Aprendizagem 2023

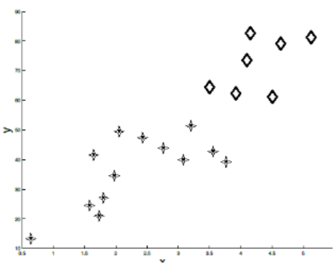
Lab 10: Dimensionality Reduction

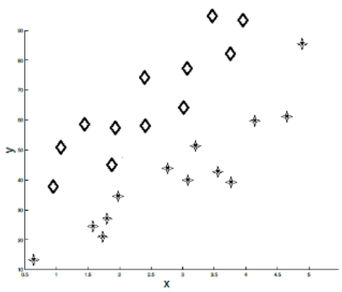
Prof: Rui Henriques

Practical exercises

1. Given the following datasets where observations are in $\mathbb{R}^2$ and belong to one of two classes:

	y_1	y_2
$\mathbf{x}_1$	0	0
$\mathbf{x}_2$	4	0
$\mathbf{x}_3$	2	1
$\mathbf{x}_4$	6	3





Which principal components can accurately discriminate the class per dataset?

Left chart: the eigenvector with highest eigenvalue is able to discriminate the given observations.

Right chart: the eigenvector with lowest eigenvalues is able to discriminate the given observation.

2. The following top-7 eigenvalues explain 90% of the variation of dataset X:

$$\lambda_1=20,\ \lambda_2=10,\ \lambda_3=5,\ \lambda_4=4,\ \lambda_5=3,\ \lambda_6=2,\ \lambda_7=1$$

What is the most accurate information regarding X:

i. X has less than 7 attributes
ii. X has 7 attributes
iii. X has more than 7 attributes
iv. X has more than 11 attributes

We know that the top seven components only explain 90% of the variation, hence there are more than 7 components, i.e. the dataset has more than 7 attributes.

The subsequent components have at most an eigenvalue of 1.

A eigenvalue of 1 is able to explain 2% of variability since $\sum_{i=1}^{7} \lambda_i = 45$ explain 90% of the variation.

In this context, there are at least 5 additional components, hence the data has 12 or more variables.

The most accurate answer is (iv).

3. Given a set of data points in $\mathbb{R}^3$, the following covariance matrix was obtained:

$$\begin{bmatrix} 91.43 & 171.92 & 297.99 \\ & 373.92 & 545.21 \\ & & 1297.26 \end{bmatrix}$$

as well as the following eigenvectors retrieved:

$$\boldsymbol{u}_1 = \begin{pmatrix} 0.2179 \\ 0.4145 \\ 0.8836 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} -0.2466 \\ -0.8525 \\ 0.4608 \end{pmatrix}, \boldsymbol{u}_3 = \begin{pmatrix} 0.9443 \\ -0.3183 \\ -0.0836 \end{pmatrix}$$

Please select the more complete answer:

i. eigenvalue λ1 is approximately 1626
ii. eigenvalue λ2 is approximately 129
iii. eigenvalues λ1 and λ2 explain >99% of the variation in data
iv. all of the above

Recovering the central property of eigenvectors $\boldsymbol{C}\mathbf{u} = \lambda\mathbf{u}$ we can retrieve the three eigenvalues:

$$\boldsymbol{C}\mathbf{u_1} = \lambda_1\mathbf{u_1} \Leftrightarrow \begin{pmatrix} 91.43 & 171.92 & 297.99 \\ 171.92 & 373.92 & 545.21 \\ 297.99 & 545.21 & 1297.26 \end{pmatrix}\begin{pmatrix} 0.2179 \\ 0.4145 \\ 0.8836 \end{pmatrix} = \begin{pmatrix} 354.49 \\ 674.20 \\ 1437.18 \end{pmatrix} = \begin{pmatrix} 0.2179 \times \lambda_1 \\ 0.4145 \times \lambda_1 \\ 0.8836 \times \lambda_1 \end{pmatrix} \Leftrightarrow \lambda_1 \approx 1626.5$$

