# Graph-regularized matrix completion

Pierre-Antoine Absil (UCLouvain)

Based on joint work with:

Nicolas Boumal (EPFL)
Shuyu Dong (INRIA Saclay)
Kyle Gallivan (Florida State University)
Benoît Loucheur (UCLouvain)
Michel Journée (Royal Meteorological Institute of Belgium)

GraphLearn: Machine Learning and Signal Processing on Graphs
CIRM, France
Nov. 7–11, 2022

# Optimization on manifolds ⇝ Graphs?

- T. P. Cason, PA, Paul Van Dooren, *Iterative methods for low rank approximation of graph similarity matrices*, Linear Algebra and its Applications, 438, pp. 1863-1882, 2013.
- Shuyu Dong, Dorina Thanou, PA, Pascal Frossard, *Learning sparse models of diffusive graph signals*, ESANN 2017.
- Shuyu Dong, PA, K. A. Gallivan, *Riemannian gradient descent methods for graph-regularized matrix completion*, Linear Algebra and its Applications, 623, pp. 193-235, 2021.

# The Netflix Prize (https://www.netflixprize.com/leaderboard.html)

## Leaderboard

**Showing Test Score.** Click here to show quiz score

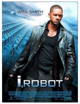
**Display top** 20 **leaders.**

| Rank | Team Name | Best Test Score | % Improvement | Best Submit Time |
|---|---|---|---|---|
| **Grand Prize - RMSE = 0.8567 - Winning Team: BellKor's Pragmatic Chaos** | | | | |
| 1 | BellKor's Pragmatic Chaos | 0.8567 | 10.06 | 2009-07-26 18:18:28 |
| 2 | The Ensemble | 0.8567 | 10.06 | 2009-07-26 18:38:22 |
| 3 | Grand Prize Team | 0.8582 | 9.90 | 2009-07-10 21:24:40 |
| 4 | Opera Solutions and Vandelay United | 0.8588 | 9.84 | 2009-07-10 01:12:31 |
| 5 | Vandelay Industries ! | 0.8591 | 9.81 | 2009-07-10 00:32:20 |
| 6 | PragmaticTheory | 0.8594 | 9.77 | 2009-06-24 12:06:56 |
| 7 | BellKor in BigChaos | 0.8601 | 9.70 | 2009-05-13 08:14:09 |
| 8 | Dace | 0.8612 | 9.59 | 2009-07-24 17:18:43 |
| 9 | Feeds2 | 0.8622 | 9.48 | 2009-07-12 13:11:51 |
| 10 | BigChaos | 0.8623 | 9.47 | 2009-04-07 12:33:59 |
| 11 | Opera Solutions | 0.8623 | 9.47 | 2009-07-24 00:34:07 |
| 12 | BellKor | 0.8624 | 9.46 | 2009-07-26 17:19:11 |
| **Progress Prize 2008 - RMSE = 0.8627 - Winning Team: BellKor in BigChaos** | | | | |
| 13 | xiangliang | 0.8642 | 9.27 | 2009-07-15 14:53:22 |
| 14 | Gravity | 0.8643 | 9.26 | 2009-04-22 18:31:32 |

# The Netflix database: ratings of movies by users

## Our convention: one row per movie, one column per user

$$X = \begin{array}{cc} \begin{array}{cccccccccccccccc} &&&&&&&&&& \text{user } u &&&&& \end{array} & \\ \left(\begin{array}{cccccccccccccccc} 5 & ? & 2 & ? & ? & 5 & ? & ? & 3 & ? & 5 & ? & ? & ? & 2 & ? \\ 3 & ? & 2 & ? & ? & 2 & ? & ? & ? & 3 & 2 & ? & 5 & ? & ? & ? \\ 1 & ? & 5 & 2 & 3 & 4 & ? & 4 & ? & ? & ? & 2 & ? & ? & ? & ? \\ ? & 1 & ? & 3 & ? & ? & ? & 3 & ? & ? & ? & ? & 2 & 1 & 5 & 5 \\ ? & 4 & ? & ? & ? & ? & 5 & ? & ? & ? & 1 & ? & ? & 1 & ? & 4 \end{array}\right) & \begin{array}{l} \\ \text{movie } i \\ \\ \\ \\ \end{array} \end{array}$$

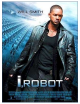


Netflix dataset: 17,770 movies and 480,189 users
Thus about 8.5 billion movie-user pairs
About 1% of the pairs are rated.

