Introduction to Machine Learning

Brown University CSCI 1950-F, Spring 2011
Prof. Erik Sudderth

Lecture 20: Principal Components Analysis & Factor Analysis

Many figures courtesy Kevin Murphy's textbook,
Machine Learning: A Probabilistic Perspective

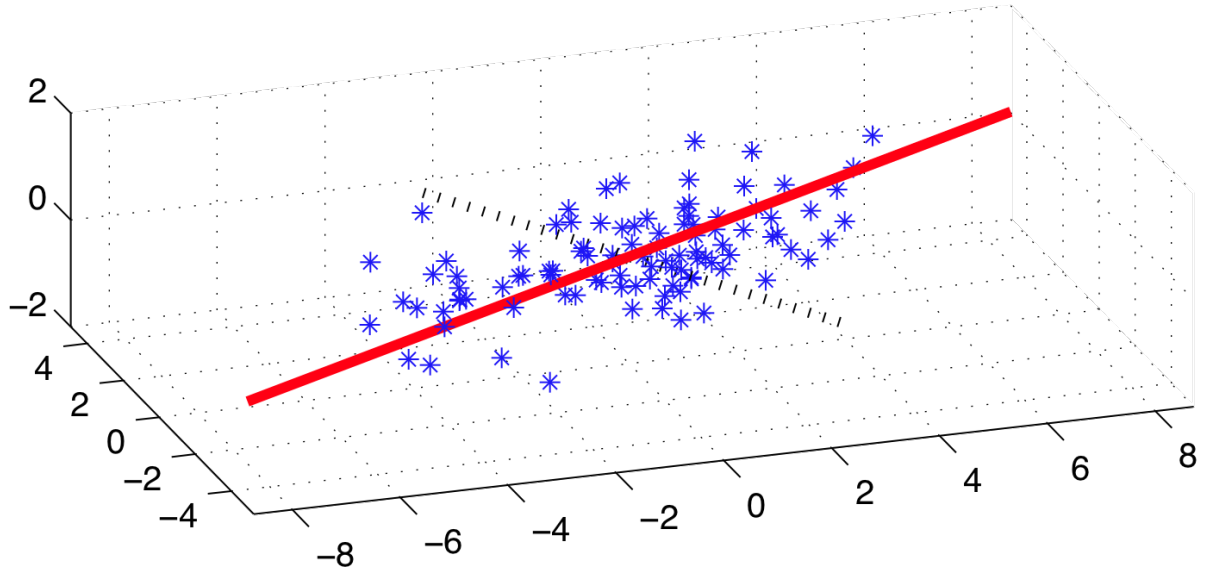
Dimensionality Reduction

	Supervised Learning	*Unsupervised Learning*
Discrete	classification or categorization	clustering
Continuous	regression	dimensionality reduction

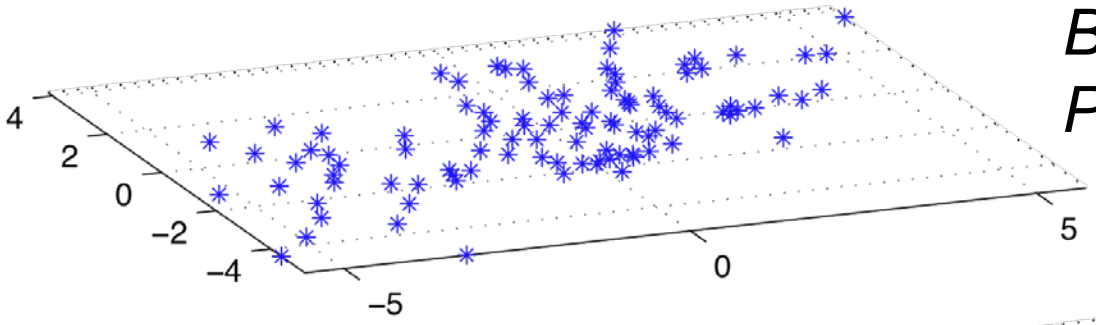
- **Goal:** Infer label/response y given only features x
- **Classical:** Find latent variables y good for *compression* of x
- **Probabilistic learning:** Estimate parameters of joint distribution p(x,y) which *maximize marginal probability* p(x)