$$\boldsymbol{C}\mathbf{u_2} = \lambda_2\mathbf{u_2} \Leftrightarrow \begin{pmatrix} 91.43 & 171.92 & 297.99 \\ 171.92 & 373.92 & 545.21 \\ 297.99 & 545.21 & 1297.26 \end{pmatrix}\begin{pmatrix} -0.2466 \\ -0.8525 \\ 0.4608 \end{pmatrix} = \begin{pmatrix} -31.79 \\ -109.93 \\ 59.50 \end{pmatrix} = \begin{pmatrix} -0.2466 \times \lambda_2 \\ -0.8525 \times \lambda_2 \\ 0.4608 \times \lambda_2 \end{pmatrix} \Leftrightarrow \lambda_2 \approx 129.0$$

$$\boldsymbol{C}\mathbf{u_3} = \lambda_3\mathbf{u_3} \Leftrightarrow \begin{pmatrix} 91.43 & 171.92 & 297.99 \\ 171.92 & 373.92 & 545.21 \\ 297.99 & 545.21 & 1297.26 \end{pmatrix}\begin{pmatrix} 0.9443 \\ -0.3183 \\ -0.0836 \end{pmatrix} = \begin{pmatrix} 6.70 \\ -2.25 \\ -0.60 \end{pmatrix} = \begin{pmatrix} 0.9443 \times \lambda_3 \\ -0.3183 \times \lambda_3 \\ -0.0836 \times \lambda_3 \end{pmatrix} \Leftrightarrow \lambda_3 = 7.1$$

Correct answer: iv).

4. Given the following dataset:

	y_1	y_2
$\mathbf{x}_1$	1	-1
$\mathbf{x}_2$	0	1
$\mathbf{x}_3$	-1	0

and the corresponding eigenvectors and eigenvalues:

λ_1=3/2 and λ_2=1/2

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \boldsymbol{u}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

a) Transform the input data using PCA

Mapping points to this new eigenspace through $\mathbf{x}' = U^T\mathbf{x}$, we get:

	c_1	c_2
$\mathbf{x'}_1$	$\frac{1}{\sqrt{2}}1 - \frac{1}{\sqrt{2}} \times -1 = \frac{2}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}1 + \frac{1}{\sqrt{2}} \times -1 = 0$
$\mathbf{x'}_2$	$\frac{1}{\sqrt{2}}0 - \frac{1}{\sqrt{2}} \times 1 = -\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}0 + \frac{1}{\sqrt{2}} \times 1 = \frac{1}{\sqrt{2}}$
$\mathbf{x'}_3$	$\frac{1}{\sqrt{2}} - 1 - \frac{1}{\sqrt{2}} \times 0 = -\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}} - 1 + \frac{1}{\sqrt{2}} \times 0 = -\frac{1}{\sqrt{2}}$

b) [*optional*] Assess the recovery error when considering the most informative component only

Let us first recovery the data. A property of U is that its inverse corresponds to the transpose matrix (orthogonal property). In this context, recovery is a simple step:

$$\mathbf{x} = (U^T)^{-1}\mathbf{x}' = (U^T)^T\mathbf{x}' = U\mathbf{x}'$$

When considering a subset of components:

$$\mathbf{x}^{recovered} = (\mathbf{u}_1 \quad \cdots \quad \mathbf{u}_p)\begin{pmatrix} x'_1 \\ \cdots \\ x'_p \end{pmatrix}$$

$$\mathbf{x_1}^{recovered} = \mathbf{u}_1 x'_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}\left(\frac{2}{\sqrt{2}}\right) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\mathbf{x_2}^{recovered} = \mathbf{u}_2 x'_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}\left(-\frac{1}{\sqrt{2}}\right) = \begin{pmatrix} -0.5 \\ 0.5 \end{pmatrix}$$

$$\mathbf{x_3}^{recovered} = \mathbf{u}_3 x'_3 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}\left(-\frac{1}{\sqrt{2}}\right) = \begin{pmatrix} -0.5 \\ 0.5 \end{pmatrix}$$

Reconstruction squared error:

$$\sum_{i=1}^{3} \|\mathbf{x}_i - \mathbf{x}'_i\|_2^{\ 2} = \|(0,0)\|_2^{\ 2} + \|(0.5,0.5)\|_2^{\ 2} + \|(-0.5,-0.5)\|_2^{\ 2} = 0.5 + 0.5 = 1$$