

MACHINE LEARNING FOR SIGNAL PROCESSING

13-1-2025

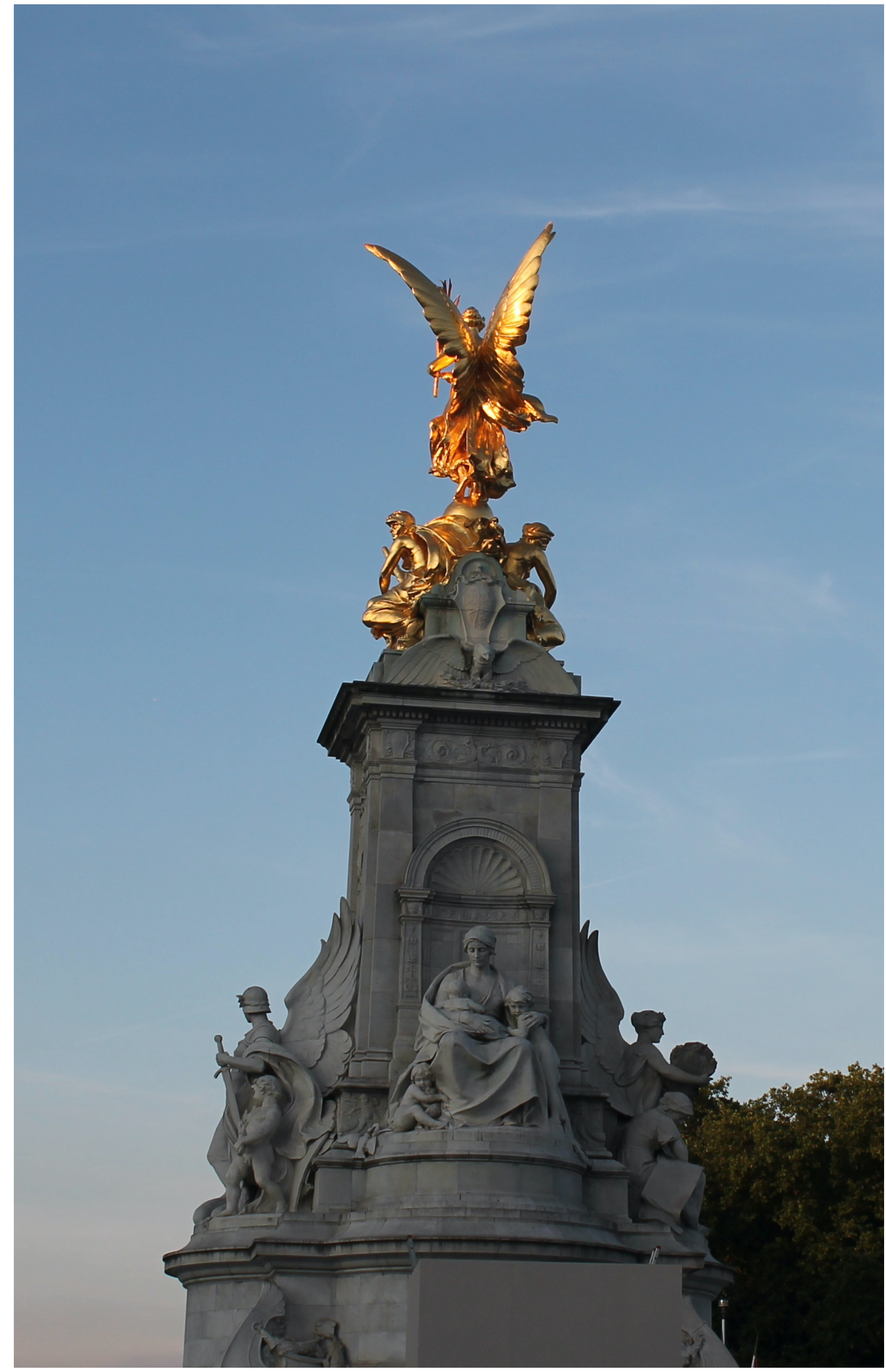
Sriram Ganapathy

*LEAP lab, Electrical Engineering, Indian Institute of Science,
sriramg@iisc.ac.in*

Viveka Salinamakki, Varada R.

*LEAP lab, Electrical Engineering, Indian Institute of Science.
viveka.sg@gmail.com varadar2000@gmail.com*

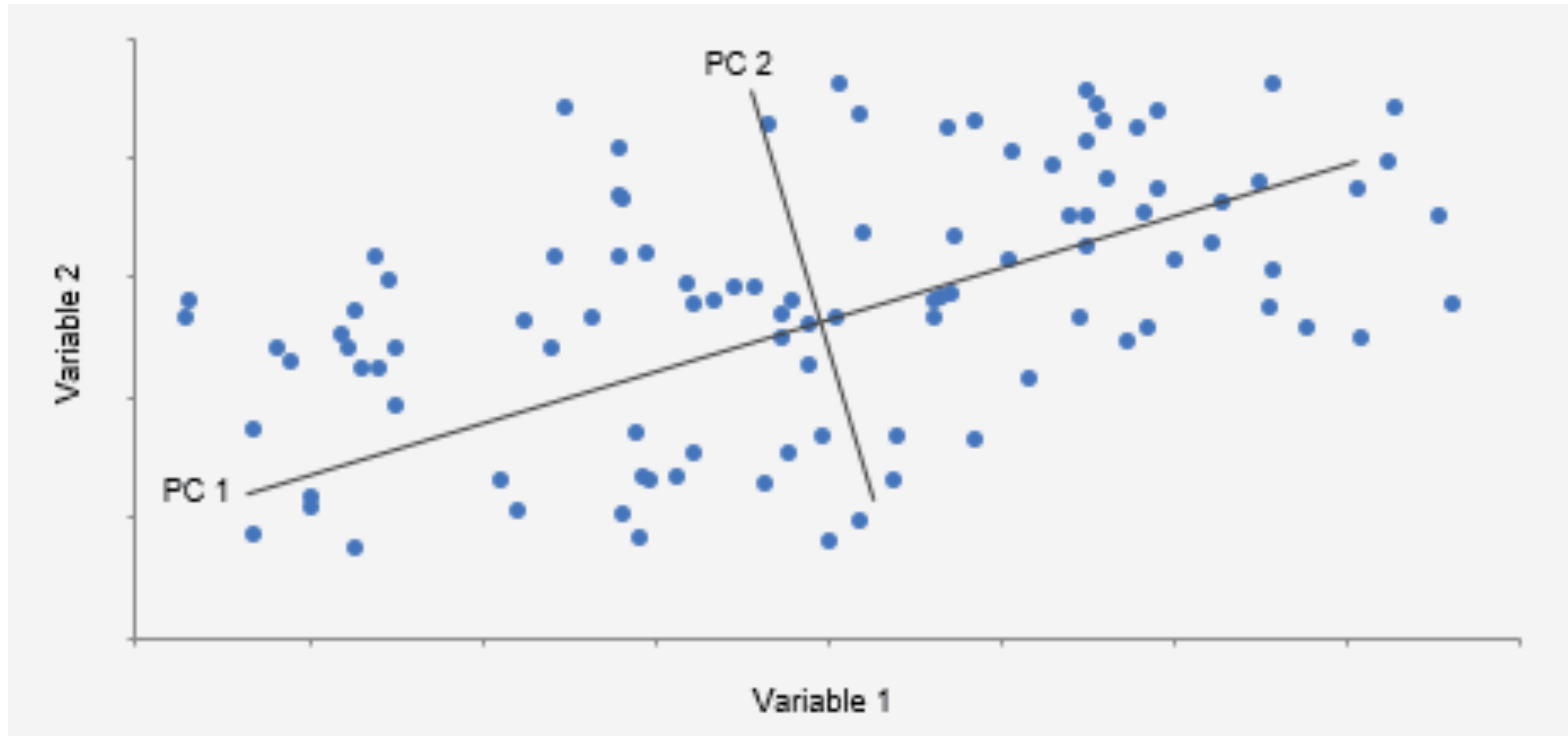
<http://leap.ee.iisc.ac.in/sriram/teaching/MLSP25/>



PRINCIPAL COMPONENT ANALYSIS

- ❖ Reducing the data \mathbf{x}_n of dimension D to lower dimension
- ❖ Projecting the data into subspace which preserves maximum data variance
- ✓ Maximize variance in projected space $M < D$
- ❖ Equivalent formulated as minimizing the error between the original and projected data points.

