Low-Rank Matrix Completion Project

Authors: Artem Bochkarev, Roman Isachenko
Ilya Zharikov, Anne-Laure Ducrocq

Skolkovo Institute of Science and Technology
Numerical Linear Algebra course

December 16, 2016.

Problem Statement

Task

Given the amount of observed matrix entries to reconstruct low-rank matrix approximation.

Applications:

- recommender systems;
- image-processing;
- imputation of NAs for genomic data;
- rank estimation for SVD.

Problem Statement

Notations:

- M — $n \times m$ unknown matrix;
- $\Omega \in \{1, \dots, n\} \times \{1, \dots, m\}$ indices of observed elements;
-
$$P_\Omega(M) = \begin{cases} M_{ij}, & \text{if } (i,j) \in \Omega; \\ 0, & \text{otherwise.} \end{cases}$$

Optimization Task (*NP*-hard)

$$\underset{X \in \mathbb{R}^{n \times m}}{\text{minimize}} \quad \text{rank}(X)$$

$$\text{subject to} \quad P_\Omega(X) = P_\Omega(M)$$

Relaxations

Original Task (NP-hard)

$$\underset{X \in \mathbb{R}^{n \times m}}{\text{minimize}} \quad \text{rank}(X), \quad \text{subject to} \quad P_\Omega(X) = P_\Omega(M)$$

SVP

$$\underset{X \in \mathbb{R}^{n \times m}}{\text{minimize}} \quad \|P_\Omega(X) - P_\Omega(M)\|_F^2$$

$$\text{subject to} \quad \text{rank}(X) \leq k$$

SVT

$$\underset{X \in \mathbb{R}^{n \times m}}{\text{minimize}} \quad \tau\|X\|_* + \|X\|_F^2$$

$$\text{subject to} \quad P_\Omega(X) = P_\Omega(M)$$

RISMF

$$\underset{\substack{U \in \mathbb{R}^{n \times k} \\ V \in \mathbb{R}^{k \times m}}}{\text{minimize}} \quad \|U\|_F^2 + \|V\|_F^2$$

$$\text{subject to} \quad P_\Omega(UV) = P_\Omega(M)$$

SoftImpute

$$\underset{X \in \mathbb{R}^{n \times m}}{\text{minimize}} \quad \|X\|_*$$

$$\text{subject to} \quad \|P_\Omega(X) - P_\Omega(M)\|_F \leq \delta$$

Related Works

1. Candes E. J., Recht B. Exact matrix completion via convex optimization. 2009.
2. Cai J. F., Candes E. J., Shen Z. A singular value thresholding algorithm for matrix completion. 2010.
3. Mazumder R., Hastie T., Tibshirani R. Spectral regularization algorithms for learning large incomplete matrices. 2010.
4. Jain P., Meka R., Dhillon I. S. Guaranteed rank minimization via singular value projection. 2010.
5. Takacs G. et al. Scalable collaborative filtering approaches for large recommender systems. 2009.
6. Vandereycken B. Low-rank matrix completion by Riemannian optimization. 2013.

SVP (Jain P., Meka R., Dhillon I. S., 2010)

Relaxation

$$\underset{X \in \mathbb{R}^{n \times m}}{\text{minimize}} \quad \|P_\Omega(X) - P_\Omega(M)\|_F^2, \quad \text{subject to} \quad \text{rank}(X) \le k$$

Input: $\Omega, P_\Omega(M), k, \eta$;
Output: X;
1: $X := 0$;
2: **repeat**
3: $\quad X := X - \eta P_\Omega^{-1}\left(P_\Omega(X) - P_\Omega(M)\right)$;
4: $\quad U, \Sigma, V := SVD(X)$;
5: $\quad X := U_k \Sigma_k V_k^T$;
6: **until** $\|P_\Omega(M) - P_\Omega(X)\|_F / \|P_\Omega(X)\|_F > \varepsilon$

SVT (Cai J. F., Candes E. J., Shen Z., 2010)

