SC4000 Machine Learning Tutorial
Dimensionality Reduction

Question 1: A dataset of five 4-dimensional instances is given in Table 1. Suppose an SVD is performed on the data matrix $\mathbf{X}$ (5-by-4) via $\mathbf{X} = \mathbf{V}\mathbf{D}\mathbf{U}^{\top}$. The matrices $\mathbf{V}$, $\mathbf{D}$, and $\mathbf{U}$ are shown in Tables 2-4, respectively. Use principal component analysis to project the 5 data points in Table 1 to 2-dimensional space.

Table 1: Data set for Question 1.

Data Points	X_1	X_2	X_3	X_4
P1	2	4	1	3
P2	1	2	3	5
P3	-2	-4	-4	-1
P4	0	-1	-2	-6
P5	-1	-1	2	-1

Table 2: The matrix $\mathbf{V}$ (5-by-5) obtained by SVD ($\mathbf{X} = \mathbf{V}\mathbf{D}\mathbf{U}^\top$)

-0.4577	0.2550	-0.5536	0.4680	0.4472
-0.5612	-0.2150	0.1896	-0.6348	0.4472
0.4497	-0.7183	-0.2749	0.0787	0.4472
0.5206	0.6063	-0.1153	-0.3849	0.4472
0.0486	0.0720	0.7542	0.4730	0.4472

Table 3: The matrix $\mathbf{D}$ (5-by-4) obtained by SVD ($\mathbf{X} = \mathbf{V}\mathbf{D}\mathbf{U}^\top$)

10.9040	0	0	0
0	4.8385	0	0
0	0	3.3973	0
0	0	0	0.3867
0	0	0	0

Table 4: The matrix $\mathbf{U}$ (4-by-4) obtained by SVD ($\mathbf{X} = \mathbf{V}\mathbf{D}\mathbf{U}^\top$)

-0.2224	0.3430	-0.3302	-0.8508
-0.4880	0.5756	-0.4046	0.5166
-0.4479	0.2924	0.8400	-0.0911
-0.7154	-0.6823	-0.1473	-0.0309