# Most ratings are unknown

## Our job is to guess what they could be

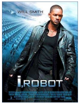
$$\begin{array}{cc} & \begin{array}{cccccccccccccccc} &&&&&&&&&& \text{user } u &&&&& \end{array} \\ X = & \begin{pmatrix} 5 & ? & 2 & ? & ? & 5 & ? & ? & 3 & ? & 5 & ? & ? & ? & 2 & ? \\ 3 & ? & 2 & ? & ? & 2 & ? & ? & ? & 3 & 2 & ? & 5 & ? & ? & ? \\ 1 & ? & 5 & 2 & 3 & 4 & ? & 4 & ? & ? & ? & 2 & ? & ? & ? & ? \\ ? & 1 & ? & 3 & ? & ? & ? & 3 & ? & ? & ? & ? & 2 & 1 & 5 & 5 \\ ? & 4 & ? & ? & ? & ? & 5 & ? & ? & ? & 1 & ? & ? & 1 & ? & 4 \end{pmatrix} \begin{array}{c} \\ \text{movie } i \\ \\ \\ \\ \end{array} \end{array}$$



# 1st approach to matrix completion: neighborhood methods

## Exploit similarities between users

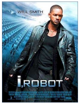
$$X = \begin{array}{cc} \begin{array}{cccccccccccccccc} &&&&&&&&&&\text{user } u&&&&& \end{array} & \\ \begin{pmatrix} 5 & ? & 2 & ? & ? & 5 & ? & ? & 3 & ? & 5 & ? & ? & ? & 2 & ? \\ 3 & ? & 2 & ? & ? & 2 & ? & ? & ? & 3 & 2 & ? & 5 & ? & ? & ? \\ 1 & ? & 5 & 2 & 3 & 4 & ? & 4 & ? & ? & 4 & 2 & ? & ? & ? & ? \\ ? & 1 & ? & 3 & ? & ? & ? & 3 & ? & ? & 1 & ? & 2 & 1 & 5 & 5 \\ ? & 4 & ? & ? & ? & ? & 5 & ? & ? & ? & 1 & ? & ? & 1 & ? & 4 \end{pmatrix} & \begin{array}{c} \\ \text{movie } i \\ \\ \\ \\ \end{array} \end{array}$$



# 2nd approach to matrix completion: low-rank methods

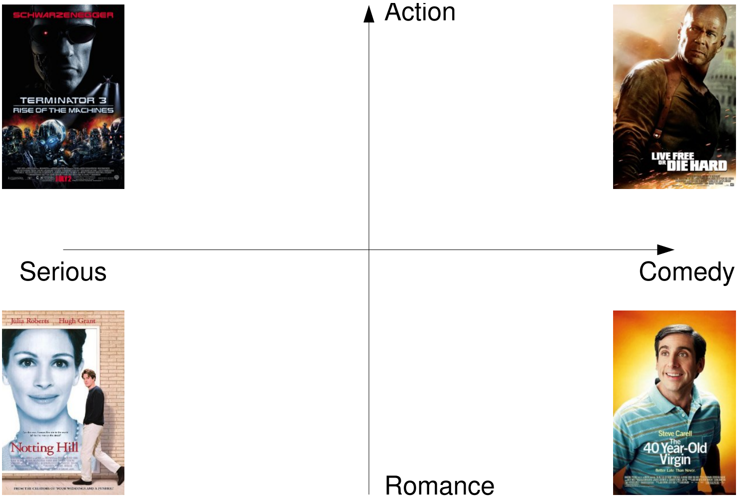
Assume that $X$ is close to matrix of low-rank $r$

$$\begin{array}{rcl}
 & \begin{array}{cccccccccccccccc} &&&&&&&&&& \text{user } u &&&&& \end{array} & \\
X = & \left(\begin{array}{cccccccccccccccc}
5 & ? & 2 & ? & ? & 5 & ? & ? & 3 & ? & 5 & ? & ? & ? & 2 & ? \\
3 & ? & 2 & ? & ? & 2 & ? & ? & ? & 3 & 2 & ? & 5 & ? & ? & ? \\
1 & ? & 5 & 2 & 3 & 4 & ? & 4 & ? & ? & ? & 2 & ? & ? & ? & ? \\
? & 1 & ? & 3 & ? & ? & ? & 3 & ? & ? & ? & ? & 2 & 1 & 5 & 5 \\
? & 4 & ? & ? & ? & ? & 5 & ? & ? & ? & 1 & ? & ? & 1 & ? & 4
\end{array}\right) & \begin{array}{l} \\ \text{movie } i \\ \\ \\ \\ \end{array}
\end{array}$$



$$\approx \begin{pmatrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{pmatrix} * \begin{pmatrix} ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \end{pmatrix}$$

