

# Low-Rank Matrix Completion project

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

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

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

## 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.

# Convex Relaxation (Candes E. J., Recht B., 2009)

## Original Task

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

## Relaxation

$$\begin{aligned} & \underset{X \in \mathbb{R}^{n \times m}}{\text{minimize}} && \|X\|_* \\ & \text{subject to} && P_{\Omega}(X) = P_{\Omega}(M). \end{aligned}$$

## Motivation:

$$\text{rank}(X) = |\{i : \sigma_i(X) \neq 0\}|; \quad \|X\|_* = \sum_{i=1}^k \sigma_i(X).$$





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

## Original Task

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

## Relaxation

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

## Motivation:

The method is the same as *SVT* with  $\delta = 0$ . When  $\delta > 0$  the overfitting is less possible.





# Riemannian Optimization