Adaptive Collaborating Filtering: The Low Noise Regime

Onkar Dabeer
School of Technology and Computer Science
Tata Institute of Fundamental Research
Mumbai, India
Email: onkar@tcs.tifr.res.in

***Abstract*— In this paper, we study collaborative filters that adapt future recommendations based on feedback from users. We consider discrete time and at each time a random user seeks a recommendation. The collaborative filter uses all past data available to make a recommendation, the user then provides binary feedback indicating whether he liked the item (rating 1) or not (rating 0), and this feedback is used by the collaborative filter for future decisions. In this setting, ideally the goal is to maximize the long run time average of the ratings, but practical considerations lead us to a moving horizon approximation. Our main result identifies a collaborative filter that optimizes a moving horizon cost in the limit as the noise in the ratings vanishes.**

I. Introduction

Collaborative filtering is commonly used to recommend relevant content to users based on their past ratings. In the case of binary ratings, the user m rates the item n as $Y_{m,n} \in \{0, 1\}$, where a 1 means the item is liked and a 0 means that it is not liked. Typically, very few entries of the matrix $\boldsymbol{Y} = [Y_{m,n}]$ are known and the collaborative filter recommends items to users based on the known ratings. Collaborative filtering is commonly used in practice [7], [11]. In the last few years a number of papers have investigated theoretical limits and identified provably good algorithms using a variety of mathematical models for the matrix $\boldsymbol{Y}$ - see for example [1], [2], [6], [9], [13] and references therein. These works do not model temporal dynamics of collaborative filtering - the performance metric (matrix completion in [1], [6], [9] and recommendation error in [2]) is for a single instance. In practice, however, we use the collaborative filter to make recommendations to users, then some of the users provide feedback in the form of ratings, which are used again for future recommendations. This is true for example in recommendation systems such as Amazon [11] and Youtube, where buys and clicks from recommended content are tracked. In this paper, we consider a simple model to study collaborative filtering that accounts for user feedback. We note that collaborative filters that exploit timestamps of the ratings have been used in the literature (see for example [10], [12]), but these works do not consider explicit user feedback.

We assume discrete time $t = 1, 2, ...$ Suppose at time t a user M_t seeks a recommendation. A collaborative filter recommends item N_t to the user M_t. The user then provides the feedback Y_{M_t,N_t} to indicate whether he liked the recommendation (a rating of 1) or not (a rating of 0). The collaborative filter is taken to be causal - at time t it relies solely on the information $\{M_s, N_s, Y_{M_s,N_s}, 1 \leq s \leq t-1\}$ and M_t. Ideally, we seek collaborative filters that maximize the long-run time average

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E[Y_{M_t,N_t}] \tag{1}$$

where the expectation is over the random variables involved in the model (see Section II for details). Thus the collaborative filter in our setting has two roles: a) providing relevant recommendations at the current time, b) sampling of the matrix $\boldsymbol{Y}$ (since the user gives feedback for every recommendation) so that future recommendations are good. Since the sampling that takes into account the response to past samples is referred to as *adaptive sampling* in the statistical and signal processing community, we refer to a collaborative filter in our setup as a "adaptive collaborative filter" (ACF). We note that our nomenclature is different than that used in [14], which uses the same name for a different setup.

The mathematical model we use for the ratings $\boldsymbol{Y}$ is similar to that in [1], [2]. In particular, the model posits that users and items are of a finite number of types, the types alone determine the ratings, and there is also noise in the ratings. The problem of finding an ACF that maximizes (1) is a sequential decision problem and it can be cast in the framework of ergodic Markov decision processes [5], [3]. However, in our case the state comprises of the entire past and the problem is intractable. Therefore, in this paper, we consider a less ambitious goal: at time t is we seek to maximize $R_H = \sum_{h=0}^{H} E[Y_{M_{t+h},N_{t+h}}]$ assuming that the system stops at time $t + H$. This leads to a finite horizon optimization problem and is referred to as the moving horizon approximation [8], [4]. Our main result identifies an ACF, which maximizes R_H for any finite H in the limit as the noise in the ratings vanishes. The asymptotically optimal ACF first assigns types to users and items so as to maximize a specified weight, and for a user of type k, it recommends items of a type that is most popular amongst users of type k. Thus the asymptotically optimal ACF keeps track of user and item clusters and recommends locally popular items. This "cluster + local popularity" structure is commonly seen in neighborhood collaborative filters without feedback studied in the literature (see [2] and references therein). Our results show that the same structure is optimal in the low noise regime even with explicit user feedback.