# A priori justification for the low-rank model



Action

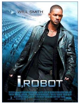
Serious

Comedy

Romance

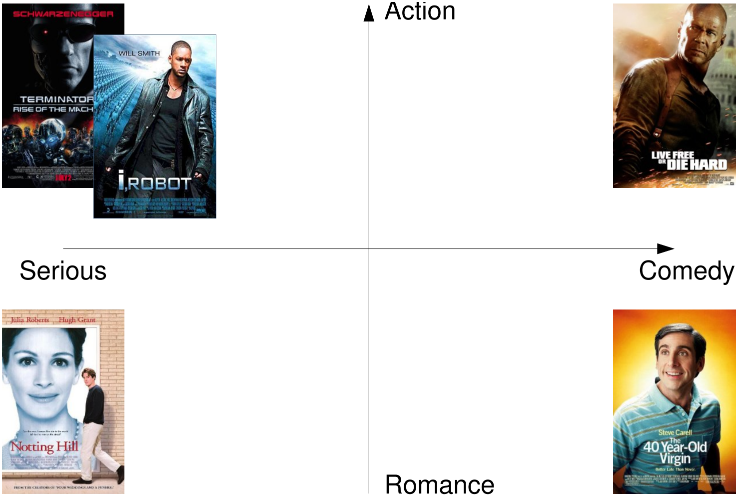
# A priori justification for the low-rank model

$$X = \begin{array}{cc} \begin{matrix} & & & & & & & & & & \text{user } u & & & & & \end{matrix} & \\ \begin{pmatrix} 5 & ? & 2 & ? & ? & 5 & ? & ? & 3 & ? & 5 & ? & ? & ? & 2 & ? \\ 3 & ? & 2 & ? & ? & 2 & ? & ? & ? & 3 & 2 & ? & 5 & ? & ? & ? \\ 1 & ? & 5 & 2 & 3 & 4 & ? & 4 & ? & ? & ? & 2 & ? & ? & ? & ? \\ ? & 1 & ? & 3 & ? & ? & ? & 3 & ? & ? & ? & ? & 2 & 1 & 5 & 5 \\ ? & 4 & ? & ? & ? & ? & 5 & ? & ? & ? & 1 & ? & ? & 1 & ? & 4 \end{pmatrix} & \begin{matrix} \\ \text{movie } i \\ \\ \\ \\ \end{matrix} \end{array}$$



$$\approx \begin{pmatrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{pmatrix} * \begin{pmatrix} ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \end{pmatrix}$$

# A priori justification for the low-rank model



Action

Serious

Comedy

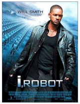
Romance

# A priori justification for the low-rank model

user $u$

$$X = \begin{pmatrix} 5 & ? & 2 & ? & ? & 5 & ? & ? & 3 & ? & 5 & ? & ? & ? & 2 & ? \\ 3 & ? & 2 & ? & ? & 2 & ? & ? & ? & 3 & 2 & ? & 5 & ? & ? & ? \\ 1 & ? & 5 & 2 & 3 & 4 & ? & 4 & ? & ? & ? & 2 & ? & ? & ? & ? \\ ? & 1 & ? & 3 & ? & ? & ? & 3 & ? & ? & ? & ? & 2 & 1 & 5 & 5 \\ ? & 4 & ? & ? & ? & ? & 5 & ? & ? & ? & 1 & ? & ? & 1 & ? & 4 \end{pmatrix}$$

movie $i$



$$\approx \begin{pmatrix} ? & ? \\ - & + \\ ? & ? \\ ? & ? \\ ? & ? \end{pmatrix} * \begin{pmatrix} ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \end{pmatrix}$$

# A priori justification for the low-rank model

Action

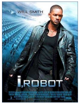
Serious

Comedy

Romance

# A priori justification for the low-rank model

$$X = \begin{array}{c}\\ \left(\begin{array}{cccccccccccccccc} & & & & & & & & & & \text{user } u & & & & & \\ 5 & ? & 2 & ? & ? & 5 & ? & ? & 3 & ? & 5 & ? & ? & ? & 2 & ? \\ 3 & ? & 2 & ? & ? & 2 & ? & ? & ? & 3 & 2 & ? & 5 & ? & ? & ? \\ 1 & ? & 5 & 2 & 3 & 4 & ? & 4 & ? & ? & ? & 2 & ? & ? & ? & ? \\ ? & 1 & ? & 3 & ? & ? & ? & 3 & ? & ? & ? & ? & 2 & 1 & 5 & 5 \\ ? & 4 & ? & ? & ? & ? & 5 & ? & ? & ? & 1 & ? & ? & 1 & ? & 4 \end{array}\right)\end{array} \begin{array}{l} \\ \\ \text{movie } i \\ \\ \\ \\ \end{array}$$

