**Homework 12**

**Instructions**

This homework contains **5** concepts and **3** programming questions. In MS word or a similar text editor, write down the problem number and your answer for each problem. Combine all answers for concept questions in a single PDF file. Export/print the Jupyter notebook as a PDF file including the code you implemented and the outputs of the program. Make sure all plots and outputs are visible in the PDF.

Combine all answers into a single PDF named andrewID\_hw12.pdf and submit it to Gradescope before the due date. Refer to the syllabus for late homework policy. Please assign each question a page by using the “Assign Questions and Pages” feature in Gradescope.

Here is a breakdown of the points for programming questions:

|  |  |
| --- | --- |
| Name | Points |
| Concept 1 | 6 |
| Concept 2 | 6 |
| Concept 3 | 6 |
| Concept 4 | 6 |
| Concept 5 | 6 |
| M12-L1-P1 | 15 |
| M12-L2-P1 | 25 |
| M12-HW1 | 50 |
| **Total** | **120** |
| Bonus | 6 |

Problem 1 (2 points)

What would the dimension of the covariance matrix be for the following data:

(Choose one)

1. 2 x 2
2. 6 x 6
3. 12 x 12
4. 20 x 20

A black background with a black square

Description automatically generated with medium confidence

Problem 2 (2 points)

Provided the following eigenvalues and eigenvectors e1 and e2, what are the values *i*, *j*, *k*, that comprise the unit normalized third eigenvector, e3?

(Text entry for each i, j, k)

A black background with a black square

Description automatically generated with medium confidence

Problem 3 (2 points)

The eigenvalues of the covariance matrix from the data in the first concept question are included below. Which components should be used to explain at least 80% of the variance in the data?

A black background with a black square

Description automatically generated with medium confidence

Multiple choice (select all that apply)

* PC1
* PC2
* PC3
* PC4
* PC5
* PC6

Problem 4 (2 points)

What should the dimension of the covariance matrix be for the following data:

(Choose one)

1. 2 x 2
2. 6 x 6
3. 10x10
4. 20 x 20

A black background with a black square

Description automatically generated with medium confidence

Problem 5 (2 points)

Select the following statements about t-SNE which are true:

(Multiple choice, select all that apply)

1. t-SNE can be used to project unseen high dimensional data into a reduced feature space
2. t-SNE preserves global structure and distances between data points by computing pairwise similarities
3. Like PCA, t-SNE is a linear dimensionality reduction technique that is used to reduce high dimensional data to a low dimensional feature space
4. t-SNE is a non-linear dimensionality technique that can learn embeddings of manifolds