Principal Components Analysis (PCA)

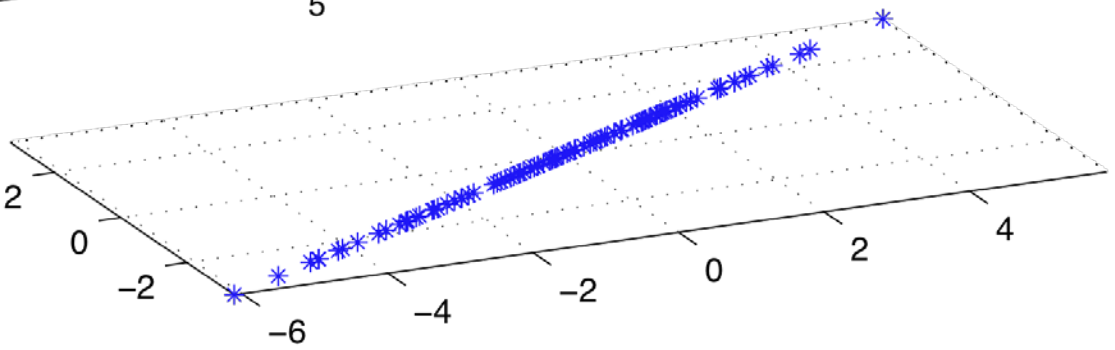
3D Data



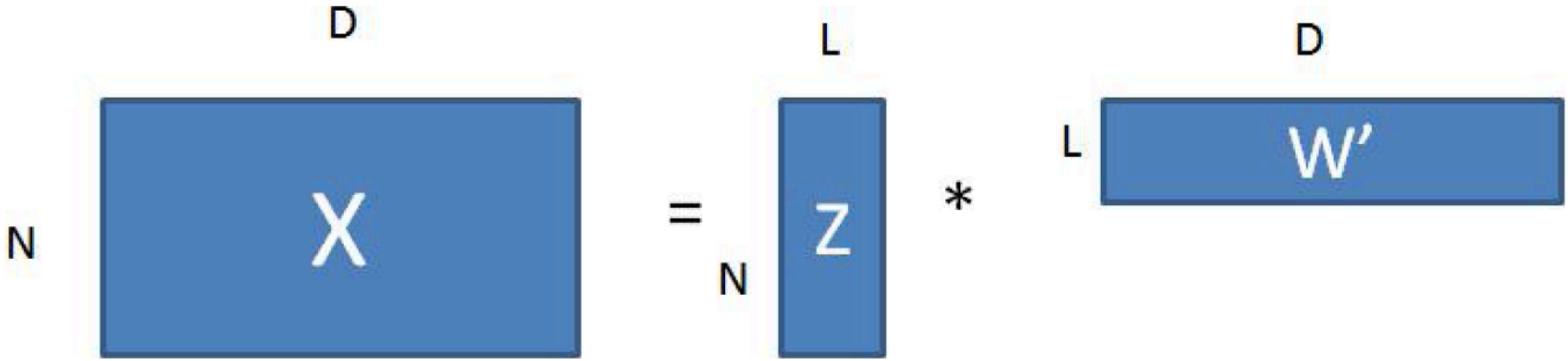
Best 2D Projection



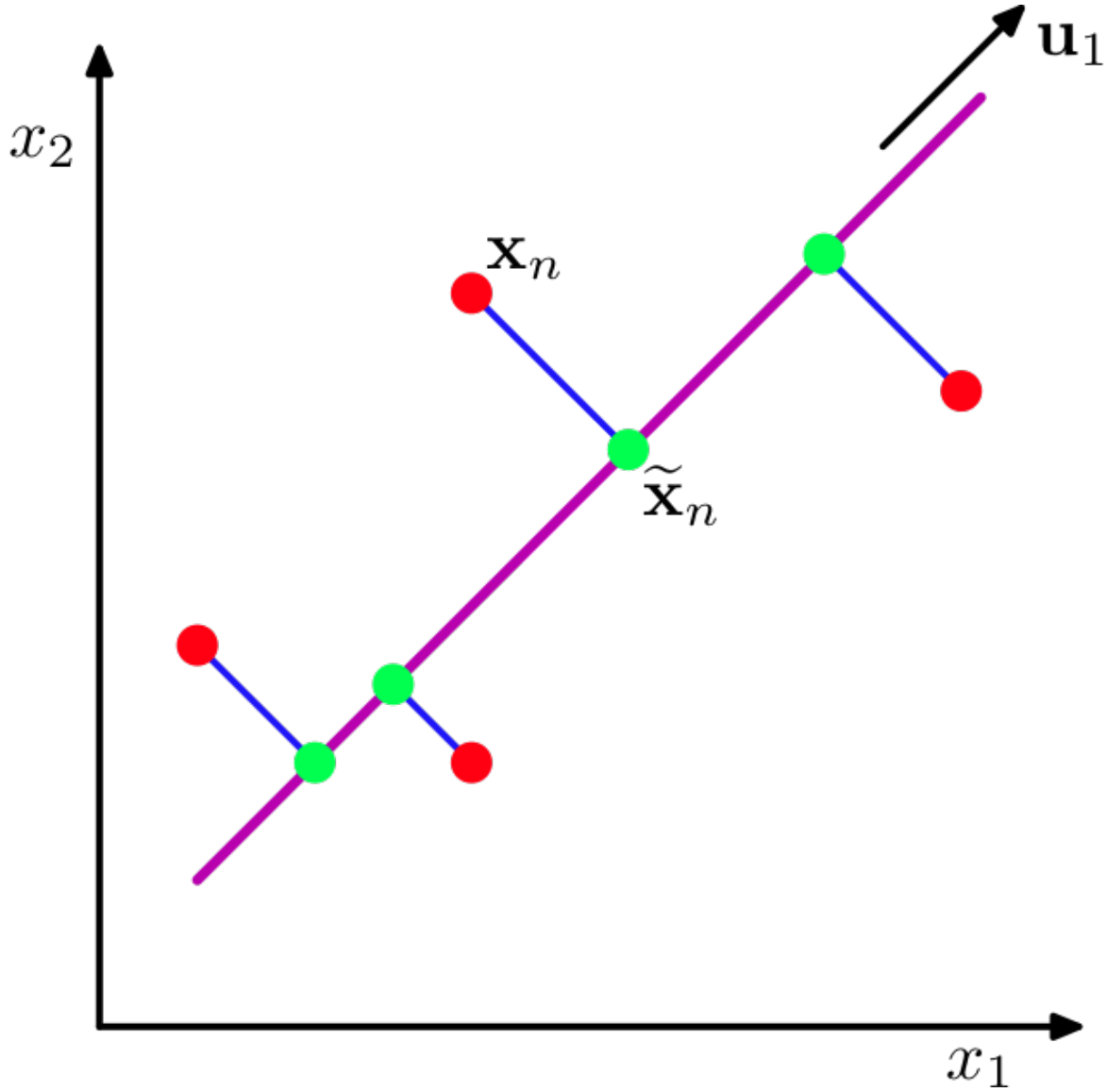
Best 1D Projection