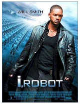


$$\approx \begin{pmatrix} ? & ? \\ - & + \\ ? & ? \\ ? & ? \\ ? & ? \end{pmatrix} * \begin{pmatrix} ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & + & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & 0 & ? & ? & ? & ? & ? \end{pmatrix}$$

$+$ mean

# 2nd approach to matrix completion: low-rank methods

Task: find $A$ and $B$



$$X = \begin{array}{cc} \begin{array}{cccccccccccccccc} &&&&&&&&&&\text{user } u&&&&& \end{array} & \\ \begin{pmatrix} 5 & ? & 2 & ? & ? & 5 & ? & ? & 3 & ? & 5 & ? & ? & ? & 2 & ? \\ 3 & ? & 2 & ? & ? & 2 & ? & ? & ? & 3 & 2 & ? & 5 & ? & ? & ? \\ 1 & ? & 5 & 2 & 3 & 4 & ? & 4 & ? & ? & ? & 2 & ? & ? & ? & ? \\ ? & 1 & ? & 3 & ? & ? & ? & 3 & ? & ? & ? & ? & 2 & 1 & 5 & 5 \\ ? & 4 & ? & ? & ? & ? & 5 & ? & ? & ? & 1 & ? & ? & 1 & ? & 4 \end{pmatrix} & \begin{array}{c} \\ \text{movie } i \\ \\ \\ \\ \end{array} \end{array}$$

$$\approx \underset{A}{\begin{pmatrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{pmatrix}} * \underset{B^T}{\begin{pmatrix} ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \end{pmatrix}}$$

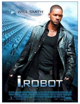
# Low-rank matrix completion (LRMC)

## Mean square error formulation

user $u$

$$X = \begin{pmatrix} 5 & ? & 2 & ? & ? & 5 & ? & ? & 3 & ? & 5 & ? & ? & ? & 2 & ? \\ 3 & ? & 2 & ? & ? & 2 & ? & ? & ? & 3 & 2 & ? & 5 & ? & ? & ? \\ 1 & ? & 5 & 2 & 3 & 4 & ? & 4 & ? & ? & ? & 2 & ? & ? & ? & ? \\ ? & 1 & ? & 3 & ? & ? & ? & 3 & ? & ? & ? & ? & 2 & 1 & 5 & 5 \\ ? & 4 & ? & ? & ? & ? & 5 & ? & ? & ? & 1 & ? & ? & 1 & ? & 4 \end{pmatrix}$$

movie $i$



$$\approx \underbrace{\begin{pmatrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{pmatrix}}_{A} * \underbrace{\begin{pmatrix} ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \end{pmatrix}}_{B^T}$$

Mean square formulation:

$$\min_{\hat{X}\in\mathbb{R}^{m\times n}} \sum_{\underset{\text{known entries}}{(i,u)\in\Omega}} \left(\hat{X}_{iu} - \overset{\text{known ratings}}{X_{iu}}\right)^2, \quad \text{subject to } \operatorname{rank}(\hat{X}) \le r.$$

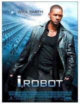
# Can we combine the neighborhood and low-rank approaches?

# Can we combine the neighborhood and low-rank approaches?

Ref: Vassilis Kalofolias, Xavier Bresson, Michael Bronstein, Pierre Vandergheynst, *Matrix completion on graphs*, arXiv:1408.1717v3

# A graph-based objective function

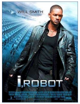
$$X = \begin{array}{ccccccccccccccccc} & & & & & & & & & & \text{user } u & & & & & & \\ 5 & ? & 2 & ? & ? & 5 & ? & ? & 3 & ? & 5 & ? & ? & ? & 2 & ? & \\ 3 & ? & 2 & ? & ? & 2 & ? & ? & ? & 3 & 2 & ? & 5 & ? & ? & ? & \text{movie } i \\ 1 & ? & 5 & 2 & 3 & 4 & ? & 4 & ? & ? & ? & 2 & ? & ? & ? & ? & \\ ? & 1 & ? & 3 & ? & ? & ? & 3 & ? & ? & ? & ? & 2 & 1 & 5 & 5 & \\ ? & 4 & ? & ? & ? & ? & 5 & ? & ? & ? & 1 & ? & ? & 1 & ? & 4 & \end{array}$$



$$\approx \underset{A}{\begin{pmatrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{pmatrix}} * \underset{B^T}{\begin{pmatrix} ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \end{pmatrix}}$$

