Graph-Regularized Generalized Low Rank Models

Mihir Paradkar & Dr. Madeleine Udell
Cornell University

Properties of Images

- High Dimensionality

Properties of Images

- High Dimensionality

- Noise and Occlusions

Noise and Occlusions

Properties of Images

- High Dimensionality

- Noise and Occlusions

- Graph Structure

Graph Structure

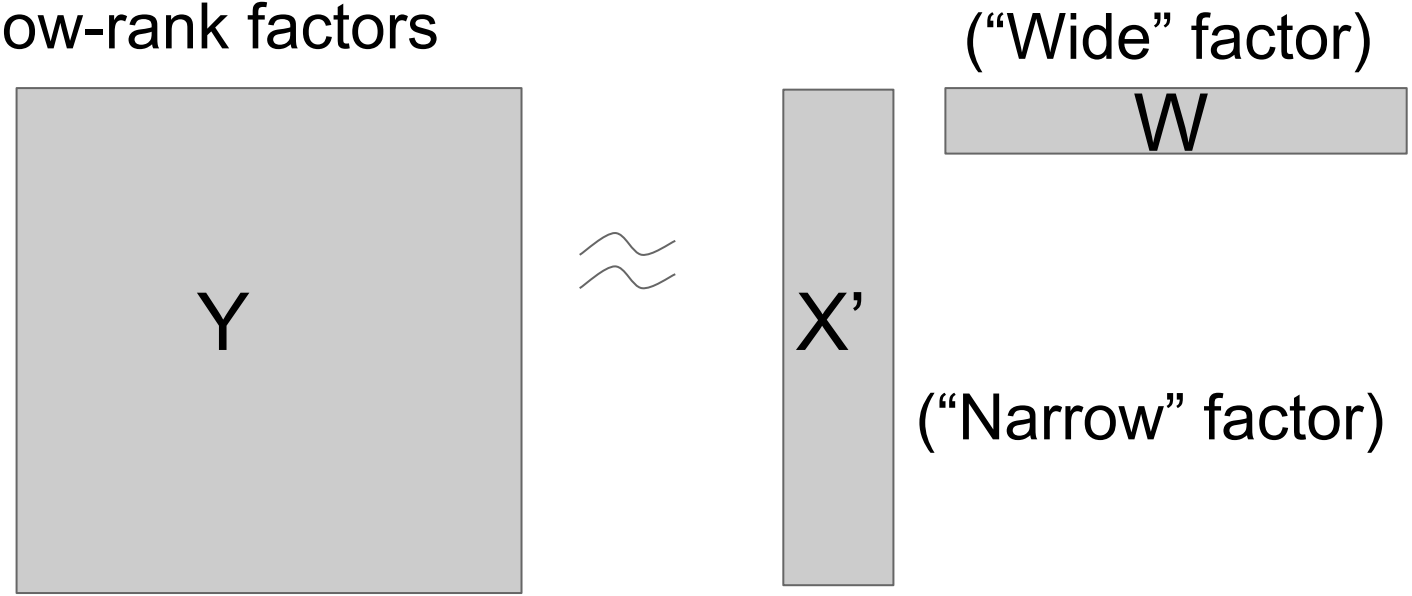
Previous Work

- Generalized Low Rank Models (GLRMs)

- Spectral Embedding

Low Rank Models

- Approximate a data matrix as the product of two low-rank factors



Low Rank Models

- Objective Function:

$$\min_{X,W} \sum_{(i,j)\in\Omega} l_j(Y_{ij}, x_i^T w_j) + r(X) + \tilde{r}(W)$$

Low Rank Models

- Using squared error recovers truncated SVD (PCA if centered and scaled)

$$\underset{X,Y}{\operatorname{argmin}} \|A - X^T Y\|_F^2$$

Spectral Embedding

- Maximizes similarity along a graph using Laplacian matrix L

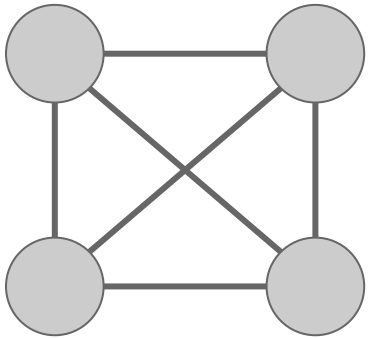
Spectral Embedding - Laplacian Matrix

(Path Graph)

1	-1		
-1	2	-1	
	-1	2	-1
		-1	1

Spectral Embedding - Laplacian Matrix

(Complete Graph)



3	-1	-1	-1
-1	3	-1	-1
-1	-1	3	-1
-1	-1	-1	3

Spectral Embedding

- Maximizes similarity along a graph using Laplacian matrix L

$$\min_{X} \ \operatorname{tr}(XLX^T) \ \ s.t. \ \ XX^T = I.$$

Graph GLRM

- Objective Function:

$$\min_{X,W} \sum_{(i,j)\in\Omega} l_j(Y_{ij}, x_i^T w_j) + \alpha \mathrm{tr}(XLX^T) + \tilde{r}(W)$$

