RTRMC: low-rank matrix completion via trust-regions on the Grassmannian

Code online!

Nicolas Boumal and Pierre-Antoine Absil, UCLouvain (Belgium)

We consider the low-rank matrix completion problem, where X is a huge low-rank matrix of which we observe a few entries, and we wish to recover X.

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

$$= U \cdot W$$

known entries

$$\min_{U \in \mathbb{R}^{m\times r}, W \in \mathbb{R}^{r \times n}} \sum_{(i,j)\in\Omega} \left((UW)_{ij} - X_{ij} \right)^2$$

observation set

This essentially comes down to recovering the column space of X,

For fixed U, finding the best W is a least-squares problem:

$$W_U = \operatorname*{argmin}_{W \in \mathbb{R}^{r\times n}} \sum_{(i,j)\in\Omega} \left((UW)_{ij} - X_{ij} \right)^2$$

It remains to minimize this function of U:

$$f(U) = \sum_{(i,j)\in\Omega} \left((UW_U)_{ij} - X_{ij} \right)^2$$

but many matrices U yield the same model UW_U.

U and U' are *equivalent* for our purpose if:

$$UW_U = U'W_{U'} \,.$$

This happens if and only if:

$$\exists M \in \mathbb{R}^{r\times r}, \text{ invertible, such that } U' = UM \,.$$

That is, U and U' are equivalent if and only if they share the same column space.

Our search space is thus the set of column spaces (subspaces of $\mathbb{R}^m$ of dimension r), i.e., the Grassmann manifold $\mathrm{Gr}(m, r)$.

which in turn is an optimization problem on the Grassmann manifold.

Unfortunately, discontinuities arise when W_U is not uniquely defined.

To ensure smoothness, we regularize the objective function:

$$f(U) = \min_{W \in \mathbb{R}^{r\times n}} \sum_{(i,j)\in\Omega} \left((UW)_{ij} - X_{ij} \right)^2 + \lambda \sum_{(i,j)\notin\Omega} (UW)_{ij}^2$$

but it seems that we now need to compute the full product UW :(.

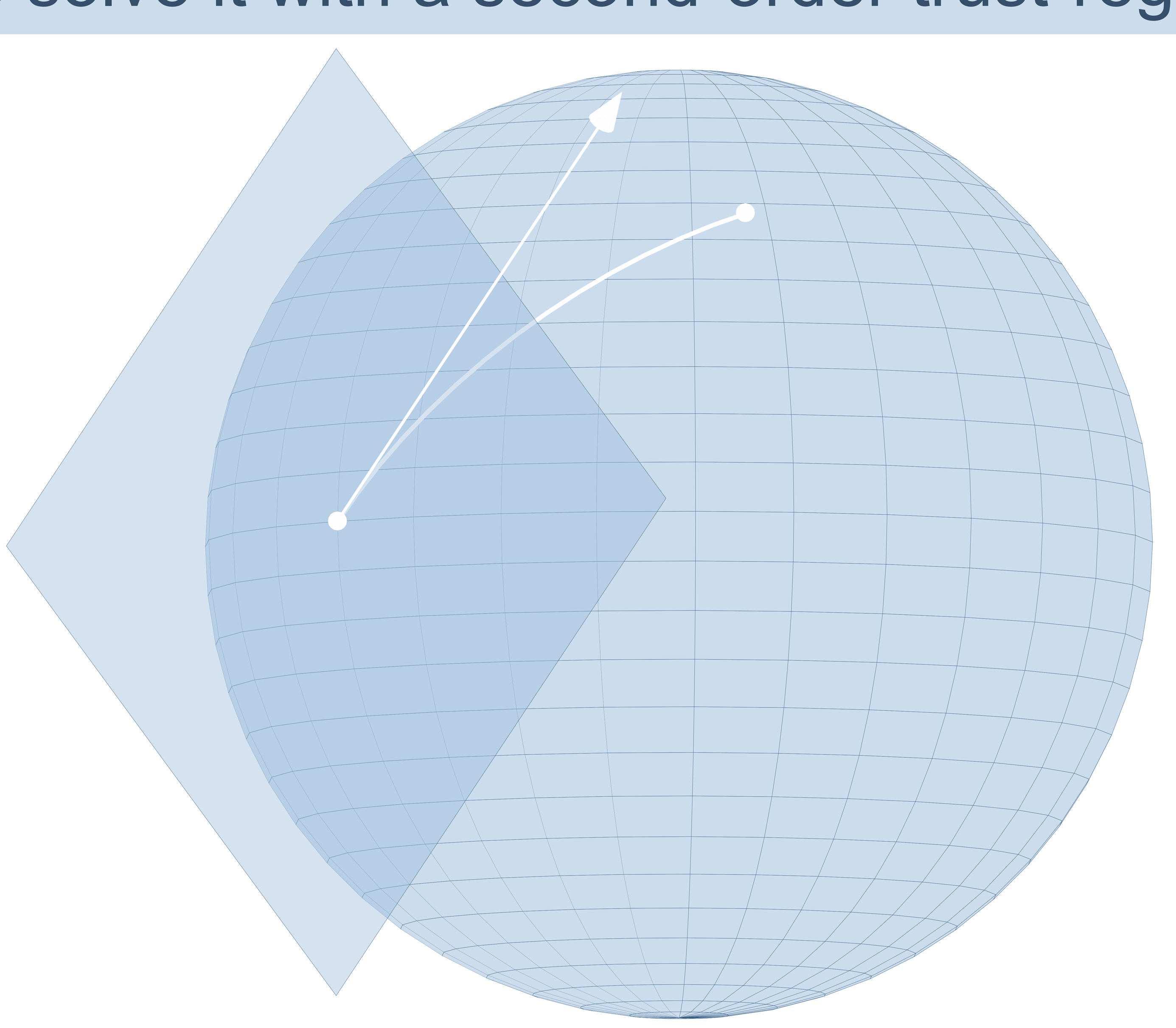
Choosing to work with orthonormal matrices U, this holds:

$$\sum_{(i,j)\notin\Omega} (UW)_{ij}^2 + \sum_{(i,j)\in\Omega} (UW)_{ij}^2 = \sum_{(i,j)} (UW)_{ij}^2 = \|UW\|_\mathrm{F}^2 = \|W\|_\mathrm{F}^2$$

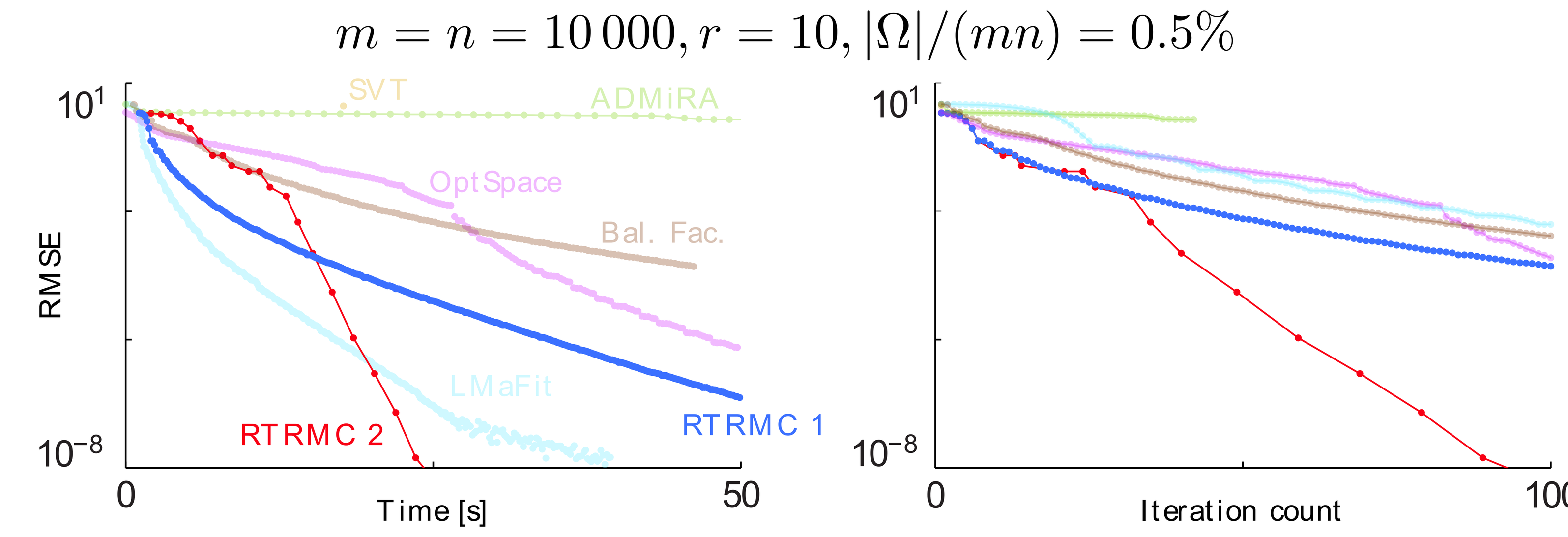
A cheap way of computing f, $\operatorname{grad} f$ and $\operatorname{Hess} f$ ensues.

We thus need to minimize a smooth function f defined over the smooth manifold $\mathrm{Gr}(m, r)$, which is a well studied problem.

We solve it with a second-order trust-region algorithm for Riemannian manifolds.



$m = n = 10\,000, r = 10, |\Omega|/(mn) = 0.5\%$



$m = 1\,000, n = 30\,000, r = 5, |\Omega|/(mn) = 2.6\%$