PCA EXAMPLE



PRINCIPAL COMPONENT ANALYSIS

- ❖ First M eigenvectors of data covariance matrix

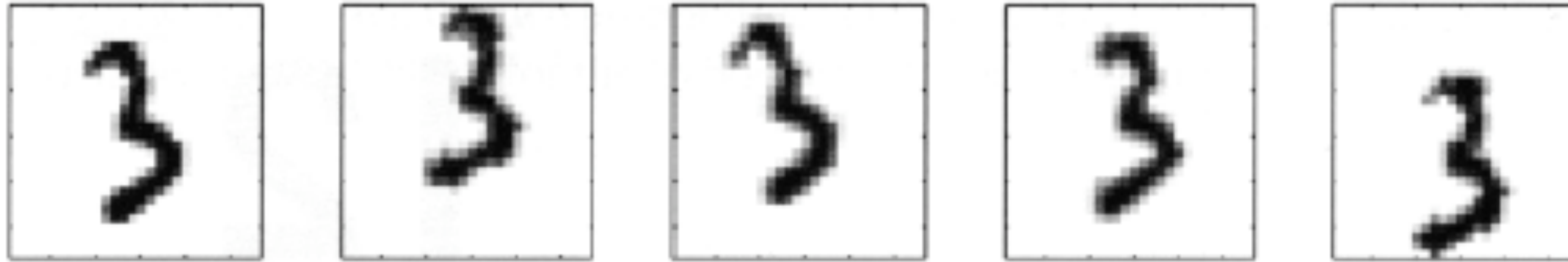
$$S = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}_n - \bar{\mathbf{x}})(\mathbf{x}_n - \bar{\mathbf{x}})^T$$

- ❖ Residual error from PCA

$$J = \sum_{i=M+1}^D \lambda_i$$

PRML - C. Bishop (Sec. 12.1)

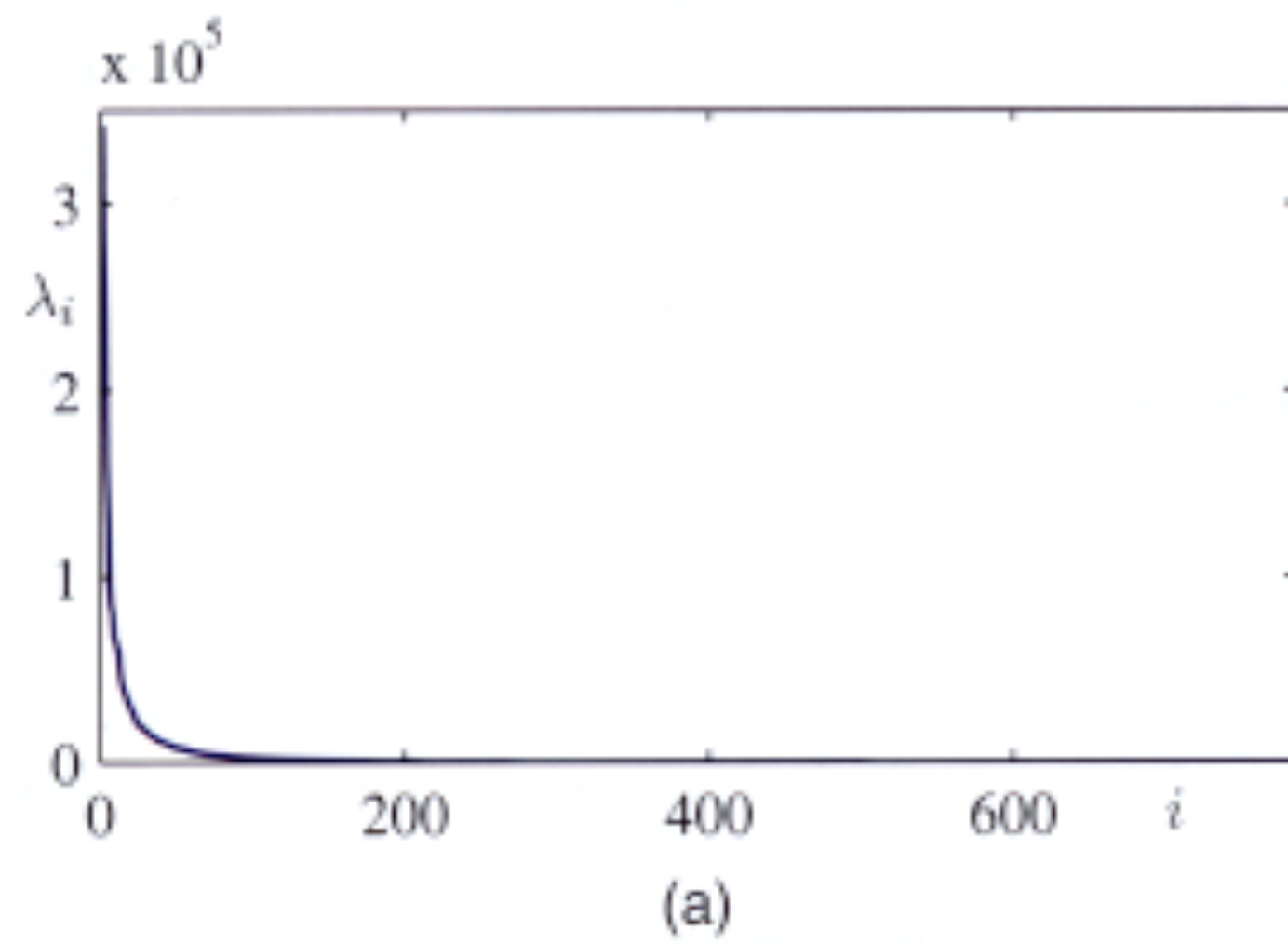
- ❖ First eigenvectors of data covariance matrix



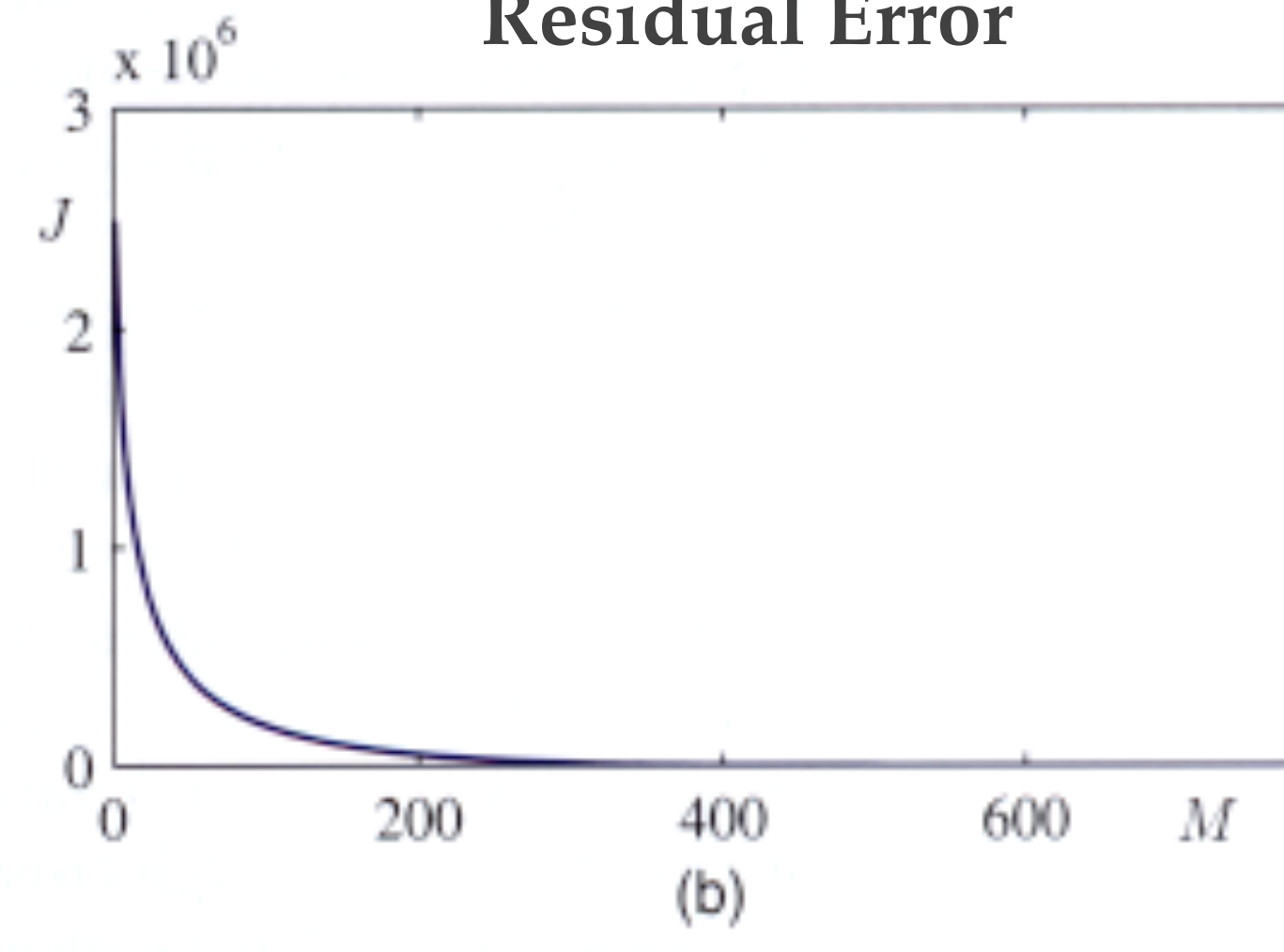
- ❖ Residual error from PCA

Handwritten digits used for PCA training...