# A graph-based objective function

$$X = \begin{array}{c c} \begin{array}{cccccccccccccccc} & & & & & & & & & & \text{user } u & & & & & \end{array} & \\ \begin{pmatrix} 5 & ? & 2 & ? & ? & 5 & ? & ? & 3 & ? & 5 & ? & ? & ? & 2 & ? \\ 3 & ? & 2 & ? & ? & 2 & ? & ? & ? & 3 & 2 & ? & 5 & ? & ? & ? \\ 1 & ? & 5 & 2 & 3 & 4 & ? & 4 & ? & ? & ? & 2 & ? & ? & ? & ? \\ ? & 1 & ? & 3 & ? & ? & ? & 3 & ? & ? & ? & ? & 2 & 1 & 5 & 5 \\ ? & 4 & ? & ? & ? & ? & 5 & ? & ? & ? & 1 & ? & ? & 1 & ? & 4 \end{pmatrix} & \begin{array}{c} \\ \text{movie } i \\ \\ \\ \\ \end{array} \end{array}$$



$$\approx \underbrace{\begin{pmatrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{pmatrix}}_{A} * \underbrace{\begin{pmatrix} ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \end{pmatrix}}_{B^T}$$

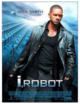
From the incomplete matrix, or from metadata on the users, build a graph with weight matrix $W$, where $W_{uv}$ indicates the "similarity" between users $u$ and $v$.

# A graph-based objective function

user $u$

$$X = \begin{pmatrix} 5 & ? & 2 & ? & ? & 5 & ? & ? & 3 & ? & 5 & ? & ? & ? & 2 & ? \\ 3 & ? & 2 & ? & ? & 2 & ? & ? & ? & 3 & 2 & ? & 5 & ? & ? & ? \\ 1 & ? & 5 & 2 & 3 & 4 & ? & 4 & ? & ? & ? & 2 & ? & ? & ? & ? \\ ? & 1 & ? & 3 & ? & ? & ? & 3 & ? & ? & ? & ? & 2 & 1 & 5 & 5 \\ ? & 4 & ? & ? & ? & ? & 5 & ? & ? & ? & 1 & ? & ? & 1 & ? & 4 \end{pmatrix}$$

movie $i$



$$\approx \underset{A}{\begin{pmatrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{pmatrix}} * \underset{B^T}{\begin{pmatrix} ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \end{pmatrix}}$$

From the incomplete matrix, or from metadata on the users, build a graph with weight matrix $W$, where $W_{uv}$ indicates the "similarity" between users $u$ and $v$.

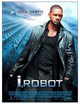
$\rightsquigarrow$ A graph-based objective function: $\displaystyle\sum_{u,v=1}^{m} W_{uv} \|\hat{X}_{:,u} - \hat{X}_{:,v}\|^2$

# Combining the low-rank and neighborhood approaches

user $u$

$$X = \begin{pmatrix} 5 & ? & 2 & ? & ? & 5 & ? & ? & 3 & ? & 5 & ? & ? & ? & 2 & ? \\ 3 & ? & 2 & ? & ? & 2 & ? & ? & ? & 3 & 2 & ? & 5 & ? & ? & ? \\ 1 & ? & 5 & 2 & 3 & 4 & ? & 4 & ? & ? & ? & 2 & ? & ? & ? & ? \\ ? & 1 & ? & 3 & ? & ? & ? & 3 & ? & ? & ? & ? & 2 & 1 & 5 & 5 \\ ? & 4 & ? & ? & ? & ? & 5 & ? & ? & ? & 1 & ? & ? & 1 & ? & 4 \end{pmatrix}$$

movie $i$



$$\approx \underbrace{\begin{pmatrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{pmatrix}}_{A} * \underbrace{\begin{pmatrix} ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \end{pmatrix}}_{B^T}$$