Our setup can also be used to gain insight into the structure of the optimal ACF in the high noise regime, but due to space constraints we do not present these results. Also, we do not consider implementation issues on real data; our focus is on identifying an optimal ACF in the low noise regime irrespective of its complexity.

The paper is organized as follows. In Section II, we describe our model and assumptions. In Section III, we present the main results and give the proofs. Some intermediate lemmas are proved in the Appendix and the conclusion is given in Section IV.

II. Data Model and Problem Formulation

We index users with $m \in \{1, 2, ...\}$ and items by $n \in \{1, 2, ...\}$. Though the number of users and items is unbounded, we only consider finite time and ratings from only finitely many users and items are exposed at any time. We take $\{M_t\}_{t\geq 1}$ to be independent and assume that $P(M_t = m) > 0$ for any m. A common belief in the collaborative filtering literature is that users share tastes and items share characteristics. We use (a minor variant of) the model in [1], [2] to capture this belief. We assume that users are of K different types and items are of L different types, but given an user (or an item), we do *not* know its type. Let U_m and V_n denote the types of user m and item n respectively. We assume $\{U_m\}$ are i.i.d. uniform over $\{1, ..., K\}$ and $\{V_n\}$ are i.i.d. uniform over $\{1, ..., L\}$. Since K, L are assumed to be finite (though typically large), it follows that we have infinitely many users and items of every type. Furthermore, we assume $\{U_m\}$ and $\{V_n\}$ are independent. The rating $Y_{m,n}$ depends only on the types of the user m and the item n in the following manner. Suppose that $\{A(k,l)\}_{k=1,l=1}^{K,L}$ are i.i.d. Bernoulli(1/2) and define the matrix $\boldsymbol{X}$ such that its (m,n)th entry $X_{m,n} = A_{U_m,V_n}$. The matrix $\boldsymbol{Y}$ is obtained by passing the entries of $\boldsymbol{X}$ through a memoryless binary symmetric channel with error probability δ, that is, given $\boldsymbol{X}$, the entries of $\boldsymbol{Y}$ are independent with

$$P(Y_{m,n} = X_{m,n}|\boldsymbol{X}) = 1 - \delta,$$
$$P(Y_{m,n} \neq X_{m,n}|\boldsymbol{X}) = \delta.$$

We note that the entries of $\boldsymbol{Y}$ are not known and the only way to reveal them is by making recommendations.

Remark: Our assumption that $\{A(k,l)\}_{k=1,l=1}^{K,L}$ are i.i.d. Bernoulli(1/2) ensures that data from a single user (that is, a single row) cannot be used to predict the rating of that user for unseen items. Thus, under this assumption, it is necessary to pool information from other users to make recommendations that are better than random guessing, that is, the "collaborative" aspect is a must.

A standard framework for maximizing (1) is that of ergodic Markov decision processes (MDP) and associated dynamic programming [5], [3]. For our model the entire past needs to be chosen as the state and the dynamic program is not tractable. Hence we consider the moving horizon approximation [4], [8]. In this case, the recommendation at time t is decided assuming that the system stops after H time ticks and applying the optimal strategy for the corresponding finite-horizon MDP.

III. Main Results

In Section III-A, we define an ACF of interest, in Section III-B we state an optimality result for it, and in Sections III-C, III-D, we prove the result.

A. An ACF Based on Maximum Type, Local Popularity

We do not know the types of the users and the items involved in the ratings available up to time $(t-1)$. Let $\mathcal{T}_{t-1}$ be the set of possible type assignments for users and items observed up to time $(t-1)$. Up to time $(t-1)$, we sample information about only finitely many users and items, and hence $\mathcal{T}_{t-1}$ is a finite set. Given a type assignment $\tau \in \mathcal{T}_{t-1}$, we associate a weight w_τ to it as follows. Given the type τ, let

$$\begin{aligned} &J_{k,l}^i(t-1) \\ &= \Big|\{s : 1 \leq s \leq (t-1), y_s = i, U_{m_s} = k, V_{n_s} = l\}\Big| \end{aligned} \tag{2}$$

where U_{m_s}, V_{n_s} are types of m_s, n_s respectively under the type assignment. Thus $J_{k,l}^i(t-1)$ is the number of ratings of value i given by users of type k to items of type l (as per the type assignment τ). The weight w_τ is defined as

$$w_\tau := \sum_{k=1}^{K} \sum_{l=1}^{L} \max\{J_{k,l}^0(t-1), J_{k,l}^1(t-1)\}. \tag{3}$$