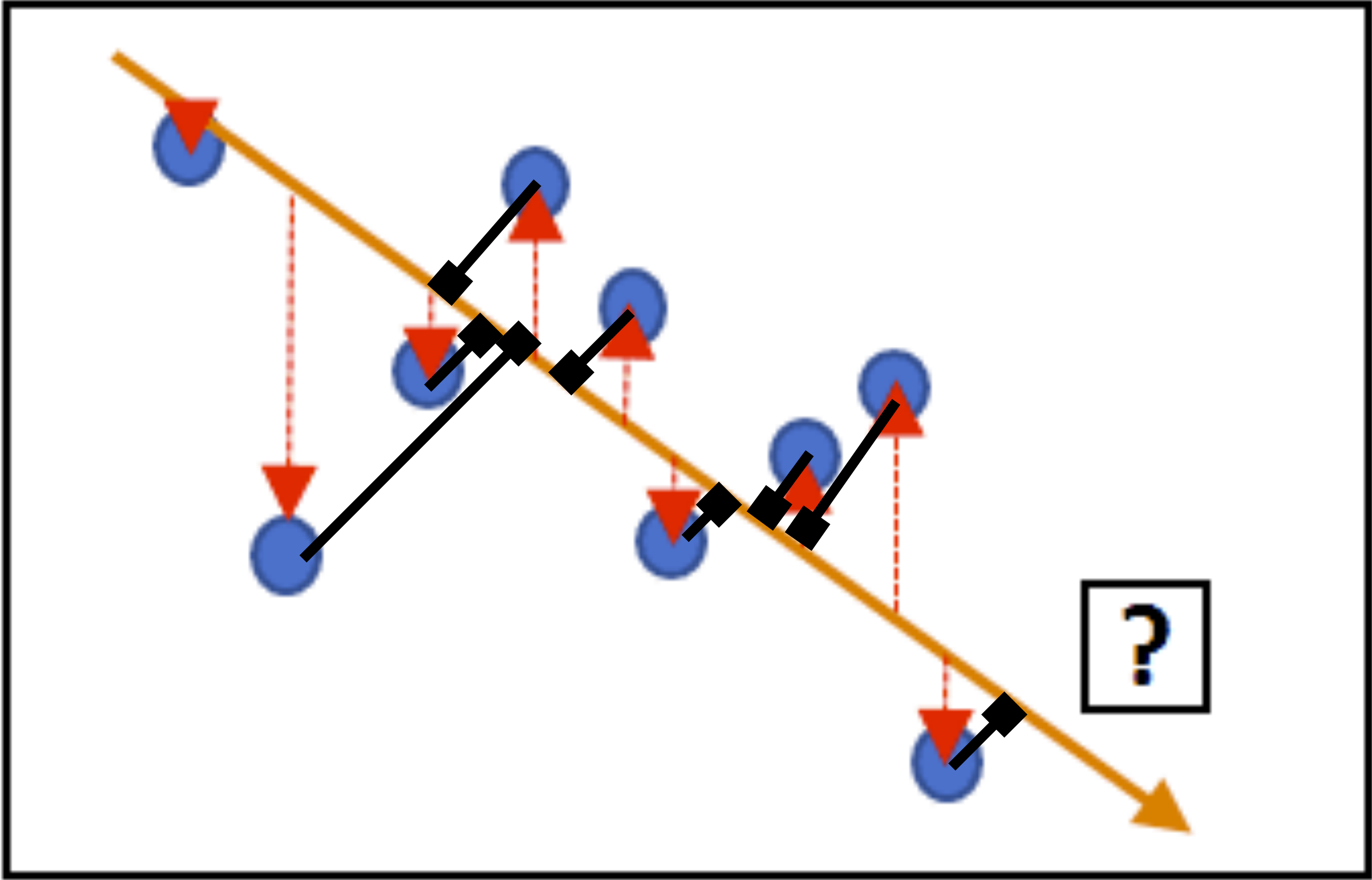
Eigen Values



Residual Error



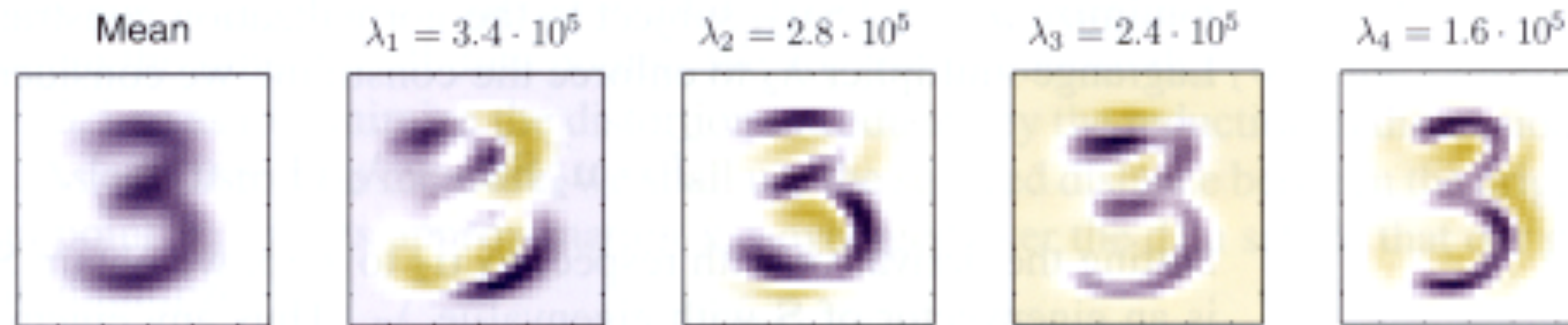
DIRECTION OF MAXIMUM VARIANCE



PCA
Regression

