Deng Cai

- Ph.D 2008
- University of Illinois at Urbana Champaign
- Advisor: Jiawei Han
- Thesis topic:
  - Spectral Regression for Dimensionality Reduction
Graph Embedding & Extension for Dimensionality Reduction

Deng Cai

Department of Computer Science
University of Illinois at Urbana Champaign

Joint work with Xiaofei He, Jiawei Han
Dimensionality Reduction

• **Question**
  – How can we detect low dimensional structure in high dimensional data?

• **Applications**
  – Digital image and speech processing
  – Gene expression microarray data
  – Visualization of large networks
  – Analysis of neuronal populations
Given high dimensional data sampled from a low dimensional manifold, how to compute a faithful embedding?
Handwritten Digit
Pose and Expression
Hand Images
Problem

Input:

\[ x_i \in \mathbb{R}^n \quad \text{with} \quad i = 1, \ldots, m \]

Output:

\[ y_i \in \mathbb{R}^d \quad \text{with} \quad d \not\geq m \]

Embedding:

Nearby points remain nearby, distant points remain distant. 
Estimate \( d \).
Spectral Methods

• Isomap (2000)
• Locally Linear Embedding (2000)
• Laplacian Eigenmap (2001)
• Hessian LLE (2003)
• Maximum Variance Unfolding (2004)
• Conformal Eigenmaps (2005)
Spectral Methods

• **Common framework**
  – Derive sparse graph from $k$-NN
  – Derive matrix from graph weights: $W$
    • Varied algorithms differ in this step
  – Derive embedding from eigenvectors

$$y^{opt} = \arg \max \frac{y^T Wy}{y^T Dy} \implies Wy = \lambda Dy$$

• **Weakness**
  – No out-of-sample extension
Linear Extension

• Require a linear projective function

\[ y = f(x) = a^T x, \quad y = X^T a \]

\[
\mathbf{a}^{opt} = \arg \max \frac{\mathbf{a}^T XWX^T \mathbf{a}}{\mathbf{a}^T XDX^T \mathbf{a}} \quad \Rightarrow \quad XWX^T \mathbf{a} = \lambda XDX^T \mathbf{a}
\]

• Examples
  – Linear Discriminant Analysis
  – Locality Preserving Projection
  – Neighborhood Preserving Embedding
  – More …
Kernel Extension

- **Require a nonlinear projective function**

\[ y = f(x) = \sum \alpha_i K(x_i, x) \quad y = Ka \]

\[ \alpha^{opt} = \arg \max \frac{\alpha^T KWK\alpha}{\alpha^T KDK\alpha} \quad \Rightarrow \quad KWK\alpha = \lambda KDK\alpha \]

- **Examples**
  - Kernel Linear Discriminant Analysis
  - Kernel Locality Preserving Projection
  - More…
Platform for Dimensionality Reduction

• Develop new algorithms
  – New objective function
  – New graph structure
    • Locality Sensitive Discriminant Analysis
    • Maximum Margin Projection
    • Semi-supervised Discriminant Analysis
    • ....

• Efficient solution
  – Cubic-time complexity  Linear-time complexity
Conclusions

- Graph embedding is a way to transform continuous manifold learning problems into discrete ones
- Graph embedding provides a general framework for dimensionality reduction
- Need efficient optimization solutions
Thank You!

http://www.ews.uiuc.edu/~dengcai2