The Max-Type, Local-Popularity (MT-LP) ACF makes a recommendation at time t as follows. First it finds the maximum weight type assignments: $\tau_* = \arg\max_{\tau\in\mathcal{T}_{t-1}} w_\tau$, where τ_* is a list of type assignments in case the maximum is not unique. Suppose $M_t = m$. If user m has appeared for the first time, then it recommends an item that is most popular [1] amongst all users observed so far. If user m has appeared before and its type under some type assignment in τ_* is k, then it recommends an unseen item of type l such that $J_{k,l}^1(t-1) > J_{k,l}^0(t-1)$ (that is, an item popular amongst items of type k). If there is more than one such type l, we resort to randomization. If no such l exists, then it recommends a new item randomly.

B. Moving Horizon Approximation

We recall that our goal at time t is to maximize $R_H = \sum_{h=0}^{H} E[Y_{M_{t+h},N_{t+h}}]$ based on the assumption that the decisions at time $t+1$,..., $t+H$ are made by the corresponding optimal schemes. For causal schemes, this reduces to maximizing the conditional expectation

$$\tilde{R}_H = \sum_{h=0}^{H} E[Y_{M_{t+h},N_{t+h}}|S_{t-1}, M_t]$$

[1] By most popular item we mean an item for which the number of ratings 1 minus the number of ratings 0 is maximum.

where S_{t-1} denotes the event $\{M_s = m_s, N_s = n_s, Y_{M_s,N_s} = y_s, 1 \le s \le t-1\}$. Our main result is the following.

Proposition 1: For the model in Section II, the following are true.

1) There exists a constant $C_{t,\delta}$ (that does not depend on N_t) such that for any ACF

$$\lim_{\delta \to 0} \frac{\log(\delta \tilde{R}_H / C_{t,\delta})}{\log((1-\delta)/\delta)} \text{ exists.} \tag{4}$$

2) Suppose that given a user type k, there exists some type assignment in τ_* and an item type l such that $J^1_{k,l}(t-1) > J^0_{k,l}(t-1)$. Similarly suppose that given an item type l, there exists some type assignment in τ_* and an user type k such that $J^1_{k,l}(t-1) > J^0_{k,l}(t-1)$. The maximum value of the limit above is achieved by the ACF MT-LP.

Proof: Due to space constraints, in this paper we only prove the result for $H = 0$ and $H = 1$. The proof for these cases is given in the following subsections. The proof for general H follows the same main ideas.

If we denote the limit in (4) by e_*, then Part 1 of the above proposition states that for small δ (the low noise regime),

$$\tilde{R}_H \approx \frac{C_{t,\delta}}{\delta} \left\{ \frac{1-\delta}{\delta} \right\}^{e_*}.$$

Since only the exponent e_* depends on the recommendation N_t at time t, it is natural to seek to maximize it, and Part 2 of the proposition states that the ACF MT-LP attains the maximum. The extra assumption needed for Part 2 of the proposition demands that each user type has a popular item type under some maximum type assignment and each item type is popular amongst some user type under some maximum type assignment. By making suitable assumptions about the initialization of the MT-LP ACF and analyzing its dynamics it seems possible to remove this assumption, but we do not pursue this direction in this paper.

C. Greedy Adaptive Collaborative Filtering

In this section we analyze the $H = 0$ case. Let $\mathcal{N}_m(t)$ be the set of items that have not been recommended to user m until time $(t-1)$ and thus are available for recommendation. The recommendation N_t can depend on all the past information, that is, on S_{t-1}. In the greedy case, we wish to choose N_t to maximize $E[Y_{M_t,N_t}]$, and the optimal choice is given by $N_t = \arg\max_{n \in \mathcal{N}_{M_t}(t)} E[Y_{m,n} | S_{t-1}, M_t = m]$. The following lemma provides an expression for the conditional expectation.