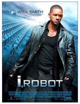
known ratings

$$\min_{\hat{X} \in \mathbb{R}^{m \times n}} \sum_{(i,u) \in \Omega} \left( \hat{X}_{iu} - X_{iu} \right)^2$$

known entries

subject to $\operatorname{rank}(\hat{X}) \leq r.$

# Combining the low-rank and neighborhood approaches



$$
\begin{array}{cc}
 & \begin{array}{cccccccccccccccc} & & & & & & & & & & \text{user } u & & & & & \end{array} \\
X = & \begin{pmatrix}
5 & ? & 2 & ? & ? & 5 & ? & ? & 3 & ? & 5 & ? & ? & ? & 2 & ? \\
3 & ? & 2 & ? & ? & 2 & ? & ? & ? & 3 & 2 & ? & 5 & ? & ? & ? \\
1 & ? & 5 & 2 & 3 & 4 & ? & 4 & ? & ? & ? & 2 & ? & ? & ? & ? \\
? & 1 & ? & 3 & ? & ? & ? & 3 & ? & ? & ? & ? & 2 & 1 & 5 & 5 \\
? & 4 & ? & ? & ? & ? & 5 & ? & ? & ? & 1 & ? & ? & 1 & ? & 4
\end{pmatrix}
\begin{array}{l} \\ \text{movie } i \\ \\ \\ \\ \end{array}
\end{array}
$$

$$
\approx \underbrace{\begin{pmatrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{pmatrix}}_{A} * \underbrace{\begin{pmatrix} ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \end{pmatrix}}_{B^T}
$$

$$
\min_{\hat{X} \in \mathbb{R}^{m \times n}} \sum_{(i,u) \in \underset{\text{known entries}}{\Omega}} \left(\hat{X}_{iu} - \overset{\text{known ratings}}{X_{iu}}\right)^2 + \beta \sum_{u,v} W_{uv} \|\hat{X}_{:,u} - \hat{X}_{:,v}\|^2
$$

subject to $\operatorname{rank}(\hat{X}) \leq r$.

# Combining the low-rank and neighborhood approaches

## Rewriting the neighborhood term

$$\sum_{u,v} W_{uv} \|\hat{X}_{:,u} - \hat{X}_{:,v}\|^2$$

$$\begin{aligned}
&= \textstyle\sum_{u,v} W_{uv} \big(\hat{X}_{:,u} - \hat{X}_{:,v}\big)^T \big(\hat{X}_{:,u} - \hat{X}_{:,v}\big) \\
&= \textstyle\sum_{u,v} W_{uv} (\hat{X}_{:,u}^T \hat{X}_{:,u} + \hat{X}_{:,v}^T \hat{X}_{:,v} - 2\hat{X}_{:,u}^T \hat{X}_{:,v}) \\
&= \textstyle\sum_{u,v} W_{uv} \big(\mathrm{diag}(\hat{X}^T \hat{X}) \mathbb{1}^T + \mathbb{1} \mathrm{diag}^T(\hat{X}^T \hat{X}) - 2\hat{X}^T \hat{X}\big)_{uv} \\
&= \mathrm{tr}\Big(W^T \big(\mathrm{diag}(\hat{X}^T \hat{X}) \mathbb{1}^T + \mathbb{1} \mathrm{diag}^T(\hat{X}^T \hat{X}) - 2\hat{X}^T \hat{X}\big)\Big) \\
&= \mathrm{tr}\Big(\big(\mathrm{Diag}(W\mathbb{1}) + \mathrm{Diag}(W^T\mathbb{1}) - 2W\big) \hat{X}^T \hat{X}\Big) \\
&= 2\mathrm{tr}( (\mathrm{Diag}(W\mathbb{1}) - W) \hat{X}^T \hat{X})
\end{aligned}$$

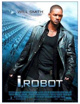
$$= 2\mathrm{tr}(\hat{X} \underset{\underset{\text{Laplacian of weighted graph } W}{|}}{L} \hat{X}^T)$$

# Combining the low-rank and neighborhood approaches

user $u$

$$X = \begin{pmatrix} 5 & ? & 2 & ? & ? & 5 & ? & ? & 3 & ? & 5 & ? & ? & ? & 2 & ? \\ 3 & ? & 2 & ? & ? & 2 & ? & ? & ? & 3 & 2 & ? & 5 & ? & ? & ? \\ 1 & ? & 5 & 2 & 3 & 4 & ? & 4 & ? & ? & ? & 2 & ? & ? & ? & ? \\ ? & 1 & ? & 3 & ? & ? & ? & 3 & ? & ? & ? & ? & 2 & 1 & 5 & 5 \\ ? & 4 & ? & ? & ? & ? & 5 & ? & ? & ? & 1 & ? & ? & 1 & ? & 4 \end{pmatrix}$$

movie $i$



$$\approx \underset{A}{\begin{pmatrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{pmatrix}} * \underset{B^T}{\begin{pmatrix} ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? & ? \end{pmatrix}}$$

known ratings

Laplacian of weighted graph $W$

$$\min_{\hat{X} \in \mathbb{R}^{m \times n}} \underbrace{\sum_{(i,u) \in \Omega}}_{\text{known entries}} \left(\hat{X}_{iu} - X_{iu}\right)^2 + \beta \; 2\mathrm{tr}(\hat{X} \, L \, \hat{X}^T)$$

known entries

subject to $\mathrm{rank}(\hat{X}) \leq r$.

