02 Learning feature combinations from multiple tasks

- MOFM: low-rank regression for learning common factors -

Abstract

Multi-output Factorization Machines (MOFM) are an extension of Convex Factorization Machines that can learn the model of several tasks simultaneously. MOFM can find combinations of factors that are predictive across tasks.

MOFM decompose the potentially very large weight matrix associated with each task using a small number of common basis vectors. Hence, MOFM are able to scale to very high-dimensional data. In addition, we propose a convex formulation for learning this decomposition with optimality guarantee.

MOFM find applications to numerous real-world problems, including medical diagnosis, recommendation systems and genomic selection of plants. In future work, we plan to further study the theoretical properties of MOFM.

Multi-Output Factorization Machines (MOFM) learn the model of several tasks simultaneously and discover common combinations of factors across tasks

- General-purpose (large-number of applications)
- Take advantage of correlations between tasks (decompose each weight matrix using common basis vectors)
- Parameter learning is easy (optimal solution is guaranteed, insensitive to initialization)
- Can handle a large number of feature combinations (MOFM scale to very high-dimensional data)

$\hat{y}$: target variable

x : features

w : 1st order weights

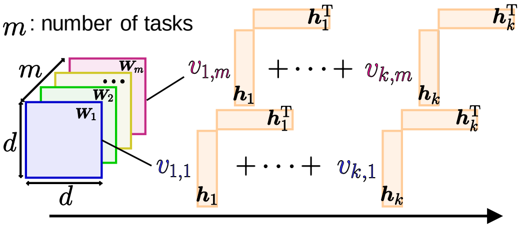
W : 2nd order weights

$$\hat{y}_m = \mathbf{x}^T w_m + \mathbf{x}^T W_m \mathbf{x} \quad \leftarrow \text{task number } m$$

$$\vdots$$

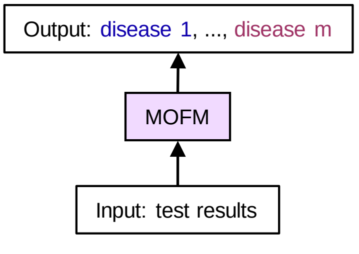
$$\hat{y}_2 = \mathbf{x}^T w_2 + \mathbf{x}^T W_2 \mathbf{x}$$

$$\hat{y}_1 = \mathbf{x}^T w_1 + \mathbf{x}^T W_1 \mathbf{x}$$



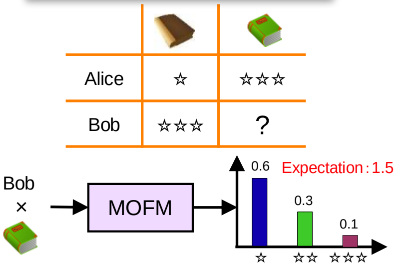
Incrementally add basis vectors $h_1,\dots,h_k$ common to each task

Application 1: Medical diagnosis



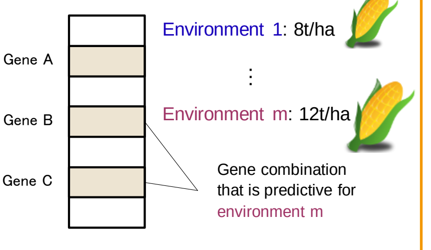
Discover common combinations of factors to predict one of m diseases

Application 2: Recommend books to users



Predict the probability of m possible ratings and recommend books based on the expected rating

Application 3: Predict corn yield from DNA



Simultaneously predict corn yield from gene combinations in m possible environments

References

[1] M. Blondel, A. Fujino, N. Ueda, "Convex Factorization Machines," in *Proc. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases,* 2015.

[2] M. Blondel, V. Niculae, T. Otsuka, N. Ueda, "Multi-output Polynomial Networks and Factorization Machines," in *Proc. Neural Information Processing Systems,* 2017.

Contact

Mathieu Blondel Ueda Research Laboratory