Lemma 1: For simplicity let $\ell = (1-\delta)/\delta$ and define

$$Q(\boldsymbol{J}^0, \boldsymbol{J}^1) := \prod_{k=1}^{K} \prod_{l=1}^{L} \left\{ \ell^{J^0_{k,l}} + \ell^{J^1_{k,l}} \right\},$$

where $J^i_{k,l}$ is the (k,l)th entry of the matrix $\boldsymbol{J}^i$. Let $\boldsymbol{E}_{k,l}$ denote the matrix with 1 at position (k,l) and zero otherwise. Then under the assumptions made in Section II,

$$\begin{aligned} \tilde{R}_0 &= E[Y_{m,n} | S_{t-1}, M_t = m] \\ &= \frac{\delta}{C_t} E_{\mathcal{T}} \left[Q(\boldsymbol{J}^0(t-1), \boldsymbol{J}^1(t-1) + \boldsymbol{E}_{U_m,V_n}) \right], \end{aligned} \tag{5}$$

where

$$C_t = E_{\mathcal{T}} \left[Q(\boldsymbol{J}^0(t-1), \boldsymbol{J}^1(t-1)) \right] \tag{6}$$

does not depend on n and the expectation $E_{\mathcal{T}}[\cdot]$ is only over the random assignment of types to users and items.

Proof: The proof is given in Appendix I.

Structure of the Greedy ACF: From Lemma 1, since δ, C_t do not depend on n, we see that the greedy policy recommends the solution of

$$\max_{n \in \mathcal{N}_m(t)} E_{\mathcal{T}} \left[Q(\boldsymbol{J}^0(t-1), \boldsymbol{J}^1(t-1) + \boldsymbol{E}_{U_m,V_n}) \right]. \tag{7}$$

Given the data up to time $(t-1)$ and a type assignment $\tau \in \mathcal{T}_{t-1}$, we can think of $Q(\boldsymbol{J}^0(t-1), \boldsymbol{J}^1(t-1))$ as a weight associated with the type assignment τ. From (7), we see that the greedy ACF assumes that its recommendation will lead to a response 1 and chooses the recommendation so that the resultant mean weight is maximum. In general, there is no simpler description of the greedy ACF. The difficulty stems from the expectation in (7), which involves a weighted sum over all type assignments. Given a single term in the summation, that is, given a type assignment τ, it is easy to check that the term is maximized by recommending an item of type l that maximizes the difference $J^1_{k,l} - J^0_{k,l}$, where user m has type k under τ. However, once we consider the entire summation involved in the expectation in (7), it is not possible to give a simple description for the solution of (7). We next consider the limiting case $\delta \to 0$, where we can give a simpler description of the greedy ACF.

Low noise regime: In the limit as $\delta \to 0$, the likelihood $\ell \to \infty$ and

$$\frac{\log(\ell^{a_1} + ... + \ell^{a_P})}{\log(l)} \to \max\{a_1, ..., a_P\} \tag{8}$$

for any $a_1, ..., a_P$. Therefore,

$$\frac{\log(Q(\boldsymbol{J}^0, \boldsymbol{J}^1))}{\log(\ell)} \to \sum_{k,l} \max\{J^0_{k,l}, J^1_{k,l}\}. \tag{9}$$

Further, since $\mathcal{T}_{t-1}$ is a finite set, the expectation in (7) has only finitely many terms, and we get

$$\begin{aligned} &\frac{\log\left(E_{\mathcal{T}}\left[Q(\boldsymbol{J}^0(t-1), \boldsymbol{J}^1(t-1) + \boldsymbol{E}_{U_m,V_n})\right]\right)}{\log(\ell)} \\ &\to \max_{\tau \in \mathcal{T}_t} \sum_{k,l} \max\{J^0_{k,l}(t-1), \tilde{J}^1_{k,l}(t-1)\} \end{aligned} \tag{10}$$

where $\tilde{J}^1_{k,l}(t-1) = J^1_{k,l}(t-1)+1$ if m, n have types k, l under τ and $\tilde{J}^1_{k,l}(t-1) = J^1_{k,l}(t-1)$ otherwise. This establishes the first part of Proposition 1 for $H = 0$.

