User Purchase Prediction

Divya Venugopalan, Maryam Tavakol,
Prof. Dr. Ulf Brefeld
TU Darmstadt, Germany

Objectives

- Predict products that user will purchase
- Identify patterns in general user purchase
- Improve recommendation system

Introduction

In the present day world, a user is inundated with many options when he wants to buy a product. It is very hard for a person to find the product that he is looking for as the market is filled with products that cater to the needs of all kinds of people. For better user satisfaction and loyalty, the online businesses need a method to display the right products to their customers. Recommender systems resolve this situation. Recommender systems aim to make user decisions easy and also help online markets improve their business by analyzing the the users purchase history and recommending products that suit the users taste.

Recommender systems are generally based on content filtering and collaborative filtering. In content filtering, the characteristics of the user or item is collected to determine the extent to which the user likes the item. Unfortunately this method requires data that is difficult to obtain or calibrate. Collaborative filtering is an alternative method that uses users previous purchase information, the relationship between users and the items for recommendation. One of the most efficient implementations of this method is using matrix factorization.

Data

We have a three way tensor, in which mode one represents the item recommended, mode two represents the item chosen and mode three represents the set of products purchased previously. If we take a tensor $\mathcal{X}$, its element x_{ijk} will denote the probability of the item j being purchased when item i was recommended based on previous item set k.

Factorization Models

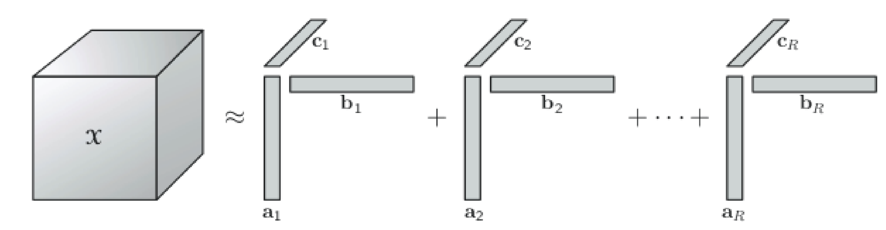

Figure 1: CP Decomposition

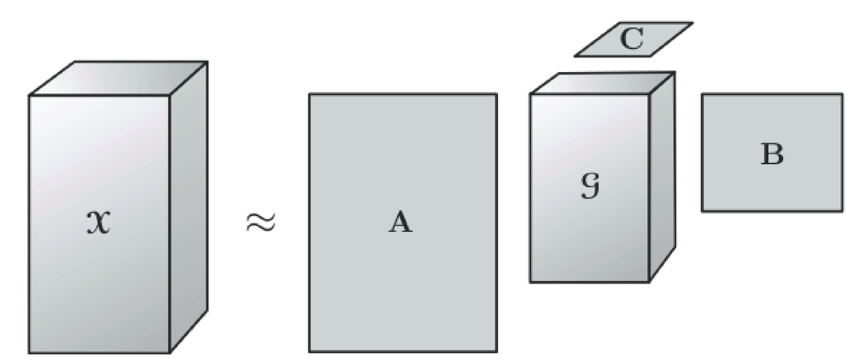

Figure 2: CP Decomposition

Methodology

Tucker Decomposition A tensor is decomposed into a core tensor which needs to be transformed by a matrix along each mode. Therefore for a three-way tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$, we have

$$\mathcal{X} \approx \hat{\mathcal{X}} = \mathcal{G} \times_1 A \times_2 B \times_3 C \qquad (1)$$

CP Decomposition This model expresses a tensor as a sum of finite number of rank-one tensor. A rank-one tensor is a N-way tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_n}$ which can be written as the outer product of N vectors, ie. ,

$$\mathcal{X} = \mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \cdots \circ \mathbf{a}^{(N)} \qquad (2)$$

Pairwsie Interaction To decompose a tensor using this method involves solving a constrained convex program which minimizes the weighted sum of nuclear norms of matrices. [1]
The Pairwise Interaction can be modeled for a tensor $\mathcal{X}$ of size $I \times J \times K$ as:

$$\hat{x}_{ijk} = \langle \mathbf{u}_i^{(a)}, \mathbf{v}_j^{(a)} \rangle + \langle \mathbf{u}_j^{(b)}, \mathbf{v}_k^{(b)} \rangle + \langle \mathbf{u}_k^{(c)}, \mathbf{v}_i^{(c)} \rangle, \qquad (3)$$

$\forall\ (i, j, k) \in [I] \times [J] \times [K]$,
where $\mathbf{u}_i^{(a)}{}_{i\in[I]}$, $\mathbf{v}_i^{(a)}{}_{j\in[J]}$ are r_1 dimensional vectors, $\mathbf{u}_j^{(b)}{}_{j\in[J]}$, $\mathbf{v}_k^{(b)}{}_{k\in[K]}$ are r_2 dimensional vectors
and $\mathbf{u}_k^{(c)}{}_{k\in[K]}$, $\mathbf{v}_i^{(c)}{}_{i\in[I]}$ are r_3 dimensional vectors, respectively [1].

Results

Table 1: Results

Method	MSE
Tucker (Core-2x2x2)	1.7
Tucker (Core-5x5x5)	1.58
Tucker (Core-10x10x10)	1.3
CP (R=3)	14.51
CP (R=5)	10.2
CP (R=10)	9.45
PITF (factors=5)	3.5
PITF (factors=10)	**0.79**
PITF (factors=15)	**0.76**

Conclusion

We demonstrate that the Pairwise Interaction tensor factorization method gives us the best prediction or least mean square error for the given probability dataset. Though Tucker and CP methods subsume PITF, the reason for PITF giving better results can be attributed to the fact that Tucker method involves finding gradients iteratively over core tensor, with complexity of $O(n^3)$ and PITF over matrices with complexity of $O(nlog^2(n))$.

References

[1] Shouyuan Chen, Michael R Lyu, Irwin King, and Zenglin Xu.
Exact and stable recovery of pairwise interaction tensors.
In *Advances in Neural Information Processing Systems*, pages 1691–1699, 2013.

[2] Guy Shani, Ronen I Brafman, and David Heckerman.
An mdp-based recommender system.
In *Proceedings of the Eighteenth conference on Uncertainty in artificial intelligence*, pages 453–460. Morgan Kaufmann Publishers Inc., 2002.

Contact Information

- Email: divya1993@gmail.com