PCA - RECONSTRUCTION

Eigenvectors

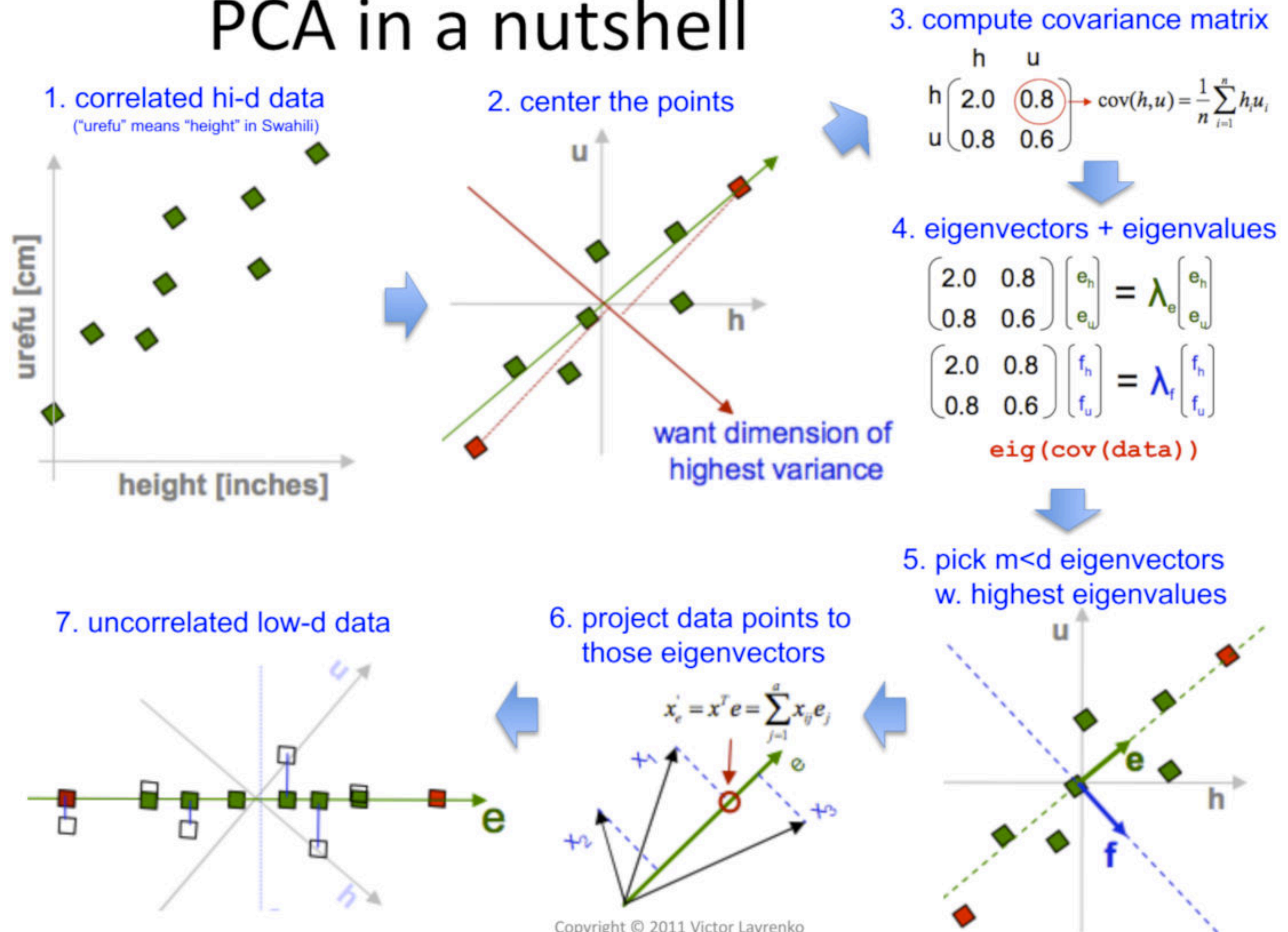


PCA - Reconstruction

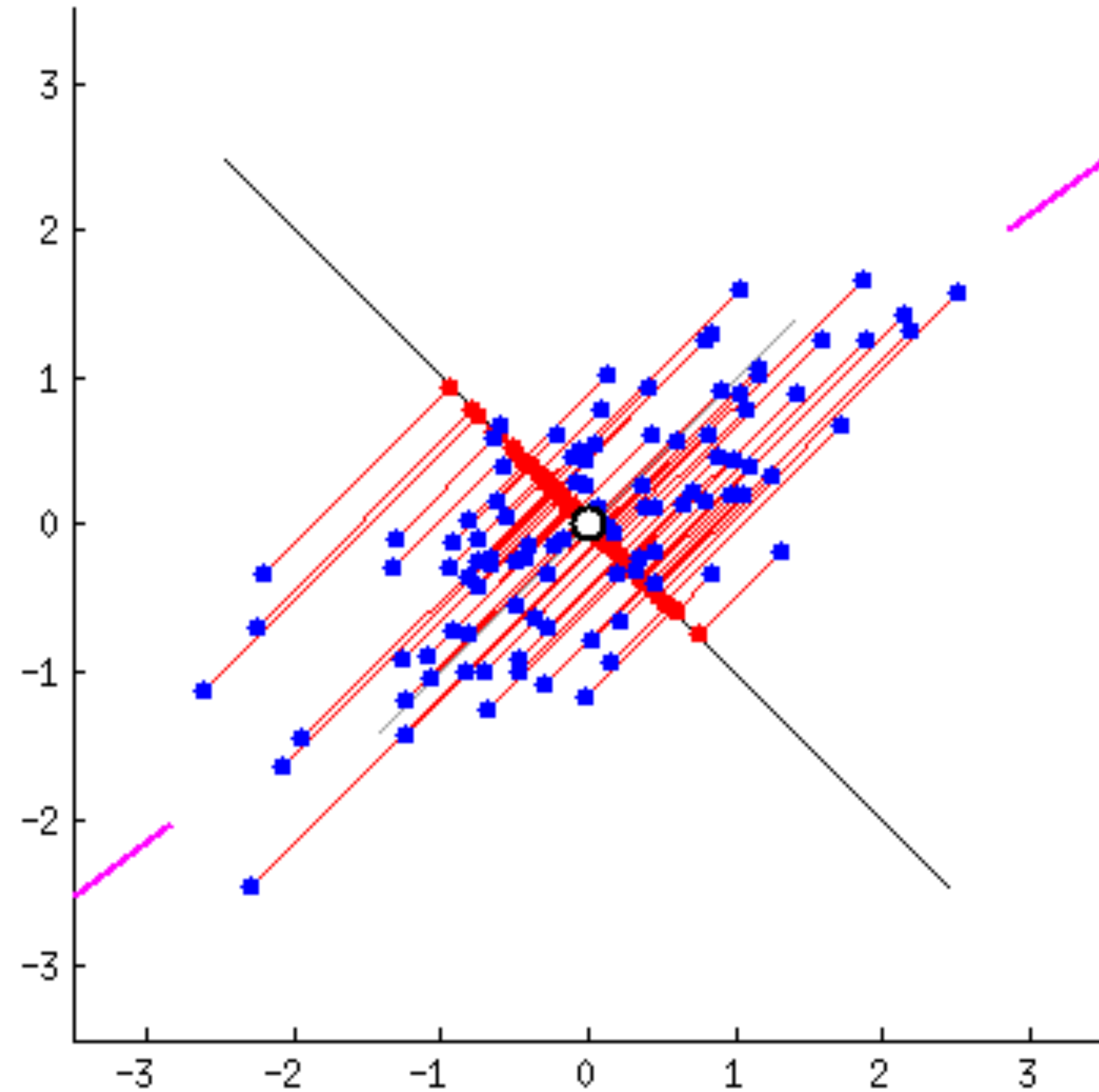
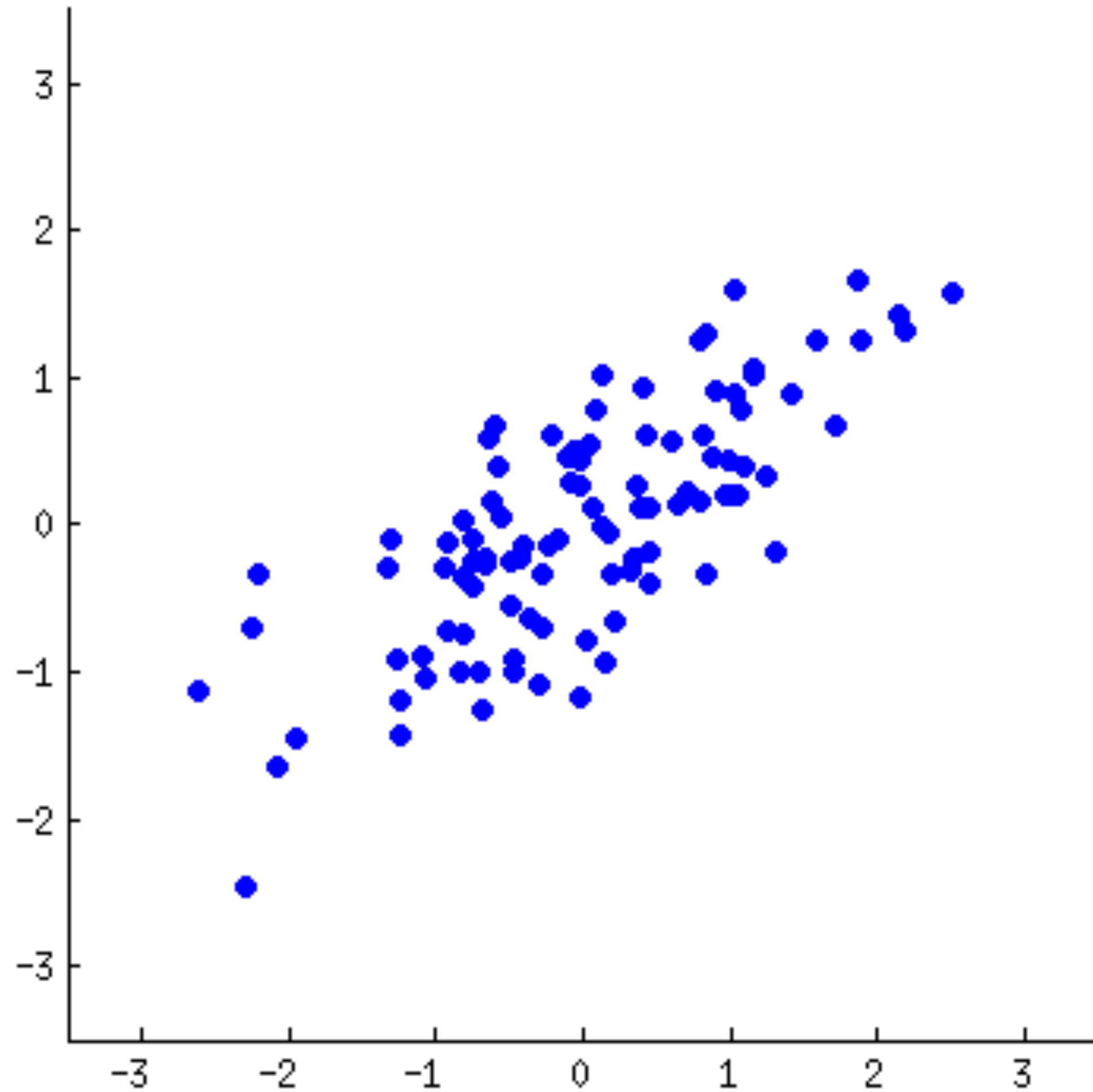