To prove Part 2 of the Proposition, we need to show that the above limit is maximized by the ACF MT-LP. Towards this end, we note that the limit in (10) is bounded by $w_{\tau_*} +$

1. We next show that the ACF MT-LP attains this bounds. Suppose $M_t = m$ has already appeared up to time $(t-1)$ and under τ_* there is an item unseen by m such that the type of the item is popular amongst users of type U_m. Then the ACF MT-LP recommends this item and the same τ_* attains the maximum in the limit (10). If there is no unseen popular item, then the ACF MT-LP recommends a new item randomly. By assumption there is at least one item type that is popular amongst users of type U_m. We can extend τ_* by assigning the popular type to the new item and we see that this extended type attains the maximum in the limit (10). Finally, consider now the case when $M_t = m$ is a new user, who is recommended the most popular available item. The most popular item is popular amongst at least one type of users, say k, and we can extend τ_* to assign type k to user m. We see that this extended type attains the maximum in the limit (10).

D. One-Step Horizon

If we take $H = 1$, then we wish to maximize the reward $R_1 = E[Y_{M_t,N_t} + Y_{M_{t+1},N_{t+1}}]$. The standard dynamic program for this finite horizon problem uses the $H = 0$ case optimal ACF at time $(t+1)$, and at time t, the recommendation is to be made to maximize

$$\tilde{R}_1 := E[Y_{m,n} + Y_{M_{t+1},N^*_{t+1}} | S_{t-1}, M_t = m]$$

where N^*_{t+1} is chosen as per the $H = 0$ optimal scheme. For this case, we have the following lemma.

Lemma 2: For simplicity of notation, we denote $\boldsymbol{J}^i(t-1)$ by $\boldsymbol{J}^i$ in this lemma. Under the assumptions made in Section II,

$$\begin{aligned} & C_t \tilde{R}_1/\delta \\ & = E_{\mathcal{T}}\left[Q(\boldsymbol{J}^0, \boldsymbol{J}^1 + \boldsymbol{E}_{U_m,V_n})\right] \\ & + \frac{\delta}{K}\sum_{k=1}^{K} E_{\mathcal{T}}\left[Q(\boldsymbol{J}^0, \boldsymbol{J}^1 + \boldsymbol{E}_{U_m,V_n} + \boldsymbol{E}_{k,l_1})\right] \\ & + \frac{\delta}{K}\sum_{k=1}^{K} E_{\mathcal{T}}\left[Q(\boldsymbol{J}^0 + \boldsymbol{E}_{U_m,V_n}, \boldsymbol{J}^1 + \boldsymbol{E}_{k,l_0})\right] \end{aligned} \tag{11}$$

where l_1 (l_0) is the type of N^*_{t+1} when $Y_{m,n} = 1$ ($Y_{m,n} = 0$ respectively) and $U_{M_{t+1}} = k$.

Proof: The proof is given in Appendix II.

Low noise regime: Next we consider the limit as $\delta \to 0$ and show that Proposition 1 is true. Since the expectation in (11) is only over a finite number of type assignments, we see from (8) that the limit $\log(C_t\tilde{R}_1/\delta)/\log(\ell)$ equals the highest exponent amongst the terms in (11). To identify the term with the highest exponent, we note that the exponent of the $Q(\cdot)$ terms is given by (9). The exponent is at least $w_\tau = \sum_{k,l}\max\{J^0_{k,l}(t-1), J^1_{k,l}(t-1)\}$ and in the two time steps $t, t+1$, the exponent can at most increase by 2. We next show that by using the MT-LP ACF at time t, $t+1$, we can always increase the exponent by 2, and thereby establish its asymptotic optimality.

Suppose $M_t = m$ has not been seen before. In this case, the ACF MT-LP recommends the most popular item (say of type l under some type assignment in τ_*) amongst all users seen so far. If the feedback $Y_{m,n} = 1$, then we see that by extending the type assignment τ_* by assigning type k to the new user m, its weight (and the exponent) can be increased by 1. This extension of τ_* is optimal in terms of the weight (3) and for this type assignment items of type l continue to be most popular. Thus for a new user of type k arriving at time $(t+1)$, an item of the most popular type l is recommended, which increases the exponent by 1 more yielding the maximum increment of 2.