Relaxation

$$\underset{X \in \mathbb{R}^{n \times m}}{\text{minimize}} \quad \lambda \|X\|_* + \|X\|_F^2, \quad \text{subject to} \quad P_\Omega(X) = P_\Omega(M)$$

Input: $\Omega, M, \varepsilon, \lambda, \eta$;
Output: X;
1: $Y := \eta P_\Omega(M)$;
2: **repeat**
3: $\quad X := S_\lambda(Y)$;
4: $\quad P_\Omega(Y) := P_\Omega(Y) + \eta P_\Omega(M - X)$;
5: **until** $\|P_\Omega(X - M)\|_F / \|P_\Omega(M)\|_F > \varepsilon$

$$S_\lambda(W) = U[D - \lambda I]_+ V^*; \quad U, D, V := SVD(W)$$

SoftImpute (Hastie T., Tibshirani R., 2010)

Lagrangian Relaxation

$$\underset{X \in \mathbb{R}^{n \times m}}{\text{minimize}} \quad \lambda \|X\|_* + \|P_\Omega(X) - P_\Omega(M)\|_F^2$$

Input: $M, \Omega, \varepsilon, \lambda_1 > \cdots > \lambda_k$;
Output: sequence $X_{\lambda_1}, \ldots, X_{\lambda_k}$;
 1: $X := 0$;
 2: **for** $i = 1, \ldots, k$ **do**
 3:    **repeat**
 4:       $X := S_{\lambda_i}(P_\Omega(M) + P_\Omega^\perp(X))$;
 5:    **until** $\|P_\Omega(M) - P_\Omega(X)\|_F / \|P_\Omega(M)\|_F > \varepsilon$
 6:   $X_{\lambda_i} := X$;

$$P_\Omega^\perp(M) = M - P_\Omega(M).$$

RIMSF

Relaxation

$$\underset{\substack{U \in \mathbb{R}^{n \times k} \\ V \in \mathbb{R}^{k \times m}}}{\text{minimize}} \quad \|P_\Omega(UV) - P_\Omega(M)\|_F^2 + \lambda \left(\|U\|_F^2 + \|V\|_F^2 \right)$$

Input: $M, \Omega, \varepsilon, \lambda, \eta$;
Output: U, V;
1: Initialization of U, V with small random numbers;
2: **repeat**
3:     **for** $(i, j) \in \Omega$ **do**
4:        Gradient step for $U[i, :]$ and $V[:, j]$ by regularized error function;
5: **until** $\|P_\Omega(M) - P_\Omega(UV)\|_F / \|P_\Omega(M)\|_F > \varepsilon$

Numerical Experiments

- Synthetic data
  Low-rank matrices with random noise.

- Low-rank images
  Visual demonstration of algorithms.