PCA as Low Rank Approximation



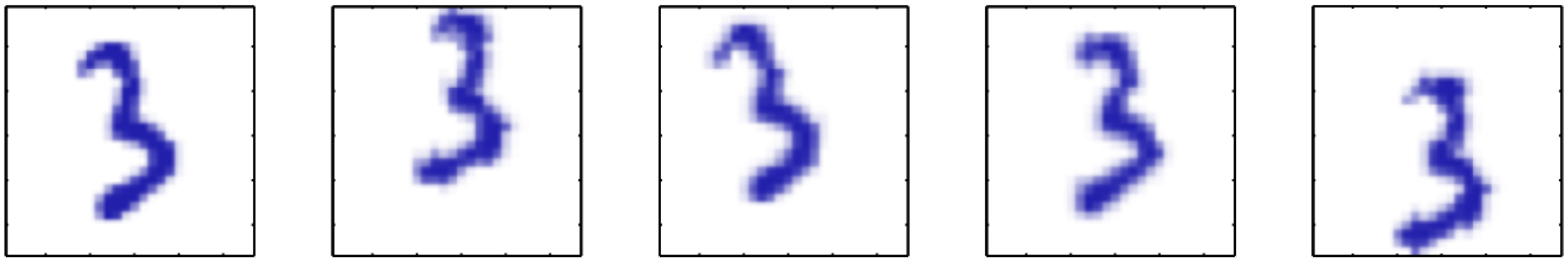
Maximizes Variance & Minimizes Error



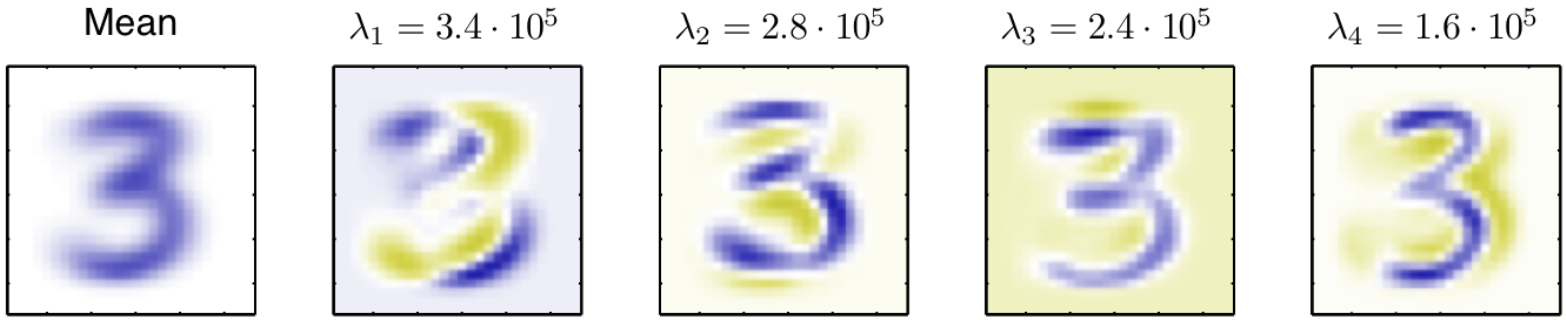
C. Bishop, Pattern Recognition & Machine Learning

