# DSC 140B - Midterm 01 Review Questions

Below is a sample of practice problems for Midterm 01. Note that the format and content of the actual exams will be similar to below, but that the number of questions will be different. While some of the questions below may require a calculator, questions on the exam will not (the numbers will be "nicer").

**Problem 1.**

True or false; the transformation $f(\vec{x}) = (2x_2, 4x_1)^T$, expressed in the standard basis, is a linear transformation.

- ◯ True
- ◯ False

**Problem 2.**

Let $\hat{u}^{(1)} = (.31, .95)^T$ and $\hat{u}^{(2)} = (-.95, .31)^T$, and let $\mathcal{U} = \{\vec{u}^{(1)}, \vec{u}^{(2)}\}$. Note that $\mathcal{U}$ is an orthonormal basis (or, at least is close enough to assume that it is).

Let $\vec{x} = (1, -3)$ in the standard basis. Let $(x_1, x_2)^T$ be the coordinates of $\vec{x}$ in the basis $\mathcal{U}$. What are these coordinates?

$x_1$: ☐

$x_2$: ☐

**Problem 3.**

Let $\hat{e}^{(1)}$ and $\hat{e}^{(2)}$ be the standard basis vectors in $\mathbb{R}^2$. Suppose that $\vec{f}$ is a linear transformation, and that $\vec{f}(\hat{e}^{(1)}) = (1, 2)^T$ and $\vec{f}(\hat{e}^{(2)}) = (2, -1)^T$.

Let $\vec{x} = (-5, 3)^T$. What are the coordinates of $\vec{f}(\vec{x})$?

Coordinate 1: ☐

Coordinate 2: ☐

**Problem 4.**

Suppose that in the standard basis, $\vec{f}(\vec{x}) = (3x_2, -3x_1)^T$, where $\vec{x} = (x_1, x_2)^T$. Let $\hat{u}^{(1)} = \frac{1}{\sqrt{2}}(1, 1)^T$ and $\hat{u}^{(2)} = \frac{1}{\sqrt{2}}(-1, 1)^T$, and let $\mathcal{U} = \{\vec{u}^{(1)}, \vec{u}^{(2)}\}$.

The formula for $\vec{f}(\vec{x})$ in the basis $\mathcal{U}$ is of the form $[\vec{f}(\vec{z})]_{\mathcal{U}} = (\alpha z_2, \beta z_1)^T$, where $[\vec{z}]_{\mathcal{U}} = (z_1, z_2)^T$. What are $\alpha$ and $\beta$?

$\alpha$: ☐

$\beta$: 

*Hint*: it may help to draw a picture.

**Problem 5.**

Let $f$ be a linear transformation in $\mathbb{R}^3$, and suppose:

- $f(\hat{e}^{(1)}) = (-2, -4, 5)$,
- $f(\hat{e}^{(2)}) = (3, 0, 4)$,
- $f(\hat{e}^{(3)}) = (1, 2, 1)$.

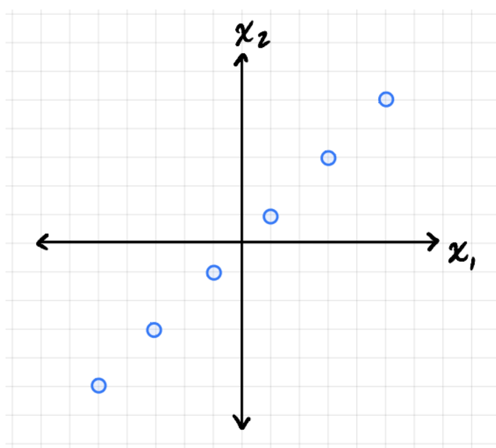
Write the matrix $A$ representing $f$ in the standard basis.

**Problem 6.**

Consider the linear transformation $\vec{f} : \mathbb{R}^2 \to \mathbb{R}^2$ which flips a vector over the $y$ axis. Write an eigenvector of this transformation.

**Problem 7.**

Consider a dataset consisting of collinear points in $\mathbb{R}^2$, as is shown below.




Consider the covariance matrix of this data. What will be its *smallest* eigenvalue?

**Problem 8.**

Suppose a $d$-dimensional dataset is reduced to $\mathbb{R}^3$ with PCA. Suppose the covariance matrix of the new dataset is computed. Which of the below could possibly be this covariance matrix?

○ $\begin{pmatrix} 5 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 2 \end{pmatrix}$

○ $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 0 \\ 3 & 0 & 3 \end{pmatrix}$

○ $\begin{pmatrix} 7 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

○ $\begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 2 & 3 & 0 \end{pmatrix}$

**Problem 9.**

Suppose we perform PCA on a dataset in order to reduce the dimensionality to 3. Let $Z$ be the data represented in the new feature representation, after performing PCA. Suppose the covariance matrix of $Z$ is computed. Which of these could possibly be the covariance matrix? Select all that apply.

○ $\begin{pmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

○ $\begin{pmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

○ $\begin{pmatrix} 6 & -1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

○ $\begin{pmatrix} 6 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

**Problem 10.**

True or False: it is possible for an off-diagonal entry in a covariance matrix to be negative.

