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

x_1	x_2	x_3	x_4	x_5	x_6
-7.55	5.85	11.88	1.99	6.39	3.05
-10.93	6.56	8.96	-0.89	7.43	4.07
-9.44	6.37	9.86	-0.62	7.73	2.88
-1.83	0.53	8.55	-6.21	-8.05	5.13
6.38	0.47	-6.72	2.71	-5.24	-2.11
7.85	-0.17	-8.48	1.40	-7.62	-3.71
9.17	0.70	-7.45	2.09	-6.13	-4.66
0.76	1.97	8.46	-5.47	-7.57	3.33
-11.58	6.13	9.34	0.21	9.00	3.03
-8.41	5.29	10.13	-0.97	7.48	5.11
-7.87	5.48	10.50	1.71	6.04	3.79
-0.84	0.23	7.99	-6.91	-7.59	3.11
1.06	-0.56	7.47	-7.12	-6.31	3.82
7.43	1.26	-8.13	1.30	-5.78	-6.79
0.59	0.88	7.85	-6.20	-8.18	3.94
7.35	1.04	-5.98	1.61	-5.69	-5.54
1.01	1.40	9.87	-5.62	-7.74	4.08
8.47	2.80	-7.24	0.93	-5.39	-4.60
8.00	1.39	-6.57	0.53	-2.77	-7.12
-10.92	7.00	8.96	-1.30	6.90	4.82

Problem 2 (2 points)

Provided the following eigenvalues and eigenvectors e_1 and e_2, what are the values *i*, *j*, *k*, that comprise the unit normalized third eigenvector, e_3?
(Text entry for each i, j, k)

$$\lambda_1 = 16$$

$$\lambda_2 = 4$$

$$\lambda_3 = 0$$

$$e_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

$$e_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

$$e_3 = \begin{bmatrix} i & j & k \end{bmatrix}$$

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?

$\lambda_1 = 160.30$

$\lambda_2 = 44.31$

$\lambda_3 = 1.86$

$\lambda_4 = 1.47$

$\lambda_5 = 0.62$

$\lambda_6 = 0.49$

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

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
-9.25	2.84	-9.38	0.66	5.71	-2.23	8.76	-5.37	-2.56	1.25
-10.24	3.23	-8.34	-0.70	5.53	-2.72	8.70	-4.77	-2.61	0.44
2.36	-10.36	5.22	-2.26	7.44	-4.88	-4.87	1.83	-8.76	-7.48
-7.84	5.72	-2.35	8.14	-6.54	10.40	-2.19	-2.51	-3.84	-1.19
-7.51	5.07	-2.21	6.73	-7.42	8.83	-4.00	-2.65	-3.57	-0.89
0.49	-8.68	4.84	0.05	6.40	-4.71	-4.96	2.05	-7.59	-6.18

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