Fitting

GraphGLRM uses modified Proximal Alternating Linearized Minimization (PALM) to fit factors; it takes two alternating proximal gradient steps per iteration, one per factor

Software Implementation

- Implementation in Julia language available at

`https://github.com/mihirparadkar/GraphGLRM.jl`

Software Implementation

- User specifies the data table, loss function, regularizers on factors, and rank, along with optional parameters like a list of indices of known values

```
gm = GGLRM(Amissing, loss, rx, ry, k, obs=obs)
```

Experiments

- Imputation of block occlusions

- Classification of faces into male/female with occluded images

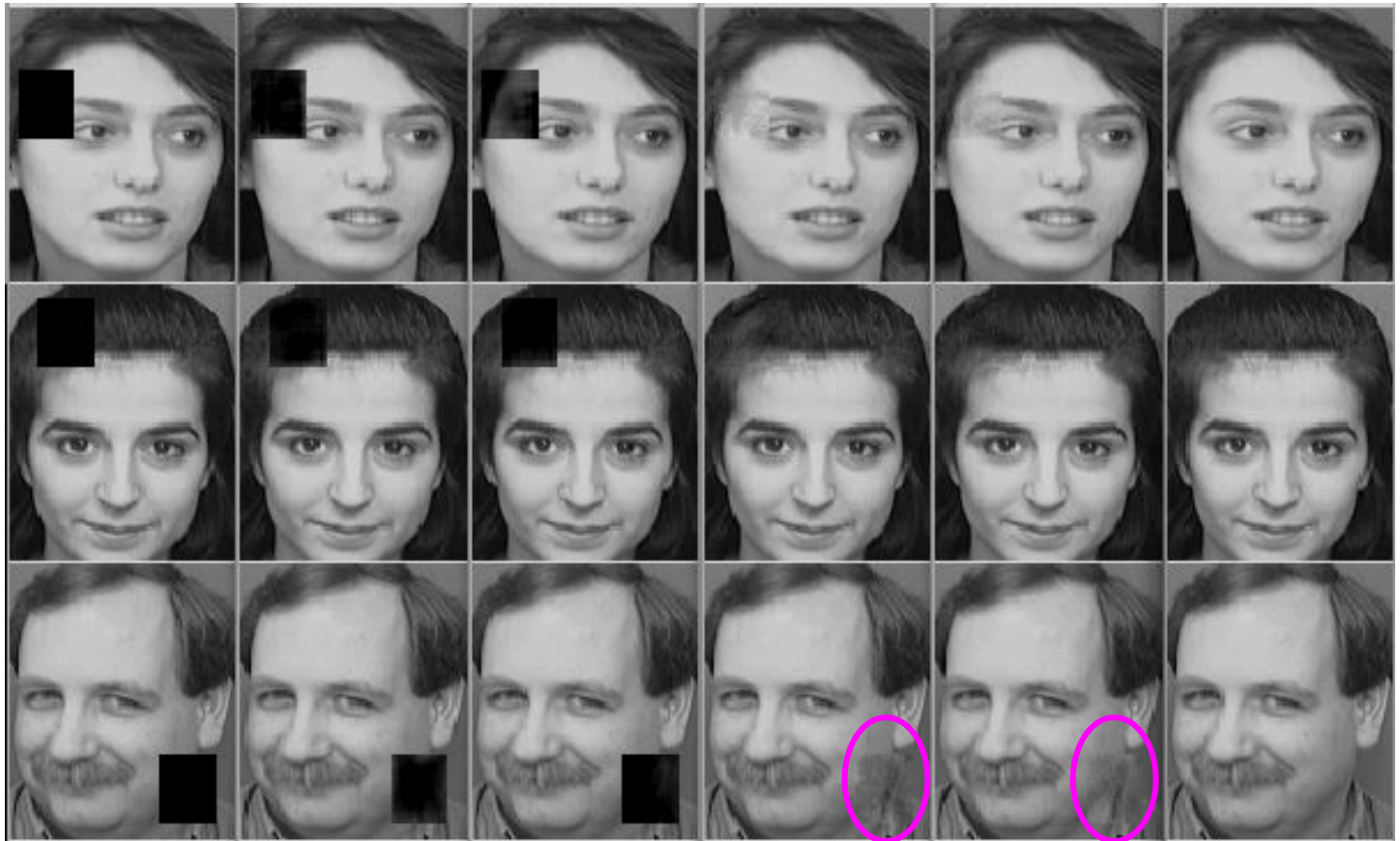
Results - Classification Experiment

Embedding Method	Precision	Recall	F1-Score
None	0.727	0.4	0.516
PCA	0.381	0.4	0.390
Spectral Embedding	1	0	0
Vanilla GLRM	0.714	0.25	0.370
Graph GLRM	**1**	**0.5**	**0.667**

