Side Information Regularized Matrix Factorization

Paul-Henri Perrin, Florian Yger, Jamal Atif, Dario Colazzo
paul-henri.perrin@easycrowd.net, florian.yger@dauphine.fr, jamal.atif@dauphine.fr, dario.colazzo@dauphine.fr

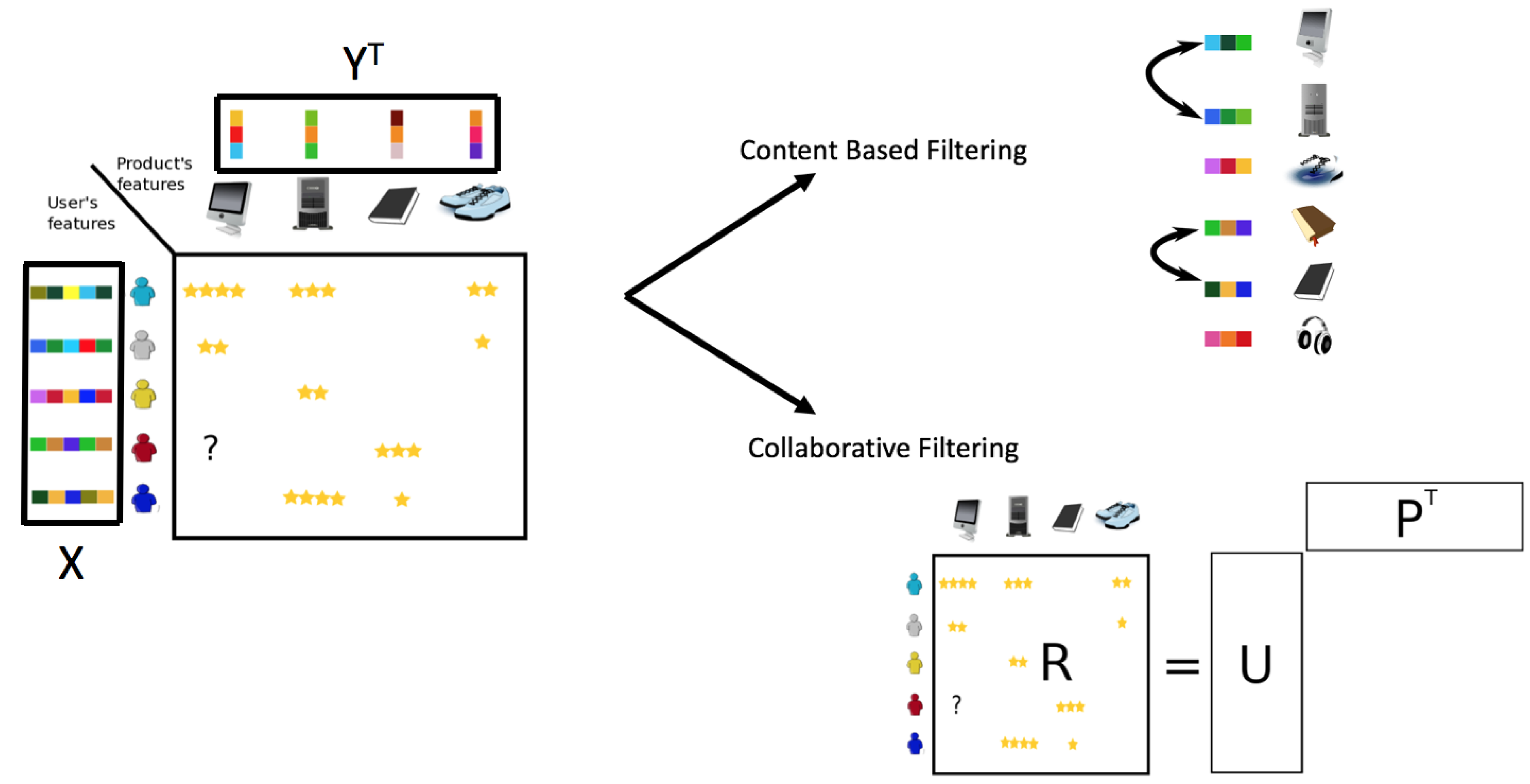
Problem

There are two main approaches to user-item recommender systems[1].

- Content-based filtering relies on the characteristics of users and items to produce recommendations.
- Collaborative filtering relies on the relationship between users and items that emerges from the ratings given by the former to the latter.

Our approach aims to use side-information to better constraint the matrix factorization problem[3] of collaborative filtering.

Recommender Systems



Ongoing works

Available data

- Synthetic Dataset designed to study the charcteristics of our model
- Movielens[5] Dataset to confront the model to real-world data
- In-house Dataset to apply the model to EasyCrowd use cases

Goals

- Improve the performances on learning and ranking metrics (RMSE, TopN[4])
- Address the cold-start problem using side information

Possible Interpretations

- Data Augmentation: Factorization of an extended matrix $\tilde{R} = \begin{bmatrix} R & X \\ Y^\top & 0 \end{bmatrix}$
- Joint Embedding
- Probabilistic Bayesian Network

Difficulties

- Many hyper-parameters to fine tune
- Long computation time

Matrix Factorization with λ Weighted Regularization

The model we use as a baseline is the simple approach described in [2]. The loss function we aim to minimize is

$$\mathcal{L}(U,P) = \|M \circ (R - UP^\top)\|_F^2 + \lambda(\|U\|_F^2 + \|P\|_F^2)$$