Suppose $M_t = m$ has appeared before. In this case, under the type assignments in τ_*, a type (say k) is already assigned to m. If there is an unseen item of type l such that $J^1_{k,l}(t-1) > J^0_{k,l}(t-1)$, then the ACF MT-LP can recommend the item and the exponent increases by 1 for the case $Y_{m,n} = 1$. An argument similar to above yields that the exponent can be further increased at time $(t+1)$. In the event that we are unable to find a popular unseen item (because m has seen all popular items), the ACF MT-LP recommends a new item randomly. The maximum type τ_* is then extended to this item by assigning it a type that is popular amongst type k users (which is guaranteed by the assumption). This increases the exponent by 1 and again by considering a new user of type k at time $(t+1)$ we get another increment of 1 in the exponent. This completes the proof for the $H = 1$ case.

IV. Conclusion

In this paper, we proposed a simple model to account for explicit feedback in collaborative filtering. While the problem of maximizing the long run average rating appears to be hard, it is possible to study the moving horizon approximation. In particular, we showed that the MT-LP ACF is optimal in the $\delta \to 0$ regime for any $H > 0$. In previous works ignoring explicit feedback, neighborhood methods based on the "cluster + local popularity" (see [2] and references therein) are quite popular. The MT-LP ACF has the same structure and is optimal even under explicit feedback on the finite horizon in the low noise regime. This is the main structural result of this paper. In a similar fashion, it is possible to identify ACFs optimal in the $\delta \to 1/2$ regime. These ACFs need to adapted for practical implementation and tested on real datasets. Other future directions include addressing the infinite horizon problem.

Appendix I
Proof of Lemma 1

To simplify notation, define the events

$$A_{t-1} := \{Y_{m_s,n_s} = y_s, 1 \le s \le (t-1)\},$$

$$B_{t-1} := \{M^{t-1} = m^{t-1}, N^{t-1} = n^{t-1}\},$$

where m^s denotes the sequence $m_1, ..., m_s$. Thus $S_{t-1} = \{A_{t-1}, B_{t-1}\}$. Since $Y_{m,n}$ is a binary random variable

$$\begin{aligned} & E[Y_{m,n}|S_{t-1}, M_t = m] \\ & = P(Y_{m,n} = 1|S_{t-1}, M_t = m) \\ & = \frac{P(Y_{m,n} = 1, A_{t-1}|B_{t-1}, M_t = m)}{P(A_{t-1}|B_{t-1}, M_t = m)}. \end{aligned} \tag{12}$$

First we calculate the numerator. For simplicity of notation, let $n_t = n$, $m_t = m$, and $y_t = 1$. If we condition on $\boldsymbol{X}$, then we know that the $Y_{m,n}$ depends only on $X_{m,n}$ and is independent of all other random variables. Therefore, conditioning on $\boldsymbol{X}$ and using the fact that the noise is i.i.d., the numerator equals,

$$\text{Num} = E\left[\prod_{s=1}^{t}(1-\delta)^{1(X_{m_s,n_s}=y_s)}\delta^{1(X_{m_s,n_s}\neq y_s)}\right],$$

where the indicator function $1(\cdot)$ is 1 if the argument is true and is 0 otherwise. We recall that $X_{m,n} = A_{U_m,V_n}$. So if $I_{k,l}$ denotes all the indices in $\{1, ..., t\}$ such that m_s, n_s have types k, l respectively, then we can write

$$\text{Num} = E\left[\prod_{k,l}\prod_{s\in I_{k,l}}(1-\delta)^{1(A_{k,l}=y_s)}\delta^{1(A_{k,l}\neq y_s)}\right].$$

But $\{A_{k,l}\}$ are i.i.d. Bernoulli(1/2), and hence, conditioning on the types of the users and items,

$$\begin{aligned} & \text{Num} \\ & = E_{\mathcal{T}}\left[\prod_{k,l}E\left[\prod_{s\in I_{k,l}}(1-\delta)^{1(A_{k,l}=y_s)}\delta^{1(A_{k,l}\neq y_s)}\Big|\mathcal{T}\right]\right] \\ & = E_{\mathcal{T}}\left[\prod_{k,l}\frac{1}{2}E\left[(1-\delta)^{J^0_{k,l}}\delta^{J^1_{k,l}} + (1-\delta)^{J^1_{k,l}}\delta^{J^0_{k,l}}\right]\right] \end{aligned}$$

where the expectation is only over the types of the users and items. Pulling out the factor $\delta^{J^0_{k,l}+J^1_{k,l}}$ and noting that $\sum_{k,l}(J^0_{k,l} + J^1_{k,l}) = t$, we get

$$\text{Num} = \frac{\delta^t}{2^{KL}}E_{\mathcal{T}}\left[\prod_{k,l}\left(\ell^{J^0_{k,l}} + \ell^{J^1_{k,l}}\right)\right].$$

The denominator of (12) has a similar expression and equations (5), (6) follow.

