1 & 0 & 1 & 5 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We can write the general form solution of this system as

$$x_1 = 5 - x_3$$
$$x_2 = -1 + x_3$$
$$x_3 \text{ is free}$$

We can also write this as a linear combination of vectors

$$\begin{bmatrix} 5 \\ -1 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

we can then take the solutions

$$\begin{bmatrix} 5 \\ -1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -4 \\ 0 \\ 1 \end{bmatrix}$$

F. We might remember from a previous homework problem that a solution set is an *affine* subspace which can be viewed as the null space of A translated by some vector. Alternatively, when we generate linearly independent vectors from a general solution, it amounts to choosing a free variable to make nonzero, and setting all the others to be 0. This means we can choose at least $\dim(\operatorname{Nul} A)$ many linearly independent solutions. But there is also the solution which sets all the free variables to 0. So all together we can choose $\dim(\operatorname{Nul} A) + 1$ many linearly independent solutions. This gives the expression $n - k + 1$ by the rank-nullity theorem.