PCA for Handwritten Digits

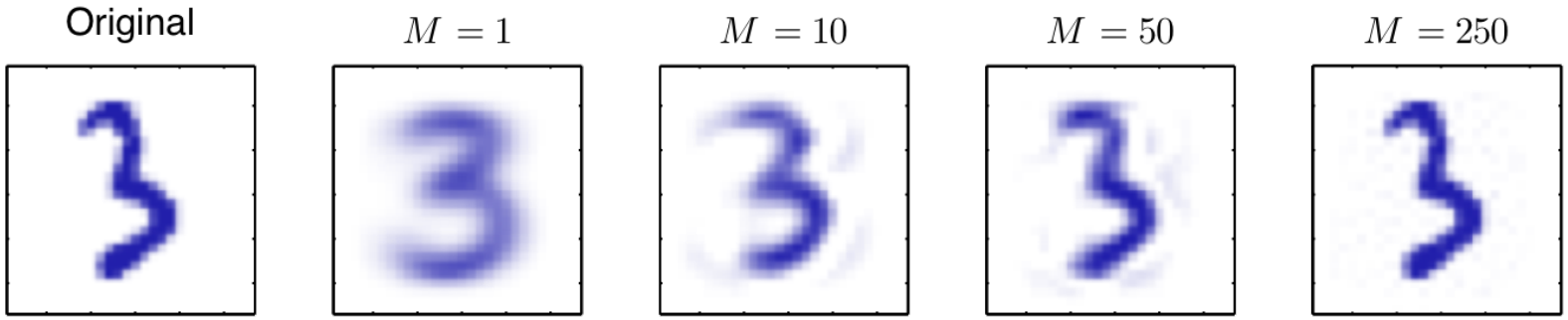
Data:
Randomly translated & rotated 3's



PCA Model:
Mean & top four eigenvectors

