Problem 1 (10 pts.)

When answering a numerical value, enter only the numerical value in the answer box. No unit required. If it says "up to N decimal places", please answer by rounding off the N+1 decimal places.

1.

Select the words in the following text about principal component analysis and dimensionality reduction correctly.

Principal component analysis is a kind of dimension reduction method based on [(unselected) ⌄] .
Dimensionality reduction is important not only for visualization, but also for avoiding [(unselected) ⌄] that there are few other data points near one in high dimension.

2.

dm-end1-1.csv records the results of 40 students in Japanese, English, math, physics and chemistry. Answer the following questions about this. If necessary, use dm-end1-1.ipynb.

1. Perform a principal component analysis (PCA) on this data. Standardize the data so that the average is 0 and the variance is 1 before PCA. Also, note that the maximum number of principal components is 5.

Answer the contribution rates of the first and second principal components (values from 0 to 1, up to the third place). Also, from which number of the principal components the cumulative contribution rate exceeds 0.8 for the first time.

The contribution rate of the first principal component (values from 0 to 1, up to the third decimal place)

[   ]

The contribution rate of the second principal component (values from 0 to 1, up to the third decimal place)

[   ]

The number of the principal components from which the cumulative contribution rate exceeds 0.8 for the first time (Integer only)

[   ]

2. Draw a biplot using the results of the principal component analysis. From the results, for the second principal component axis, choose the most appropriate one from the following explanations.

◉ (unselected)
○ The second principal component axis is positively correlated with the performance of language subjects such as Japanese and English.
○ The second principal component axis is negatively correlated with the performance of language subjects such as Japanese and English.
○ The second principal component axis is positively correlated with the results of science subjects such as mathematics, physics, and chemistry.
○ The second principal component axis is negatively correlated with the results of science subjects such as mathematics, physics, and chemistry.

upload ipynb file:

[参照...] ファイルが選択されていません。

upload html file:

[参照...] ファイルが選択されていません。