PCA and SVD

1 What does PCA try to achieve?

Maximum variance perspective

Project remainig data to obtain maximum variance

Reconstruction perspective

Theorem 1. *Suppose we want to find an orthonormal set of L basis vectors $\boldsymbol{w}_j \in \mathbb{R}^D$ (they form columns of $\boldsymbol{W} \in \mathbb{R}^{D \times L}$) such that the following reconstruction objective is minimized:*

$$J(\boldsymbol{W}, \boldsymbol{Z}) = \|\boldsymbol{X} - \boldsymbol{W}\boldsymbol{Z}\|_F^2.$$

The solution is to set $\hat{\boldsymbol{W}}$ to be the L eigenvectors of the empirical covariance matrix $\hat{\boldsymbol{\Sigma}} = \boldsymbol{X}^T\boldsymbol{X}$.

One can show that the errors can be decomposed due to triangles, then it's again the maximum variance perspective kicks in.

Again, in either cases, the principal components are by definition orthonormal.y

2 How to compute the principal components?

Eigendecomposition of covariance matrix

3 Computing principal compoents using SVD

4 SVD computes more than just the principal components