Recommender Systems via Matrix Completion: Frank-Wolfe, Pairwise Frank-Wolfe, and Projected Gradient Methods

Ilaria Boschetto, Federico Clerici, Raffaele D'Agostino, Gabriele Villa

Dipartimento di Matematica "Tullio Levi-Civita"
Università degli Studi di Padova

A.A. 2024-2025

Matrix Completion: Problem Formulation

- Goal: recover a low-rank matrix $X \in \mathbb{R}^{n_1 \times n_2}$ from a subset of observed entries $\{U_{ij}\}_{(i,j) \in J}$.
- Original non-convex formulation:

$$\min_{X \in \mathbb{R}^{n_1 \times n_2}} \sum_{(i,j) \in J} (X_{ij} - U_{ij})^2 \quad \text{s.t.} \quad \mathrm{rank}(X) \leq \delta$$

- Relaxation: replace rank constraint with nuclear norm constraint

$$\min_{X \in \mathbb{R}^{n_1 \times n_2}} \sum_{(i,j) \in J} (X_{ij} - U_{ij})^2 \quad \text{s.t.} \quad \|X\|_* \leq \delta$$

- Feasible set: convex hull of rank-one matrices

$$C = \mathrm{conv}\{\delta\, \mathrm{u}\mathrm{v}^\top : \|\mathrm{u}\| = \|\mathrm{v}\| = 1\}$$

Algorithms

1. **Projected gradient method**: performs gradient descent steps followed by projection onto the feasible set to enforce constraints.
2. **Frank-Wolfe method**: a projection-free algorithm that moves toward a solution of a linear minimization oracle over the constraint set.
3. **Pairwise Frank-Wolfe method**: improves standard FW by selecting a pair of atoms (away and standard FW) and performing an update in their direction, accelerating convergence.

Linear Minimization Oracle (LMO)

- At each Frank-Wolfe iteration, solve:

$$\mathrm{LMO}_C(\nabla f(X_k)) = \arg\min_{\|X\|_* \le \delta} \mathrm{tr}(\nabla f(X_k)^\top X)$$

- Optimal solution:

$$S_k = \delta \cdot \mathsf{u}_1 \mathsf{v}_1^\top$$

  where u_1, v_1 are the top left and right singular vectors of $-\nabla f(X_k)$.
- In code, computed via:

```
u, s, vt = svds(-grad, k=1, solver="arpack")
```

- `ARPACK` is used for efficiency: it computes only the leading singular vectors, making it well-suited for large and sparse matrices.

Performance Tracking and Stopping Criteria

- **Loss Function:** Measures the squared error over observed entries:
$$f(X) = \sum_{(i,j)\in J} (X_{ij} - U_{ij})^2$$
- **Duality Gap:** Quantifies sub-optimality in Frank-Wolfe methods:
$$g_k = \langle \nabla f(X_k), X_k - S_k \rangle$$
where S_k is the solution of the Linear Minimization Oracle (LMO).
- **Relative Change Between $\hat{X}_k$ and X_k:** Used as a stopping criterion for the Projected Gradient method:
$$\frac{\|\hat{X}_k - X_k\|_F}{\|X_k\|_F}$$

Stepsizes

We implemented the following stepsize strategies:

- **Diminishing stepsize**
- **Exact line search stepsize**
- **Lipschitz constant dependent stepsize**
- **Armijo line search stepsize**

$\Rightarrow$ The *Exact line search stepsize* has been used in the whole experiment

Tuning the δ parameter

- A validation-based approach was adopted to select the optimal value of the parameter δ.
- δ controls the threshold in the matrix completion model.
- A range of candidate δ values was tested.
- The δ yielding the best validation performance (lowest RMSE) was selected.

$\Rightarrow$ Maintaining a low-rank matrix has been preferred over lower error

Datasets

Amazon Gift Cards

- Sparse (4.9% density) with 377 users, 129 items (1-5 ratings)
- Latest two user interactions per user are used for training, the (N-1)-th for validation and the N-th for testing

Netflix

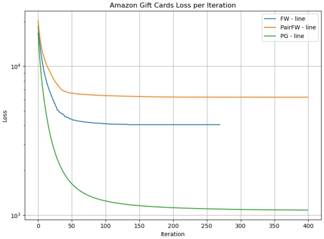
- Integer movie ratings (1-5) from 1999-2005
- Subset created by selecting top 95% movies and 99% users (by rating count), further restricted to 2005 ratings to reduce bias
- Final dataset consisting of 4694 users and 225 movies

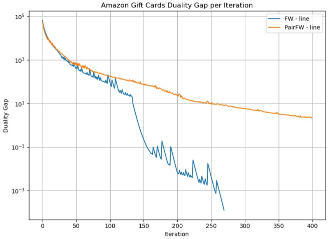
MovieLens 100K

- 100,000 integer movie ratings (1-5) from 943 users and 1682 films, with a 6.3% density

Amazon Gift Cards

Observed matrix rank: 129

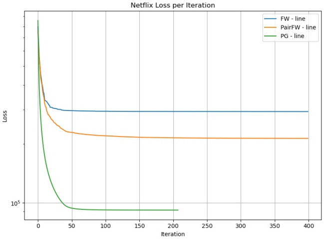


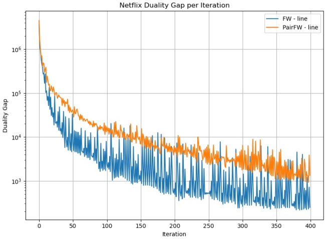


Algorithm	**RMSE Test**	**Rank**	**Time(sec)**
Projected Gradient	0.930	4	3.210
Pairwise FW	1.857	87	4.162
Standard FW	1.537	2	0.653

Netflix

Observed matrix rank: 225

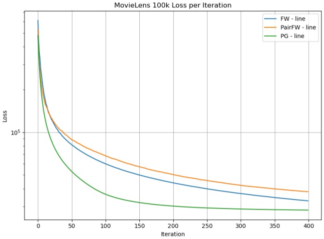


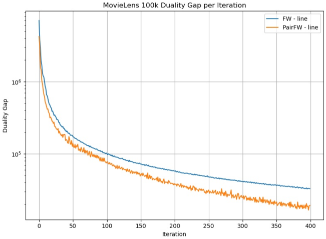


Algorithm	RMSE Test	Rank	Time(sec)
Projected Gradient	0.889	96	19.380
Pairwise FW	1.089	188	101.972
Standard FW	1.263	68	11.622

MovieLens

Observed matrix rank: 943





Algorithm	RMSE Test	Rank	Time(sec)
Projected Gradient	0.951	115	260.403
Pairwise FW	0.980	379	2760.050
Standard FW	0.980	401	47.662