Principle Component Analysis

Aim: Find a direction of projection that maximize the data's variance

How does the variance looks like?

Generally

Variance/Covariance formula:

$$Var(X) = Cov(X, X) = \frac{1}{N} \sum_{i=1}^{N} \left(X_i - \bar{X}\right) \left(X_i - \bar{X}\right)$$

In PCA, first step is to **centralize** the dataset.

$$X_i = X_i - \bar{X}$$

Therefore, for simplicity purpose, the variance formula becomes:

$$Var(X) = \frac{1}{N} \sum_{i=1}^{N} (X_i) (X_i)$$

PCA

Note that x must be **centralized**!
The projected data, y :

$$y = x_i \cdot w$$

$$\sigma_w^2 = \frac{1}{n} \sum_i (x_i \cdot w_i)^2 = \frac{1}{n} (xw)^T (xw) = \frac{1}{n} w^T x^T xw$$

$$\sigma_w^2 = w^T \frac{x^T x}{n} w = w^T C w$$

C = Covariance matrix

Find the Principle axes of the dataset

Objective function:

$$\max_{w^T w = 1} w^T C w$$

> Since it’s a maximizing function, the purpose of w has unit length to avoid w->inifinity

Lagrangian:

$$L(w, \lambda) = w^T C w - \lambda(w^T w - 1)$$

From KKT condition:

$$\frac{\partial L(w, \lambda)}{\partial w} = 2Cw - 2\lambda w = 0$$

$$Cw = \lambda w$$

> where w = eigenvectors, lambda = eigenvalue

Some final remarks…

- Eigenvector w with the largest eigenvalue(since it’s maximizing function) is the projection we’re looking for!
- Eigenvectors of the Covariance Matrix, `C` are the principle components of the dataset.
- The direction of second principle component is orthogonal to the first component with the most variance.

Reference

- CMU Lecture Notes
- Why eigenvector w taken to be unit norm?

Author

Author: Lee Zhicheng
Date: 30/04/2020