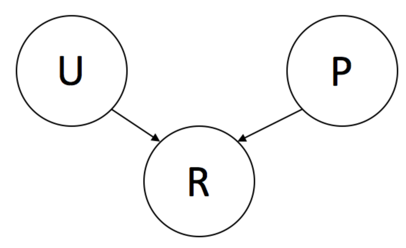


Figure 1: Schema for the baseline model: we determine U and P that best reconstruct R

Inference Rules

$$\begin{cases} \forall i \in [1, n_u], \mathbf{u}_{i\cdot} = \left(\sum_j^{n_p} m_{ij} r_{ij} \mathbf{p}_{j\cdot}\right) \times \left(\sum_j^{n_p} m_{ij} \mathbf{p}_{j\cdot}^\top \mathbf{p}_{j\cdot} + \lambda I_{n_f}\right)^{-1} \\ \forall j \in [1, n_p], \mathbf{p}_{j\cdot} = \left(\sum_i^{n_u} m_{ij} r_{ij} \mathbf{u}_{i\cdot}\right) \times \left(\sum_i^{n_u} m_{ij} \mathbf{u}_{i\cdot}^\top \mathbf{u}_{i\cdot} + \lambda I_{n_f}\right)^{-1} \end{cases}$$

Latent Variables to constraint Side Information

We introduce two latent variables W_U and W_P of respective size $k_u \times n_f$ and $k_p \times n_f$ such that

$$\begin{cases} U^\top X W_U = I_{n_f} \\ P^\top Y W_P = I_{n_f} \end{cases}$$

These matrices will be approximated in a way similar to U and P. Due to the additional degrees of freedom, we must include additional regularization terms. The loss function that we are trying to minimize takes the following form:

$$\mathcal{L}(U,P,W_U,W_P) = \|M \circ (R - UP^\top)\|_F^2 + \lambda\|U\|_F^2 + \mu\|P\|_F^2$$
$$+ \gamma\|X - UW_U^\top\|_F^2 + \eta\|W_U\|_F^2 + \xi\|Y - PW_P^\top\|_F^2 + \theta\|W_P\|_F^2 + \zeta\|W_P W_U^\top\|_F^2$$

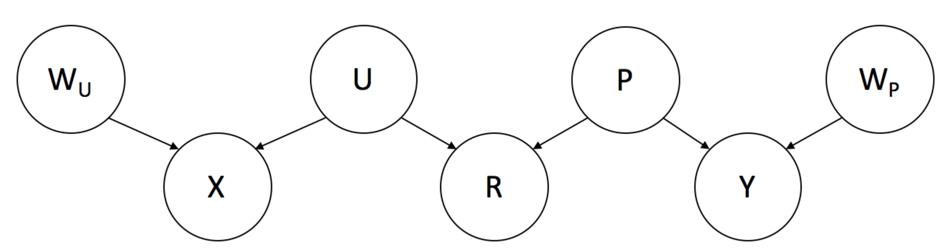


Figure 2: Schema for the SILV model: not only must U and P reconstruct R, but U and W_U (resp. P and W_P) must reconstruct X (resp. Y)

Inference Rules

$$\begin{cases} \forall i \in [1, n_u], \mathbf{u}_{i\cdot} = \left(\sum_j^{n_p} m_{ij} r_{ij} \mathbf{p}_{j\cdot} + \gamma \sum_t^{k_u} x_{it} \mathbf{w}_{U,t\cdot}\right) \left(\sum_j^{n_p} m_{ij} \mathbf{p}_{j\cdot}^\top \mathbf{p}_{j\cdot} + \gamma \sum_t^{k_u} \mathbf{w}_{U,t\cdot}^\top \mathbf{w}_{U,t\cdot} + \lambda I_{n_f}\right)^{-1} \\ \forall t \in [1, k_u], \mathbf{w}_{U,t\cdot} = \left(\gamma \sum_i^{n_u} x_{it} \mathbf{u}_{i\cdot}\right) \left(\gamma \sum_i^{n_u} \mathbf{u}_{i\cdot}^\top \mathbf{u}_{i\cdot} + \eta I_{n_f} + \zeta \sum_s^{k_p} \mathbf{w}_{P,s\cdot}^\top \mathbf{w}_{P,s\cdot}\right)^{-1} \end{cases}$$

References

[1] M. Montaner and B. Lopez and J. Lluis de la Rosa: *A Taxonomy of Recommender Agents on the Internet*, Artifical Intelligence Review (2003)

[2] Y. Zhou and D. Wilkinson and R. Schreiber and R. Pa *"Large-Scale Parallel Collaborative Filtering for the Netflix Prize*, Proceedings of AAIM (2008)

[3] Gulyas, Laszlo et al.: *Matrix Completion with Noise*, Proceedings of IEEE (2010)

[4] J. Delporte and A. Karatzoglou and T. Matuszczyk and S. Canu: *Socially Enabled Preference Learning from Implicit feedback data*, Proceedings of ECML/PKDD, Singapore (2013)

[5] Gulyas, Laszlo et al.: *The Movielens Datasets: History and Context*, ACM Trans. Interact. Intell. Syst. (2016)

Paris-BD 2017

TELECOM PARISTECH
46 rue Barrault, 75013 Paris
9 Mai 2017