PCA SUMMARY

PCA in a nutshell

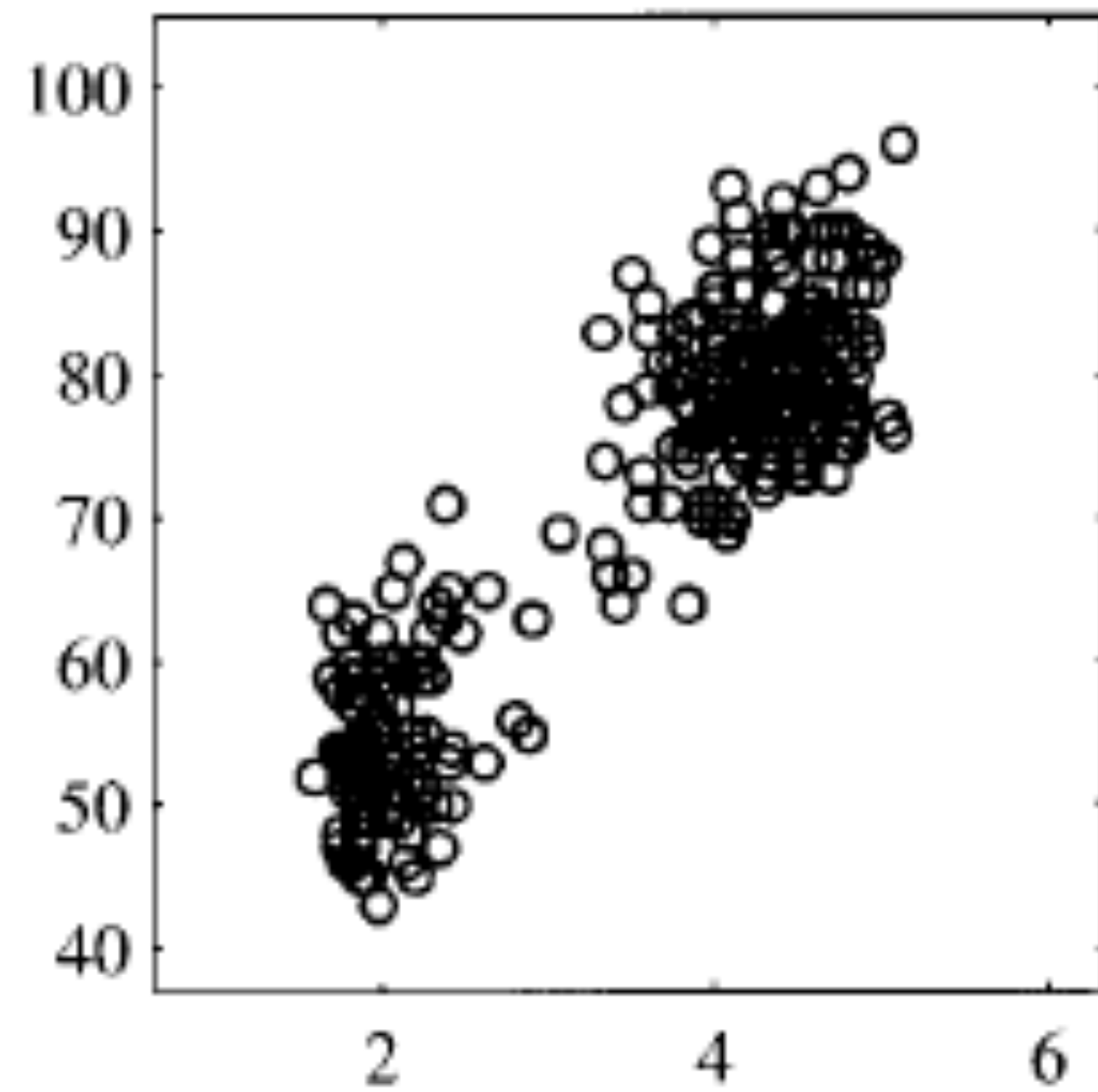


VISUALIZING PCA

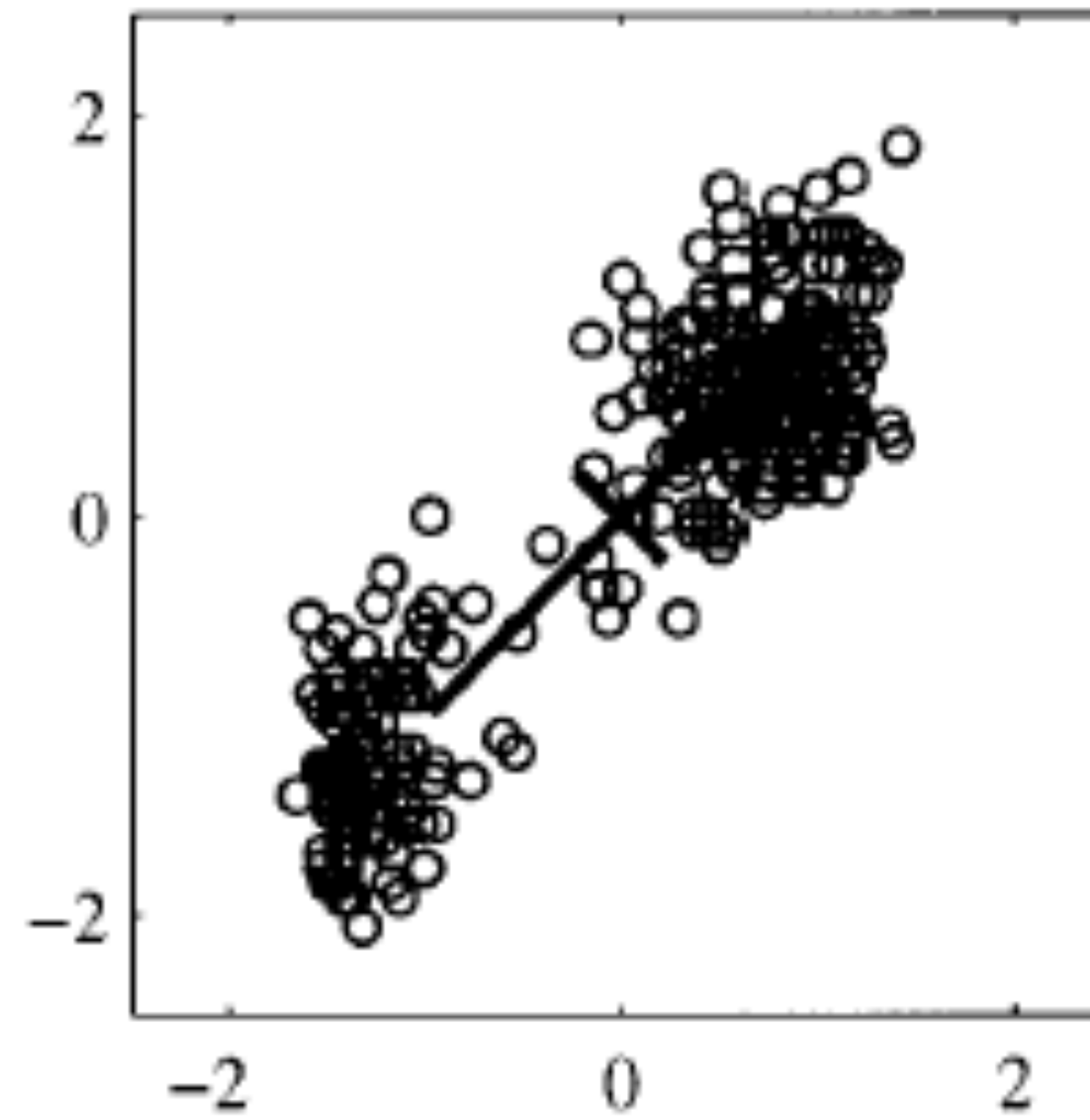


WHITENING THE DATA

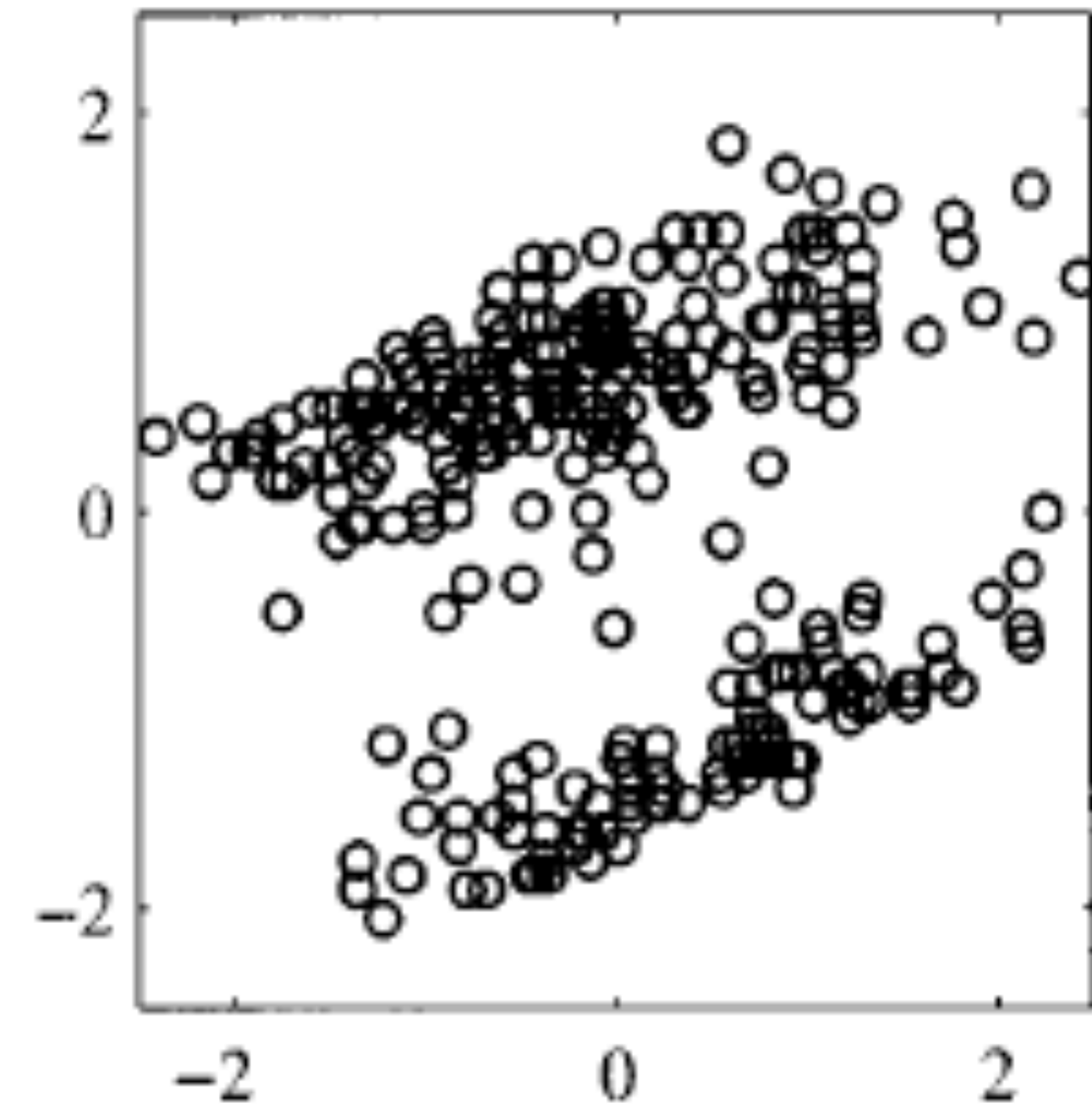
Original Data



Standardization



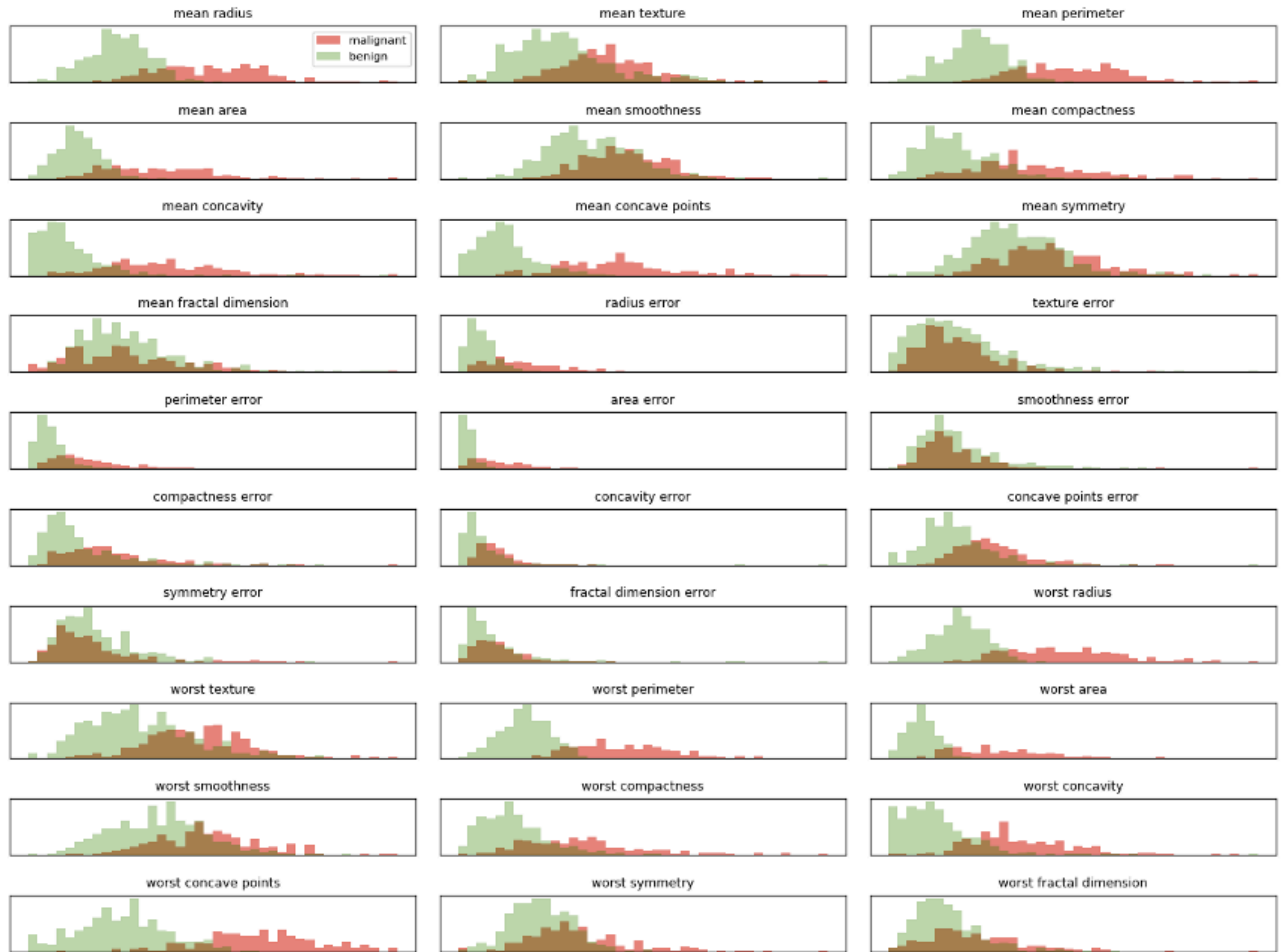
Whitening



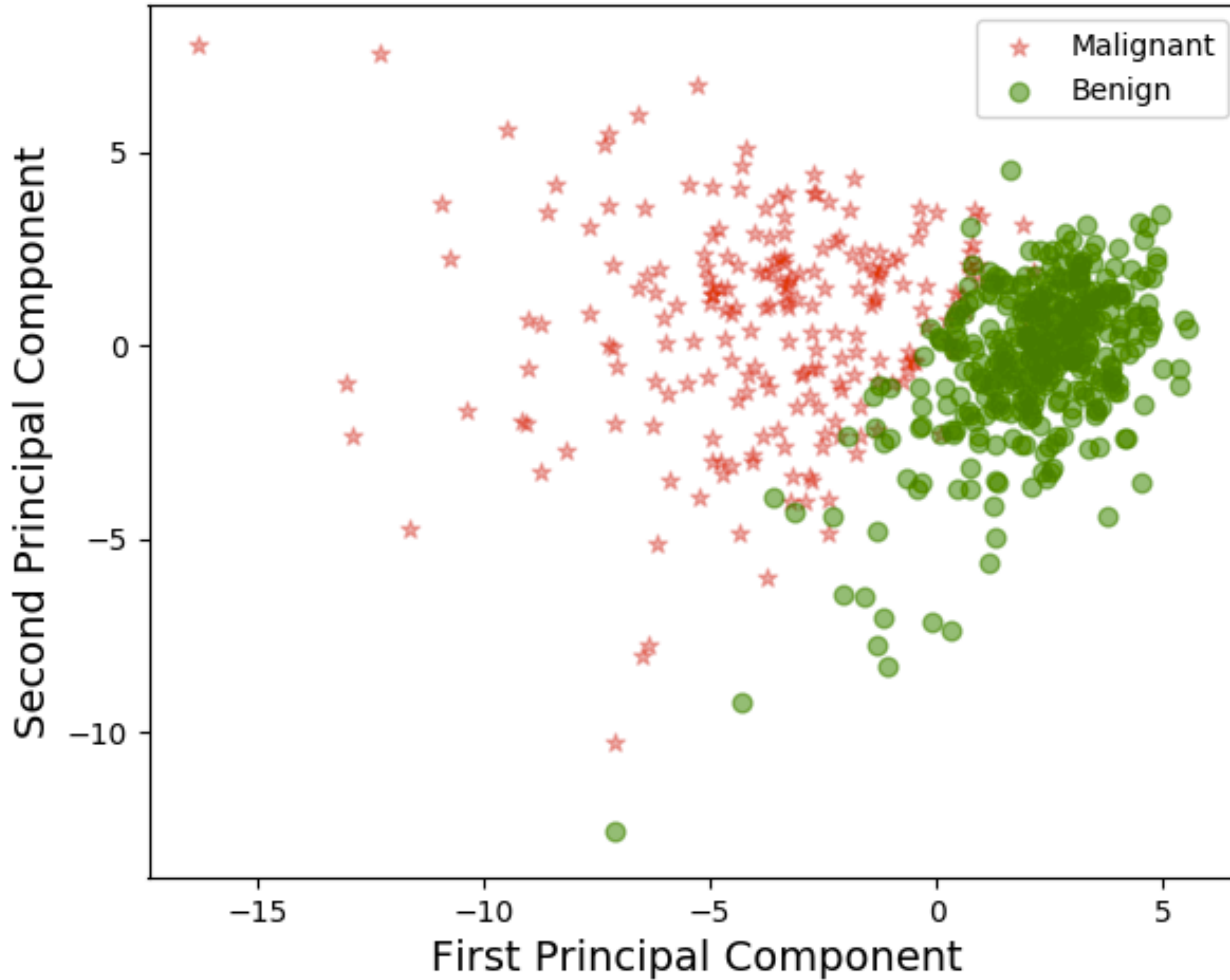
APPLICATION