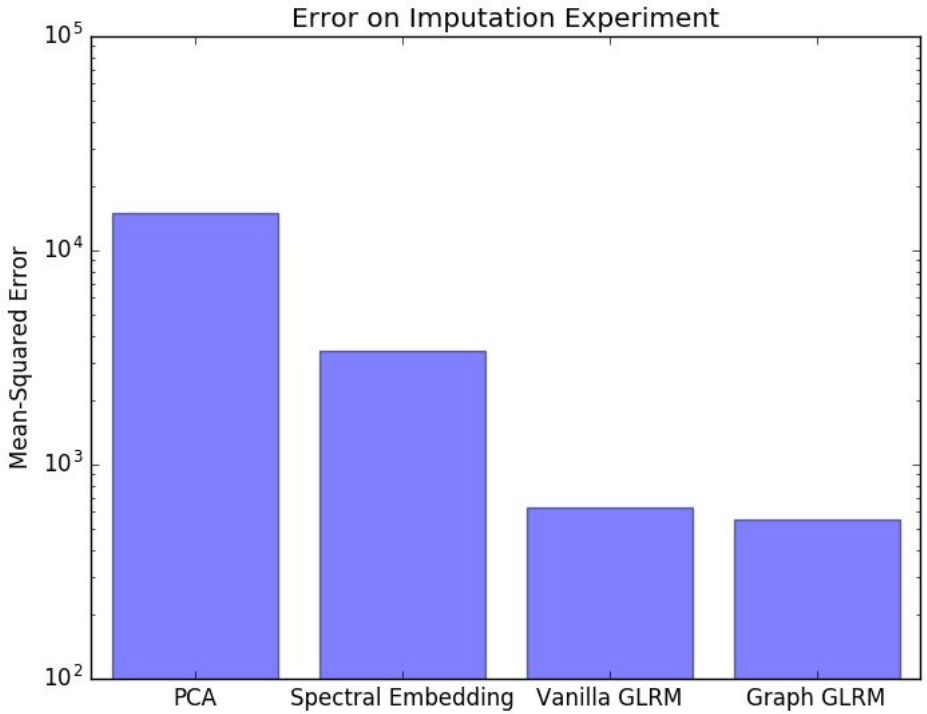
Results - Imputation Experiment

Method	MSE
PCA	15032
Spectral Embedding	3415.4
Vanilla GLRM	634.63
Graph GLRM	**554.48**

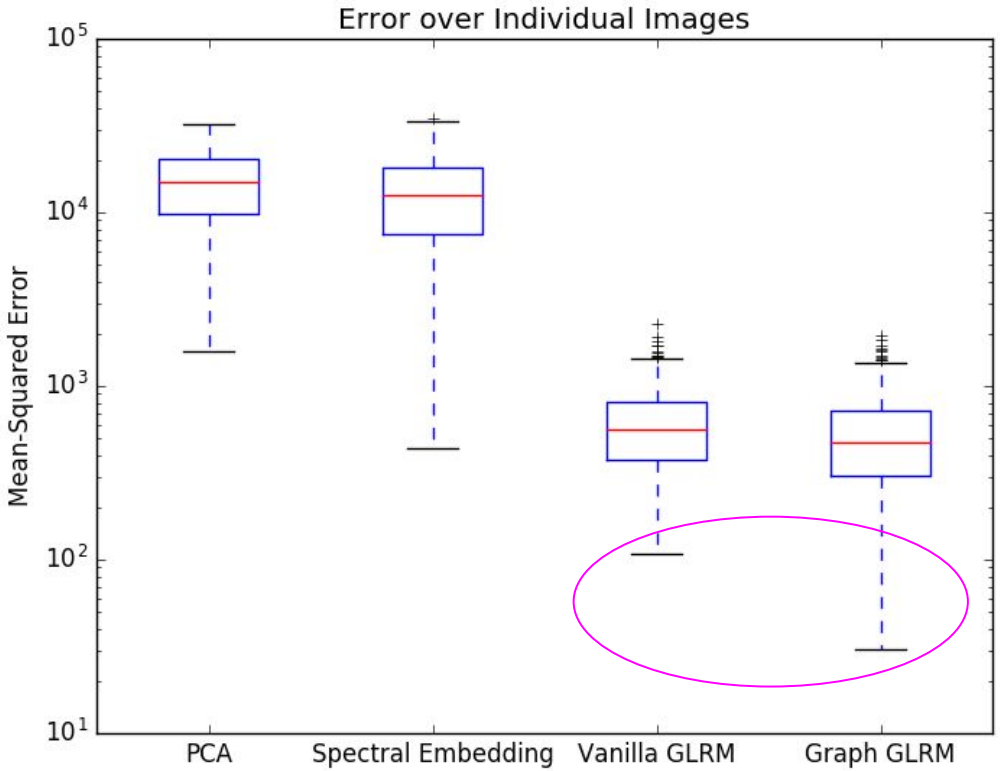
Results - Imputation Experiment



Results - Imputation Experiment



Results - Imputation Experiment



Conclusions

- Combines linear embedding of GLRM with non-linear dimensionality reduction through graph Laplacian

- Freely-available and performant software implementation for fitting

- Can improve performance in reconstructing missing values and in classification with noisy data

Thank You!