

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.

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)

$$\begin{array}{ll} \underset{X \in \mathbb{R}^{n \times m}}{\text{minimize}} & \text{rank}(X) \\ \text{subject to} & P_{\Omega}(X) = P_{\Omega}(M) \end{array}$$

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

$$\underset{X \in \mathbb{R}^{n \times m}}{\text{minimize}} \quad \|X\|_*$$
$$\text{subject to} \quad \|P_{\Omega}(X) - P_{\Omega}(M)\|_F \leq \delta$$

- 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.

Relaxation

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

Input: $\Omega, P_{\Omega}(M), k, \eta$;

Output: X ;

1: $X := 0$;

2: **repeat**

3: $X := X - \eta P_{\Omega}^{-1}(P_{\Omega}(X) - P_{\Omega}(M))$;

4: $U, \Sigma, V := \text{SVD}(X)$;

5: $X := U_k \Sigma_k V_k^T$;

6: **until** $\|P_{\Omega}(M) - P_{\Omega}(X)\|_F / \|P_{\Omega}(X)\|_F > \varepsilon$

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: $X := S_\lambda(Y)$;

4: $P_\Omega(Y) := P_\Omega(Y) + \eta P_\Omega(M - X)$;

5: **until** $\|P_\Omega(X - M)\|_F / \|P_\Omega(M)\|_F > \varepsilon$

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

Lagrangian Relaxation

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

Input: $M, \Omega, \varepsilon, \lambda_1 > \dots > \lambda_k$;

Output: sequence $X_{\lambda_1}, \dots, X_{\lambda_k}$;

1: $X := 0$;

2: **for** $i = 1, \dots, k$ **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: $X_{\lambda_i} := X$;

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

Relaxation

$$\begin{array}{l} \text{minimize} \\ U \in \mathbb{R}^{n \times k} \\ V \in \mathbb{R}^{k \times m} \end{array} \quad \|P_{\Omega}(UV) - P_{\Omega}(M)\|_F^2 + \lambda (\|U\|_F^2 + \|V\|_F^2)$$

Input: $M, \Omega, \varepsilon, \lambda, \eta$;

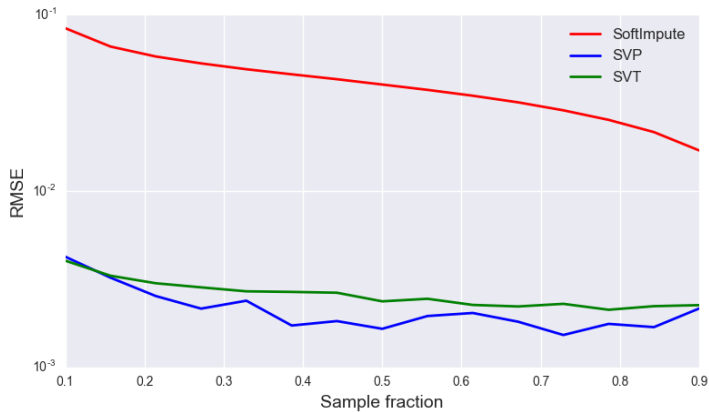
Output: U, V ;

- 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** $\|P_{\Omega}(M) - P_{\Omega}(UV)\|_F / \|P_{\Omega}(M)\|_F > \varepsilon$

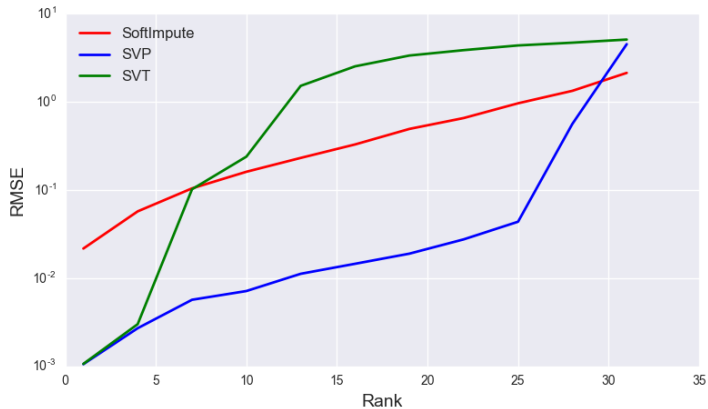
Numerical Experiments

- Synthetic data
Low-rank matrices with random noise.
- Low-rank images
Visual demonstration of algorithms.
- Assessment dataset
Real dataset for collaborating filtering.

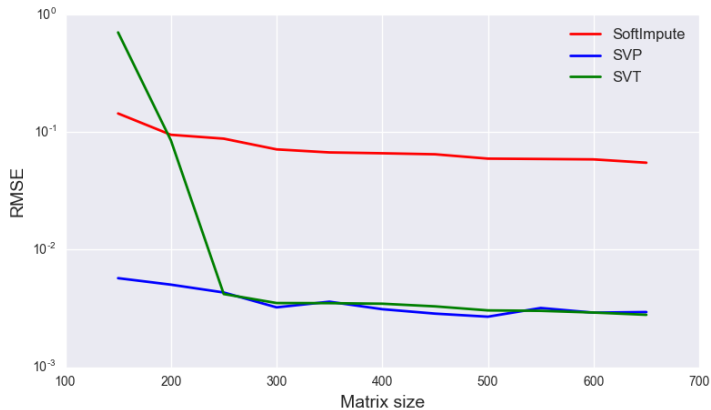
Synthetic Data



Synthetic Data



Synthetic Data



Assessment data