- ❖ Wisconsin Cancer dataset ([https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Diagnostic\)](https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)))
- ❖ 569 participants
- ❖ 212 (M) 357 (B)
- ❖ 30 features —> digitized image of a fine needle aspirate (FNA) of a breast mass. The features describe characteristics of the cell nuclei present in the image.

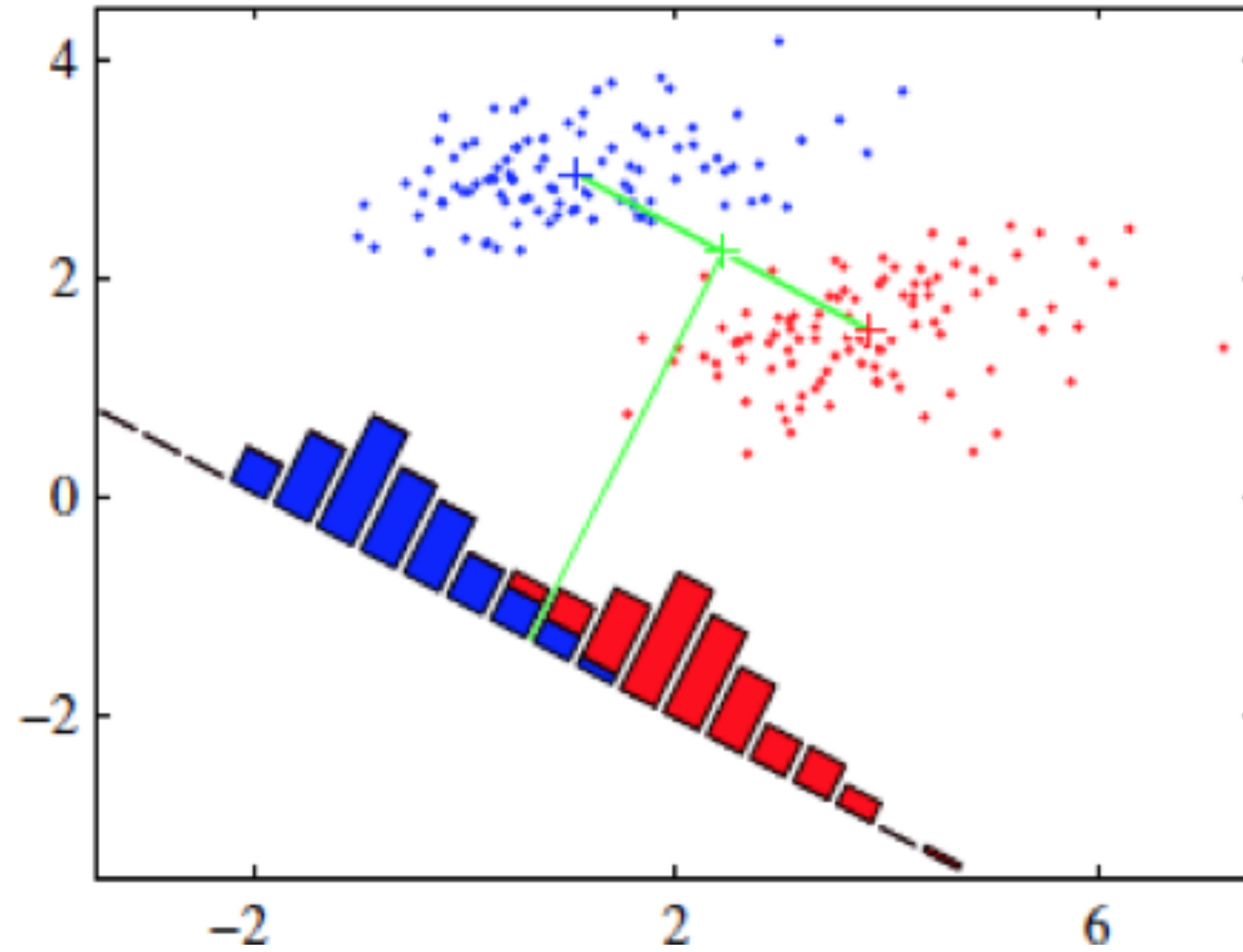
Raw Features



PCA



WITHOUT THE WITHIN CLASS FACTOR



LINEAR DISCRIMINANT ANALYSIS

❖ Generalized Eigenvalue problem

Find a linear transform $f(\mathbf{x}) = \mathbf{w}^T \mathbf{x}$ with a criterion which maximizes the class separation

- Maximize the between class distance in the projected space while minimizing the within class covariance

