Machine Learning for Sensory Signals

Dimensionality Reduction

02-02-2017

Established 1911

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Recap

- Decision rule and inference rule in machine learning
- Minimizing the classification error
  - Maximum aposteriori rule
- Loss function and expected loss in classification
- Regression with mean square error loss
  - Conditional Expectation

Recap

- Three Approaches to ML
- Generative modeling
  - Estimating the likelihood and the prior density separately.
- Discriminative modeling
  - Directly modeling the class posterior probability.
- Discriminant function
  - No probabilistic interpretation.

Pros and Cons

- Generative Modeling
  - Useful for synthetic data generation
  - Outlier Detection
- Discriminative modeling
  - Typically better for classification
- Discriminant function
  - No posterior estimation (loss function cannot be incorporated directly)
  - Combining multiple models is harder.

Principal Component Analysis

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

Principal Component Analysis

- First M eigenvectors of data covariance matrix

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

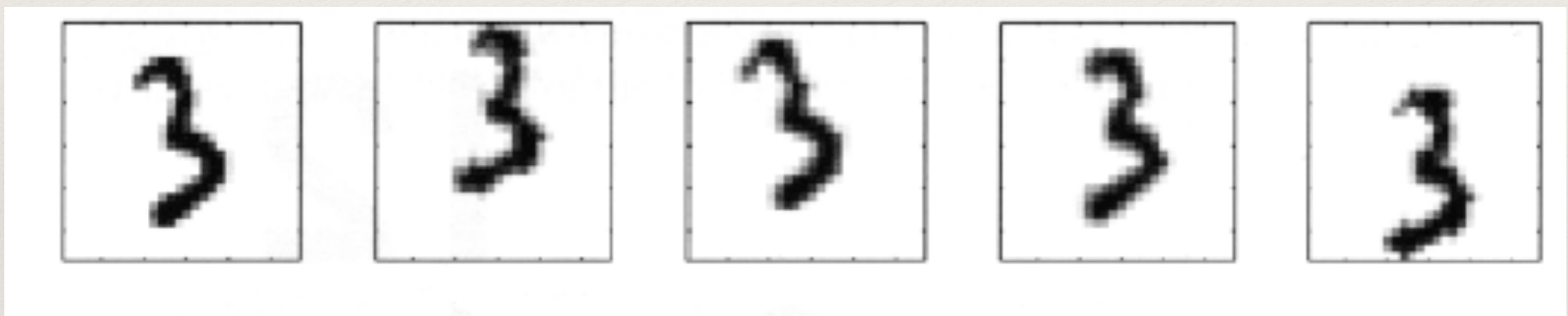
- Residual error from PCA

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

PRML - C. Bishop (Sec. 12.1)