# Combining the low-rank and neighborhood approaches

## Opportunities

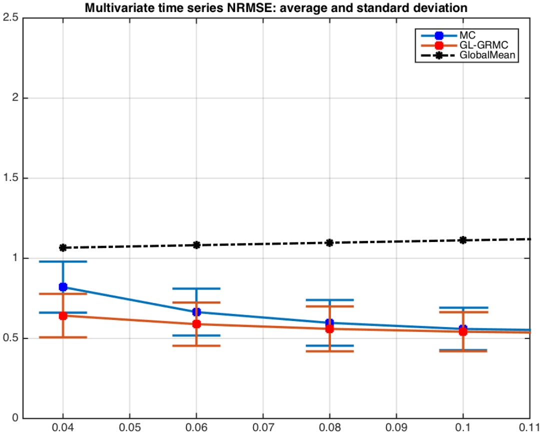
- How to choose the weights?
    - Let $W_{uv} := w(X_{:,u}, X_{:,v})$, where $w$ is a "simple" decreasing function of the distance between $X_{:,u}$ and $X_{:,v}$?
    - Set $w(X_{:,u}, X_{:,v}) = 0$ when the distance is too large, yielding a sparse graph $W$?
    - Formulate the choice of $W$ as an optimization problem, possibly with a sparsity-inducing regularizer?
    - Solve for $\hat{X}$ and $W$ alternatively?
    - Make use of side information?
- How to solve the optimization problem for $\hat{X}$?
    - Good news: When $\hat{X} = UV^T$, we have $\mathrm{tr}(\hat{X} L \hat{X}^T) = \mathrm{tr}(B^T L B A^T A)$.
    - Bad news: Alternating minimization becomes more costly with the graph term.
    - We have proposed a low-rank gradient descent methods that outperform alternating minimization.[1]

---

[1] Shuyu Dong, P.-A. Absil, K. A. Gallivan, *Riemannian gradient descent methods for graph-regularized matrix completion*, LAA, 2021, doi:10.1016/j.laa.2020.06.010.

# Combining the low-rank and neighborhood approaches

## Experimental results



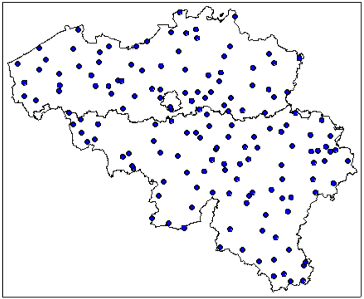
Root relative squared error (RRSE) vs sampling ratio on traffic data. Here the graph is given by $[W_{\varepsilon,\sigma}(X)]_{ij} = \mathbf{1}_{\geq\sigma}(\exp(-d(X_{i,:} - X_{j,:})/\varepsilon^2))$.

# Application to weather data completion

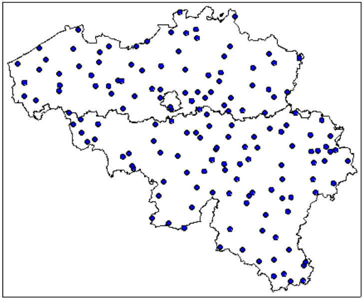
# Application to weather data completion

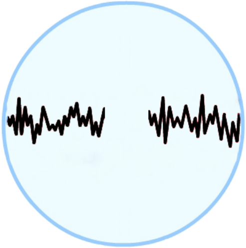
Ref: Benoît Loucheur, PA, Michel Journée, *Forthcoming paper*, 2023.

