Unsupervised Learning - Practical session
IMA205

Pietro Gori

25 March 2020

This Practical session is about unsupervised learning. We will use the dimensionality reduction and clustering techniques presented this morning to analyze toy examples and recognize faces.

You have two weeks (18/04) to update a small report (2 jupyter notebooks + theoretical questions) to the *site pédagogique* of IMA205 under the section *Reports-TP*. You can answer in French or English. The deadline is 23:59 of the 18th of April. I remind you that the report is mandatory and evaluated.

1 PCA

1. Why do we need to center the data before computing a PCA ? If you want, you can use the toy examples to answer this question.

2. Let x_p and x_q be two row-vectors representing two images, U an *orthogonal* matrix whose columns are the eigenvectors of X and $y_p = x_p U$, $y_q = x_q U$, check that $x_p x_q^T = y_p y_q^T$. This shows that $Y = XU$ is a linear transformation that preserves inner products.

3. Let X be the original data, a matrix $[N, d]$, and Y the scores of a PCA keeping all eigenvectors, which means that Y is also a matrix $[N, d]$. Are X and Y equal ? If not, why ? What would you use (generally speaking) in a machine learning problem ? Why ?

4. Let C be the covariance matrix of X and $C = UDU^T$ its eigen decomposition. Show that the covariance matrix of $Y = XU$ is D.

2 KPCA

1. Why the basis vectors $\{\alpha_i\}$ are not plotted as in PCA ?

3 ICA

1. Center and whitening the input data are two important preprocessing of ICA. Let X_C be the centered data X, Σ the covariance matrix of X and $\Sigma = UDU^T$ the eigen decomposition of X. Show that $X_W = D^{-\frac{1}{2}} U^T X_C$ and $Y = \Sigma^{-1/2} X_C$ can be used to whiten the centered data X_C. Remember that the whitening transformation $X_W = \mathcal{W} X_C$ transforms the data X_C such that $E[X_W X_W^T] = I$.

2. Why do you think it is useful whitening the data ? Hint: look at the covariance matrix of X and X_W and the number of parameters to estimate...

3. Let $X_W = \mathcal{W}X_C$ be the whitened and centered data. Let $S_W = W_W X_W$ be the ICA transformation. What would be the mixing matrix A_C for X_C, namely the centered (and not whitened) data ?

4 NNMF

1. Show that the KKT conditions for the NNMF problem:

$$\min_x f(x) \text{ s.t. } x \geq 0 \tag{1}$$

   are : $x \geq 0$, $\nabla f(x) \geq 0$ and $\nabla f(x)x = 0$

2. Looking at these equations, explain why the non-negativity constraint naturally leads to sparse factors