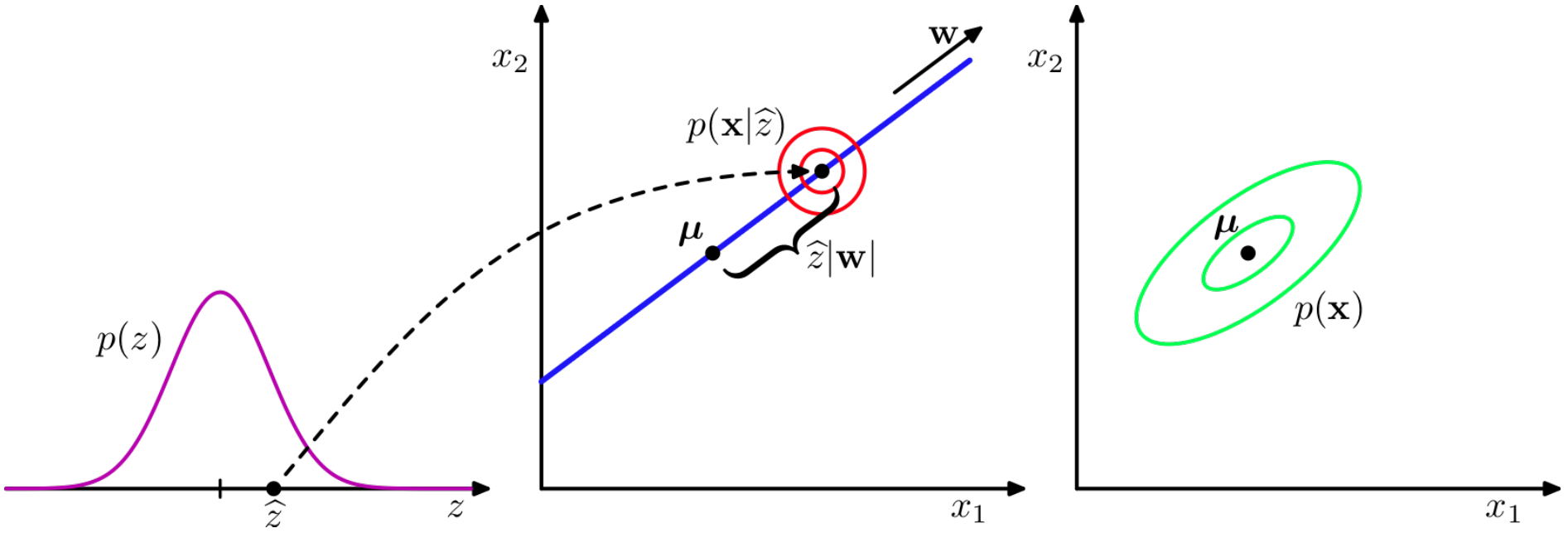


Compression:
Quality vs. latent dimension



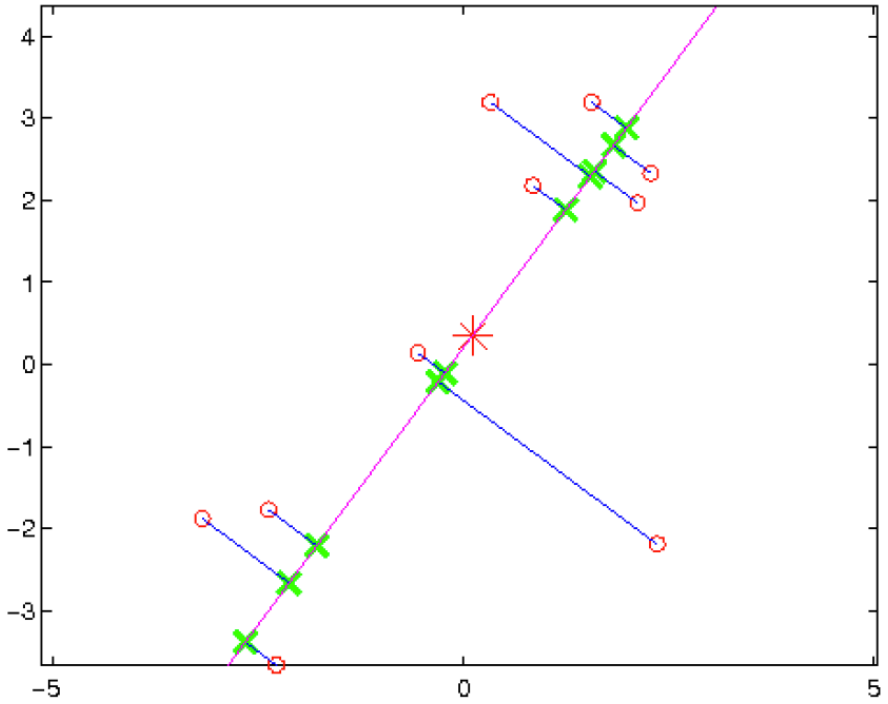
C. Bishop, Pattern Recognition & Machine Learning