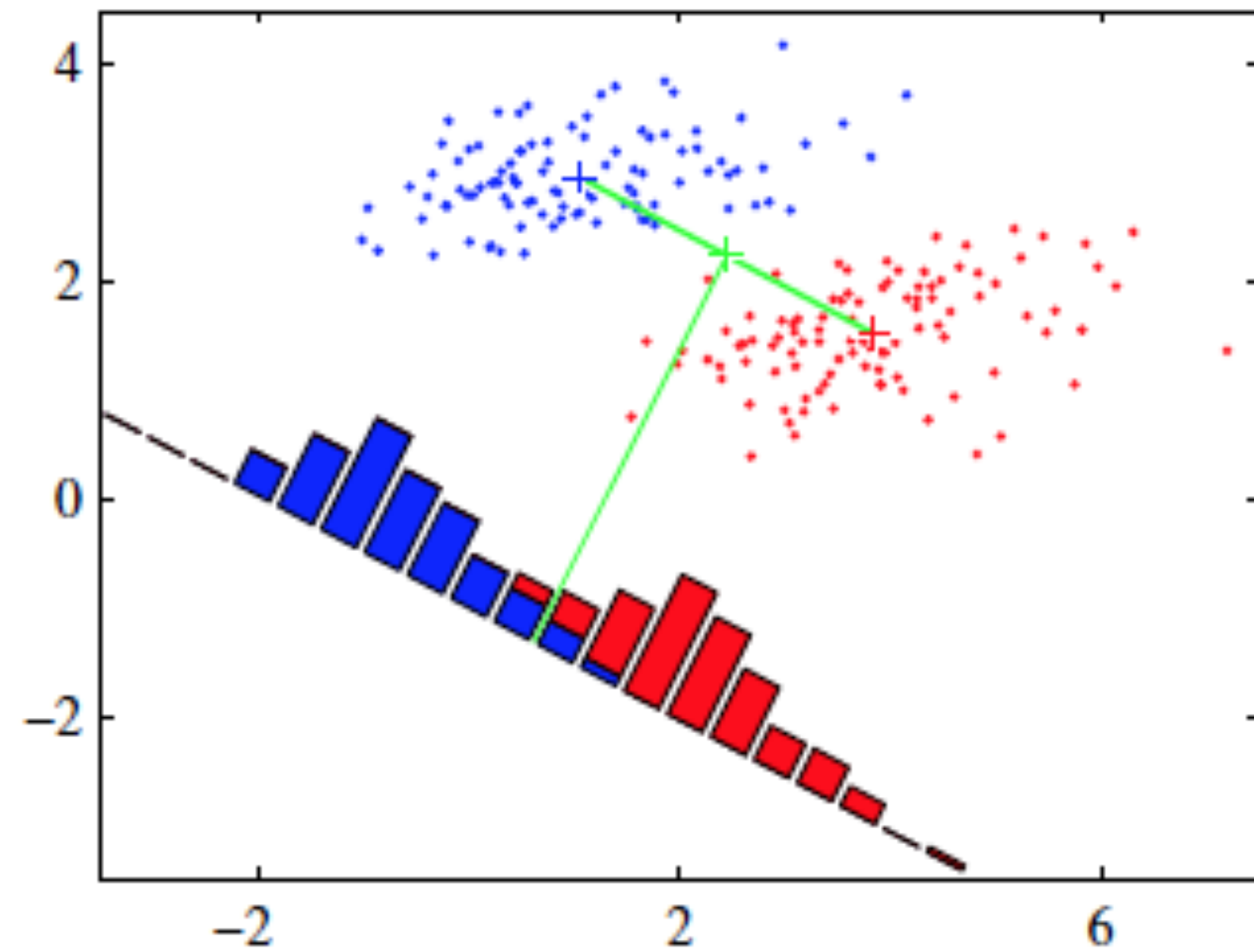
$$J = \frac{\mathbf{w}^T \mathbf{S}_b \mathbf{w}}{\mathbf{w}^T \mathbf{S}_w \mathbf{w}}$$

$$\mathbf{S}_b = \sum_{k=1}^K N_k (\mathbf{m}_k - \mathbf{m})(\mathbf{m}_k - \mathbf{m})^T \quad \mathbf{S}_w = \sum_{k=1}^K \sum_{n \in C_k} (\mathbf{x}_n - \mathbf{m}_k)(\mathbf{x}_n - \mathbf{m}_k)^T$$

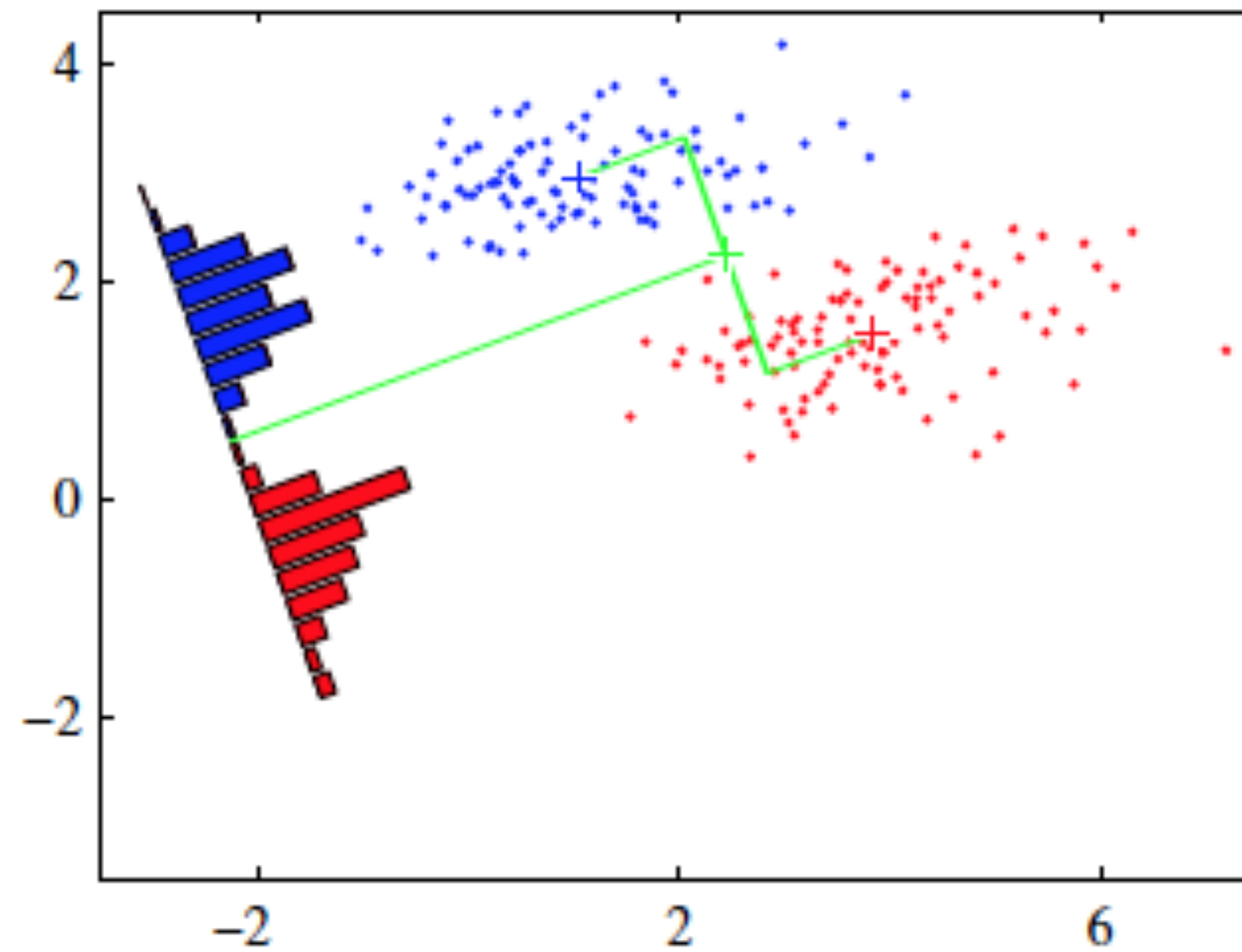
$$\mathbf{S}_w^{-1} \mathbf{S}_b$$

LINEAR DISCRIMINANT ANALYSIS

Projecting on line joining means



Fisher Discriminant



THANK YOU

Sriram Ganapathy and TA team
LEAP lab, C328, EE, IISc
sriramg@iisc.ac.in