# Weather data completion: Gaps in weather data

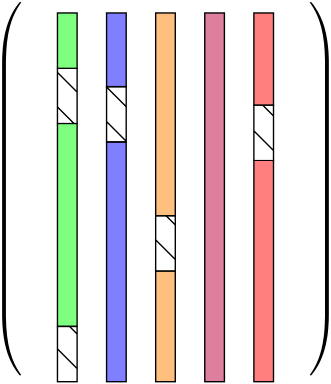


# Weather data completion: Gaps in weather data

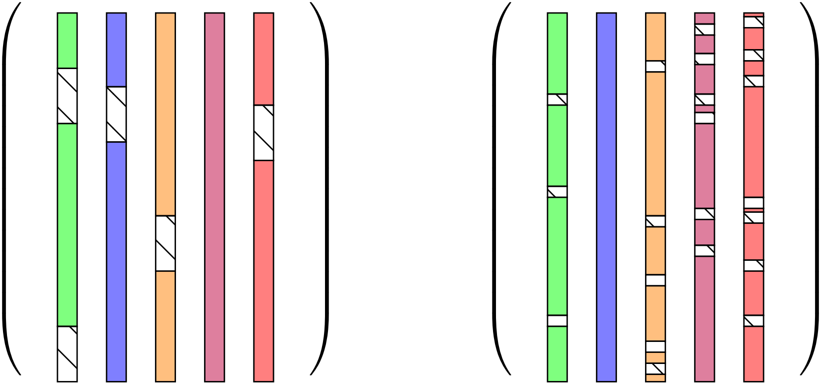




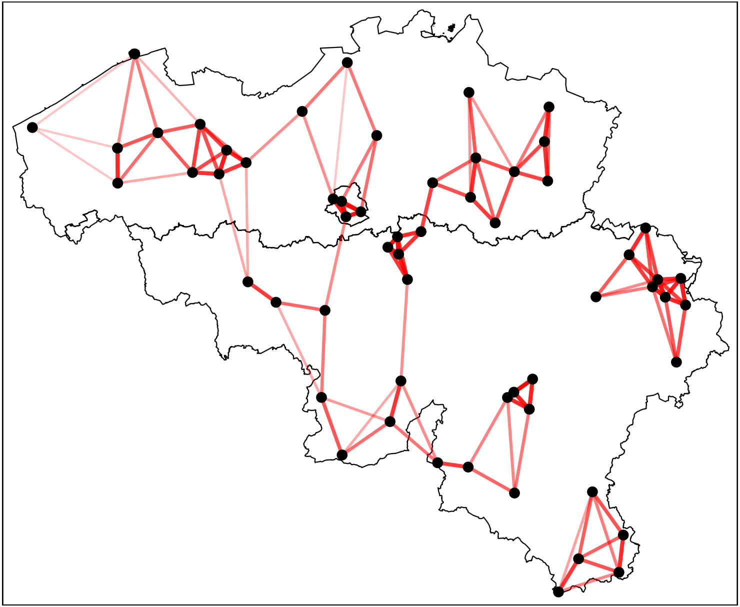
# Weather data completion: Gaps in weather data



# Weather data completion: Gaps in weather data



# Weather data completion: Example of spatial graph



# Weather data completion: Optimization model

$$\begin{aligned}\min_{\substack{A\in\mathbb{R}^{m\times r}\\ B\in\mathbb{R}^{n\times r}}} \quad & \frac{1}{2}\sum_{(i,u)\in\Omega}\left((AB^T)_{iu}-X_{iu}\right)^2 \quad + \quad \frac{\lambda_a}{2}\|A\|_F^2+\frac{\lambda_b}{2}\|B\|_F^2 \\ & +\frac{\lambda_L}{2}\left\{\operatorname{Tr}\left(A^T\mathbf{Lap}(W^a)A\right)+\operatorname{Tr}\left(B^T\mathbf{Lap}(W^b)B\right)\right\}\end{aligned}$$

# Weather data completion: Optimization model

$$\min_{\substack{A\in\mathbb{R}^{m\times r}\\ B\in\mathbb{R}^{n\times r}}} \quad \frac{1}{2}\sum_{(i,u)\in\Omega}\left((AB^T)_{iu} - X_{iu}\right)^2 \quad + \quad \frac{\lambda_a}{2}\|A\|_F^2 + \frac{\lambda_b}{2}\|B\|_F^2$$

$$+\frac{\lambda_L}{2}\left\{\operatorname{Tr}\left(A^T\mathbf{Lap}(W^a)A\right) + \operatorname{Tr}\left(B^T\mathbf{Lap}(W^b)B\right)\right\}$$

| | GRALS Conditions | Bloc | Spread |
|---|---|---|---|
| Case #1 | No constraints | 0.52 | 0.45 |
| Case #2 | $\lambda_L = \lambda_a = \lambda_b = 0$ | 0.95 | 0.93 |
| Case #3 | $\lambda_a = \lambda_b = 0$ | 0.89 | 0.91 |
| Case #4 | $\lambda_L = 0$ | 0.76 | 0.67 |
| Case #5 | $Lap(W^b) = 0$ | 0.72 | 0.61 |
| Case #6 | $Lap(W^a) = 0$ | 0.55 | 0.6 |

Table: Average RMSE (in °C) on the test set. Results for the two types of missing data generation with different constraints applied on the model.