○ True

○ False

**Problem 11.**

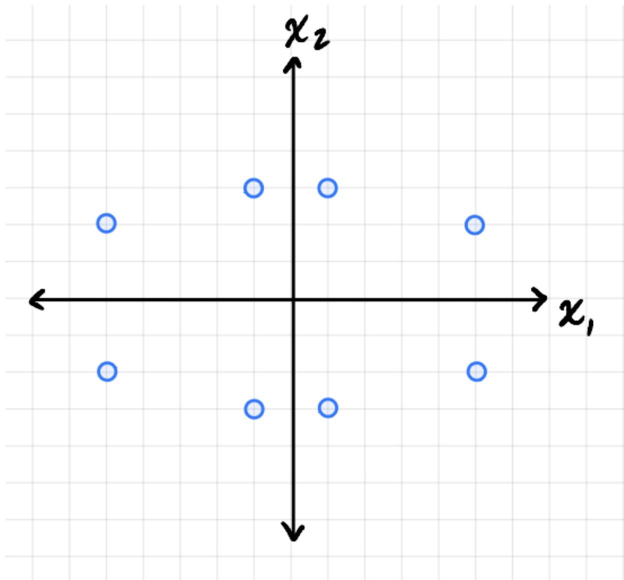
True or False: it is possible for an eigenvalue of a covariance matrix to be negative.

○ True

○ False

**Problem 12.**

Consider the data set shown below.




Suppose PCA is used to reduced the dimension to 1. What will be the reconstruction error?

**Problem 13.**

Is performing PCA once to go from 100 dimensions to 50 the same as performing PCA twice, once to go from 100 dimensions to 75, and again to go from 75 to 50? More formally, suppose $X$ is a data set in $R^{100}$. Consider the following two approaches to dimensionality reduction with PCA.

Approach 1) We create a new dataset $Z_1$ in $\mathbb{R}^{50}$ by applying PCA to $X$. Approach 2) We first create a new dataset $X'$ in $\mathbb{R}^{75}$ by applying PCA to $X$. Next, we create another new dataset, $Z_2$, by applying PCA to $X'$.

True or False: The result of both approaches must be identical. That is, $Z_1$ and $Z_2$ are the same.

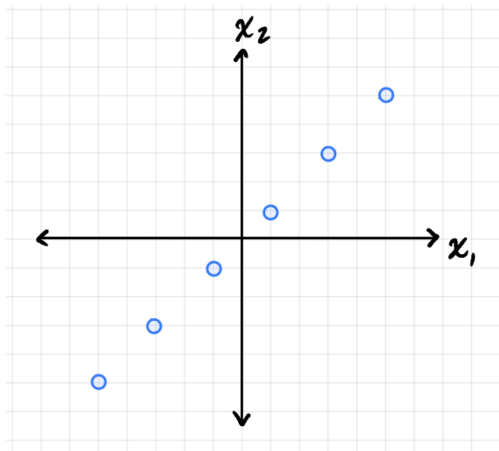
Note: you may assume that all eigenvalues of the covariance matrix of $X$ are greater than zero.

◯ True

◯ False

**Problem 14.**

Consider a dataset consisting of collinear points in $\mathbb{R}^2$, as is shown below.




Consider the covariance matrix of this data. What will be its *largest* eigenvalue?

**Problem 15.**

Let $X$ be an already-centered dataset. Recall that *standardizing* a dataset means centering each feature and dividing each feature by its standard deviation.

Now consider the following two approaches to dimensionality reduction with PCA:

- Approach 1: standardize $X$ to get $X'$, then perform PCA to get a dataset $Z_1$ - Approach 1: perform PCA on $X$ to get a dataset $Z'$, then standardize $Z'$ to get a dataset $Z_2$

True or False: these two approaches are the same. That is, $Z_1 = Z_2$.

- ◯ True
- ◯ False

**Problem 16.**

Let $X$ be a matrix of data points in $\mathbb{R}^3$ (each row is a data point, and each column is a different feature). Suppose the second column is exactly twice the first column, and the third column is exactly three times the first. Let $C$ be the covariance matrix of $X$ How many non-zero eigenvalues will $C$ have?

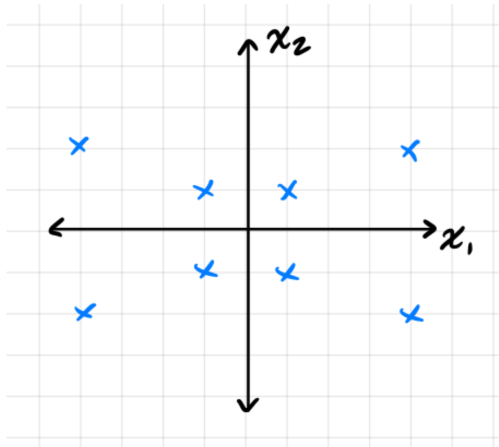
**Problem 17.**

Suppose $f$ is the linear transformation $\vec{f}(x_1, x_2) = (2x_1, 5x_2)^T$.

What is $\max_{\vec{u}} \vec{u} \cdot \vec{f}(\vec{u})$ subject to the constraint that $u$ is a unit vector?

**Problem 18.**

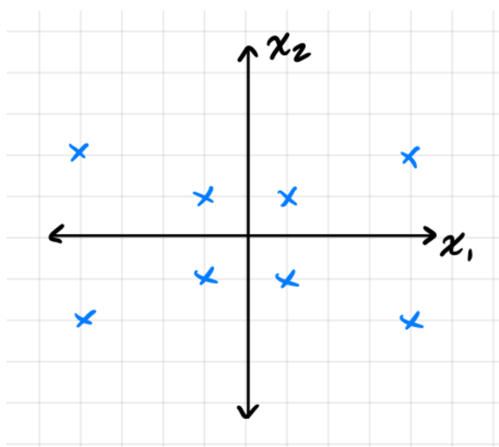
Consider the data below. What is the top eigenvalue of the data's covariance matrix?




Note that the grid in the image can be used to determine the coordinates of each point. Your answer should be accurate within 0.01

**Problem 19.**

Consider again the data from above:




Suppose PCA is used to project the data onto one principal component. What is the reconstruction error of this projection?