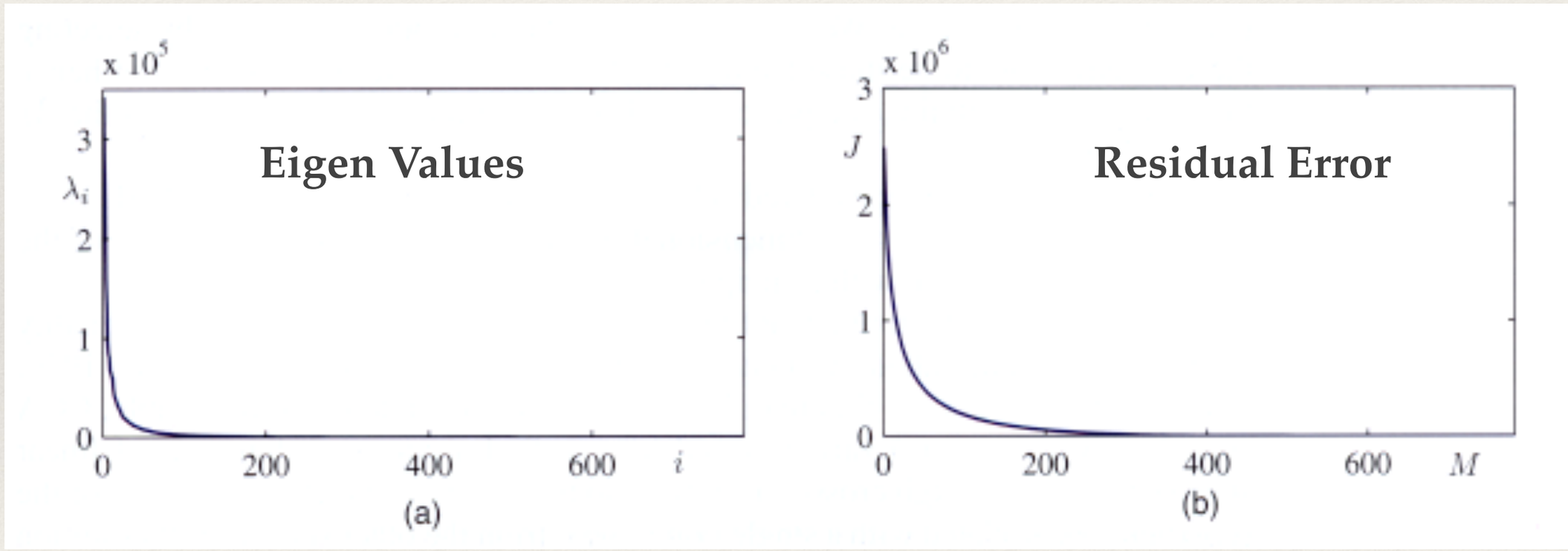
PCA

Handwritten digits used for PCA training…



PRML - C. Bishop (Sec. 12.1)

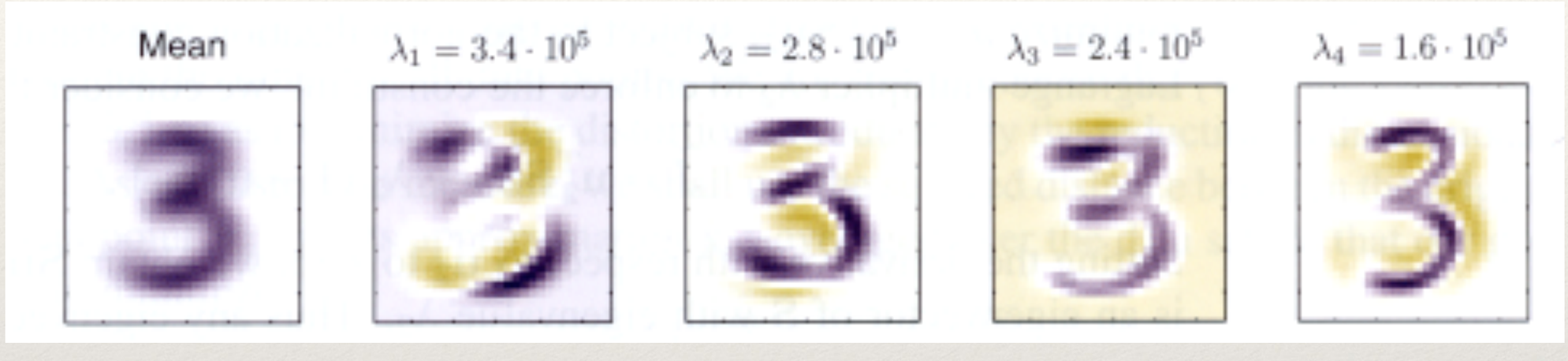
PCA



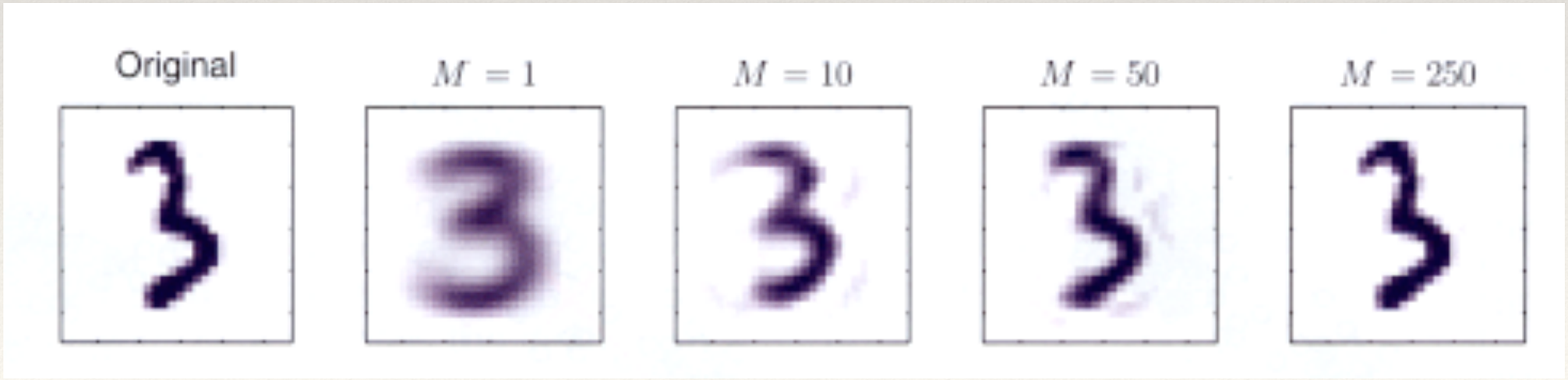
PRML - C. Bishop (Sec. 12.1)

PCA - Reconstruction

Eigenvectors



PCA - Reconstruction



PRML - C. Bishop (Sec. 12.1)