Probabilistic PCA

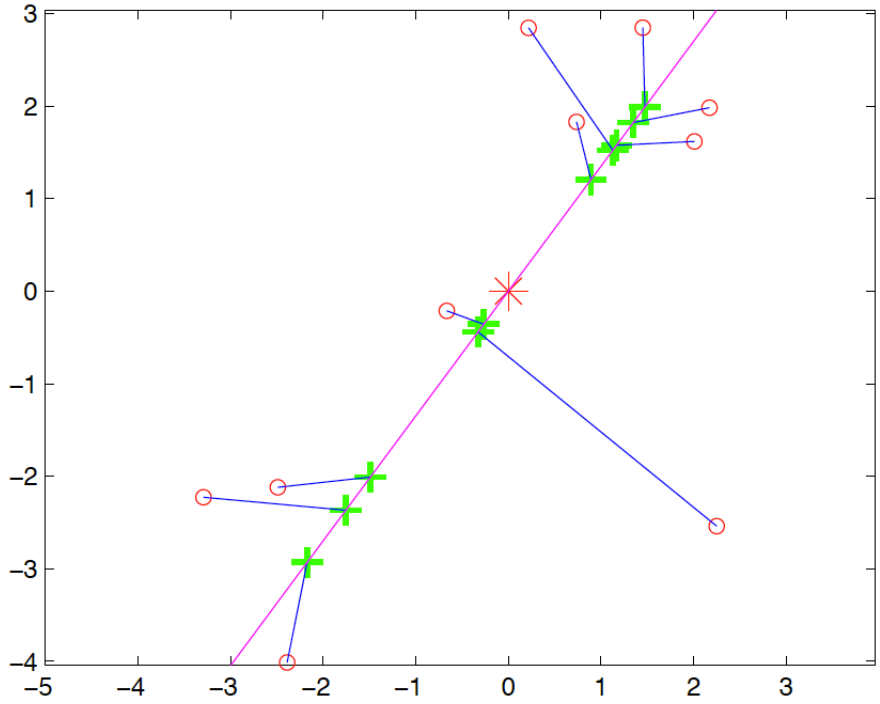


C. Bishop, Pattern Recognition & Machine Learning

PCA versus Probabilistic PCA

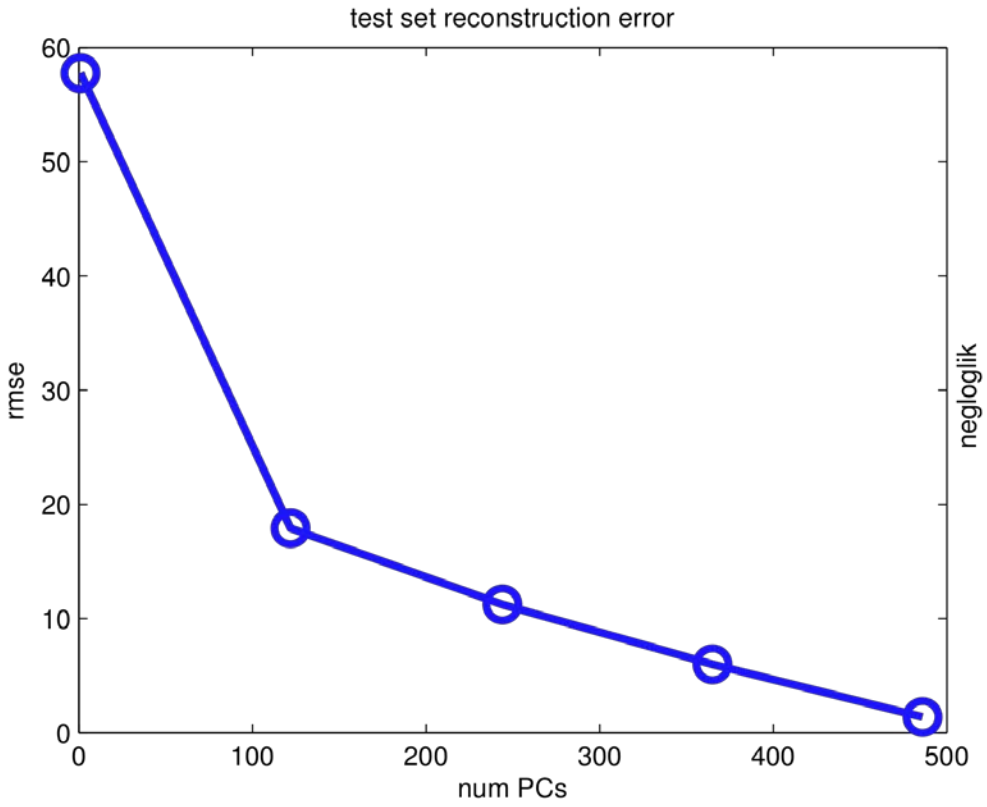


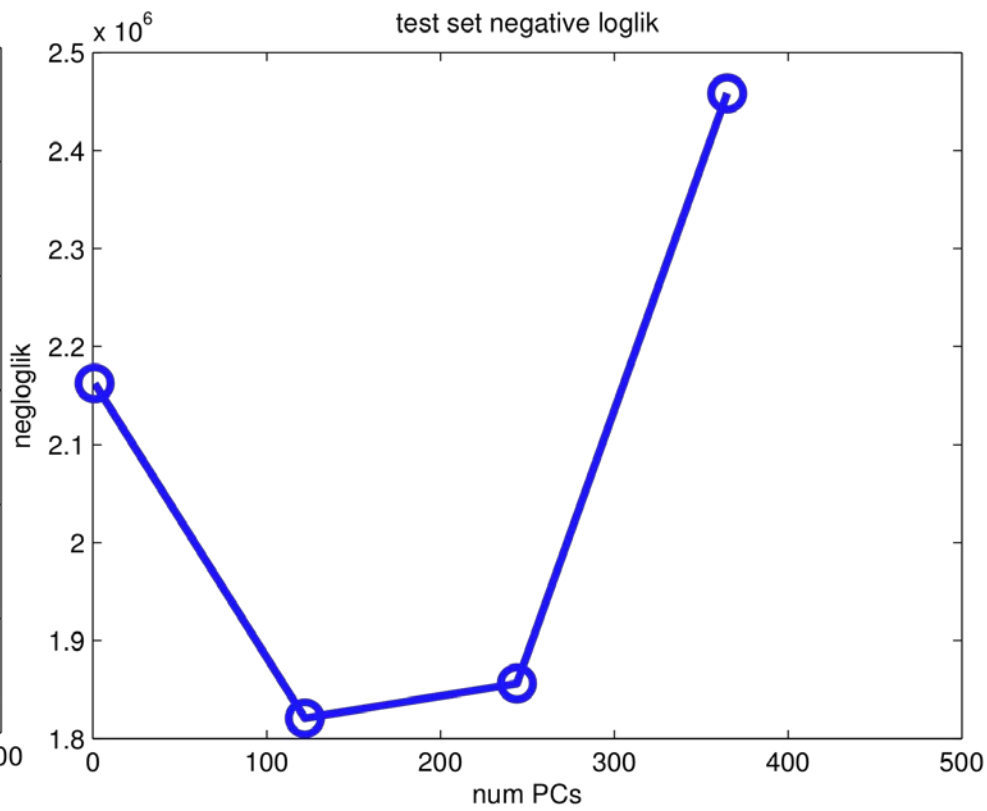
Standard PCA
(orthogonal projection)



Probabilistic PCA
(shrunk towards mean)

Prediction of Validation Data





Standard PCA
(reconstruction error)

Probabilistic PCA
(negative log likelihood)

Factor Analysis Example

Features of Cars in 2004

- Suggested retail price in USD
- Price to dealer in USD
- Engine size in liters
- Num. cylinders
- Horsepower
- City MPG
- Highway MPG
- Weight in pounds
- Wheelbase in inches
- Length in inches
- Width in inches