Appendix II
Proof of Lemma 2

Given Lemma 1, we only need to find an expression for $E[Y_{M_{t+1},N^*_{t+1}}|S_{t-1}, M_t = m]$, which equals $P(Y_{M_{t+1},N^*_{t+1}} = 1|S_{t-1}, M_t = m)$. First since M_{t+1} is independent of the past and is of type k with probability $1/K$,

$$\begin{aligned} & P(Y_{M_{t+1},N^*_{t+1}} = 1|S_{t-1}, M_t = m) \\ & = \frac{1}{K}\sum_{k=1}^{K}P(Y_{k,N^*_{t+1}} = 1|S_{t-1}, M_t = m, M_{t+1} = k). \end{aligned}$$

Given $M_t = m$, if we recommend $N_t = n$ at time t,

$$\begin{aligned} & P(Y_{k,N^*_{t+1}} = 1|S_{t-1}, M_t = m, M_{t+1} = k) \\ & = \sum_{i=0}^{1}P(Y_{m,n} = i|S_{t-1}, M_t = m) \\ & \qquad \times P(Y_{k,N^*_{t+1}} = 1|S^i_t, M_{t+1} = k) \end{aligned}$$

where $S^i_t = \{S_{t-1}, Y_{m,n} = i\}$. Now Lemma 1 provides an expression for all the terms in the above equation and Lemma 2 follows.

Acknowledgement

The author thanks Prof. Vivek Borkar, IIT Bombay, for helpful discussions.

References

[1] Aditya, S.T.; Dabeer, O.; Dey, B.K.; , “A Channel Coding Perspective of Collaborative Filtering,” *IEEE Transactions on Information Theory*, vol.57, no.4, pp.2327-2341, April 2011

[2] Barman, K.; Dabeer, O.; , “Analysis of a Collaborative Filter Based on Popularity Amongst Neighbors,” *IEEE Transactions on Information Theory*, vol.58, no.12, pp.7110-7134, Dec. 2012

[3] D. Bertsekas, “Dynamic programming and optimal control,” Athena Scientific, Belmont, MA, 1995

[4] D. Bertsekas, “Dynamic programming and suboptimal control: A survey from ADP to MPC,” *European Journal of Control*, 11.4-5: pp.310-334, 2005

[5] V. S. Borkar, “A convex analytic approach to markov decision processes,” *Probability Theory and Related Fields*, 78 (4). pp. 583-602, 1988

[6] Candes, E.J.; Recht, B.;, “Exact Matrix Completion via Convex Optimization,” *Foundations of Computational Mathematics*, vol.9, no.6, pp.717-772, 2009

[7] A. S. Das, M. Datar, A. Garg, and S. Rajaram, “Google news personalization: scalable online collaborative filtering,” *Proceedings of the 16th international conference on World Wide Web (WWW ’07)*, New York, USA, pp.271-280, 2007

[8] Keerthi, S. S., and Gilbert, E. G., 1988. “Optimal, Infinite Horizon Feedback Laws for a General Class of Constrained Discete Time Systems: Stability and Moving-Horizon Approximations, *J. Optimization Theory Appl.*, Vo. 57, pp. 265-293

[9] Keshavan, R.H.; Montanari, A.; Sewoong Oh; , “Matrix Completion From a Few Entries,” *IEEE Transactions on Information Theory*, vol.56, no.6, pp.2980-2998, June 2010

[10] Y. Koren, “Collaborative filtering with temporal dynamics,” Communications of the ACM, vol.53, no.4, pp.89-97, April 2010

[11] Linden, G.; Smith, B.; York, J.; , “Amazon.com recommendations: item-to-item collaborative filtering,” *IEEE Internet Computing*, vol.7, no.1, pp. 76- 80, Jan/Feb 2003

[12] