

Principal Component Analysis

Download the datafile *magic04.data* from the [UCI Machine Learning Repository](#). The dataset has 10 real attributes, and the last one is simply the class label, which is categorical, and which you will ignore for this assignment.

Principal Component

Compute the dominant eigenvalue and eigenvector of the covariance matrix Σ via the power-iteration method. One can compute the dominant eigenvector/-value of the covariance matrix iteratively as follows.

Let

$$\mathbf{x}_0 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

be the starting vector in \mathbb{R}^d , where $d = 10$ is the number of dimensions.

In each iteration i , we compute the new vector:

$$\mathbf{x}_i = \Sigma \mathbf{x}_{i-1}$$

We then find the element of \mathbf{x}_i that has the maximum absolute value, say at index m . For the next round, to avoid numerical issues with large values, we re-scale \mathbf{x}_i by dividing all elements by x_{im} , so that the largest value is always 1 before we begin the next iteration.

To test convergence, you may compute the norm of the difference between the scaled vectors from the current iteration and the previous one, and you can stop if this norm falls below some threshold, say 0.0001. That is, stop if

$$\|\mathbf{x}_i - \mathbf{x}_{i-1}\| < 0.0001$$

For the final eigen-vector, make sure to normalize it, so that it has unit length. Also, the ratio $\frac{x_{im}}{x_{i-1,m}}$ gives you the largest eigenvalue.

First Two Principal Components

Now compute the first two principal components (PCs) of the covariance matrix Σ using a generalization of the above iterative method.

Let \mathbf{X}_0 be a $d \times 2$ matrix with two non-zero column vectors of length d , with unit length. We will iteratively multiply \mathbf{X}_0 with Σ on the left.

The first column will not be modified, but the second column will be orthogonalized with respect to the first one by subtracting its projection along the first column (see section 1.3.3 in chapter 1). That is, let \mathbf{a} and \mathbf{b} denote the first and second column of \mathbf{X}_1 , where

$$\mathbf{X}_1 = \Sigma \mathbf{X}_0$$

Then we orthogonalize \mathbf{b} as follows:

$$\mathbf{b} = \mathbf{b} - \left(\frac{\mathbf{b}^T \mathbf{a}}{\mathbf{a}^T \mathbf{a}} \right) \mathbf{a}$$

After this \mathbf{b} is guaranteed to be orthogonal to \mathbf{a} . This will yield the final matrix \mathbf{X}_1 with the two column vectors denoting the current estimates for the first and second eigenvectors (or the principal components). Before the next iteration, normalize each column to be unit length, and repeat the whole process. That is, from \mathbf{X}_1 obtain \mathbf{X}_2 and so on, until convergence.

To test for convergence, you can look at the total absolute difference between \mathbf{X}_i and \mathbf{X}_{i-1} . If the difference is less than some threshold, use $\epsilon = 0.001$, then we stop.

Once you have obtained the two principal components: \mathbf{u}_1 and \mathbf{u}_2 , project each of the original points \mathbf{x}_i onto those two vectors, to obtain the new projected points in 2D (see Eq. (7.6)). Plot these projected points in the two PC dimensions.