- Assessment dataset
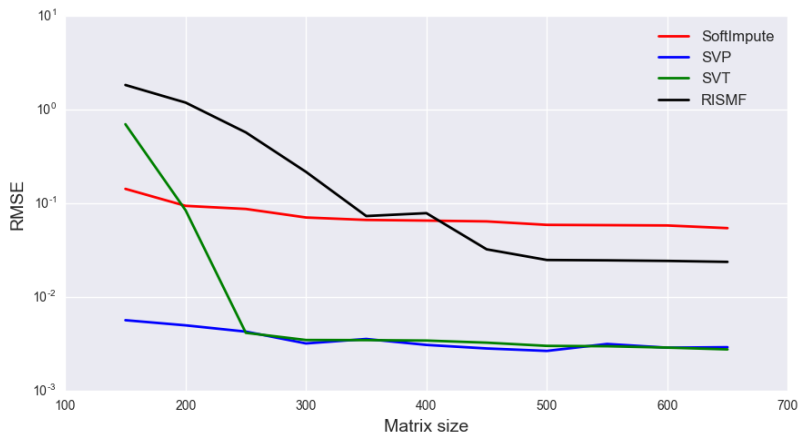 Real dataset for collaborating filtering.
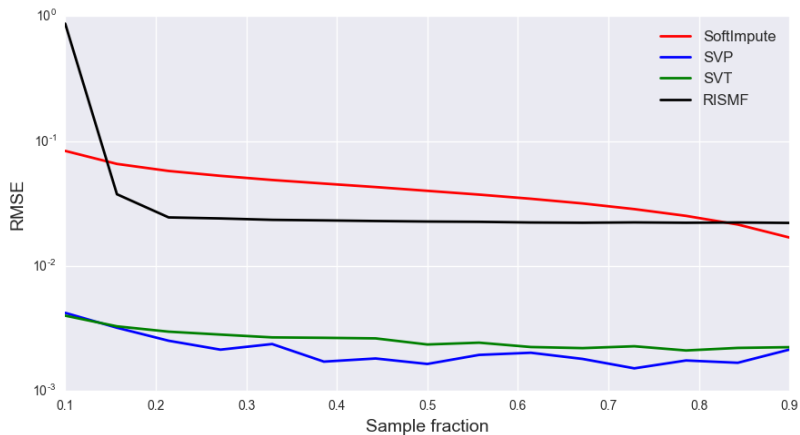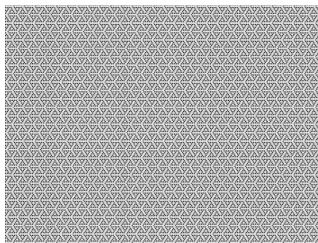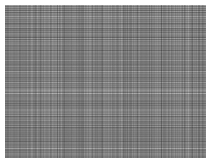
Synthetic Data

Synthetic Data
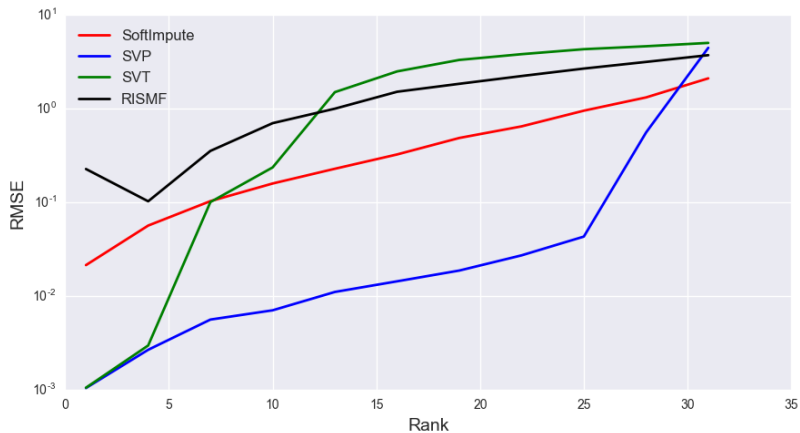
Images

Sampled Image
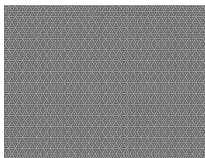

SoftImpute

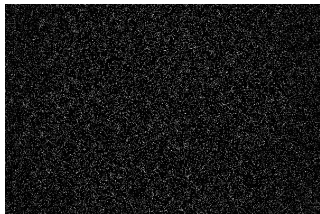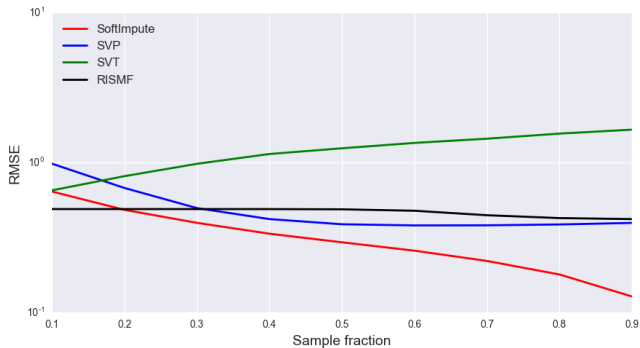### RISMF

SVT

SVP

Images

Sampled Image

SVT

RISMF

SVT

SVP

Assessment data

Assessment data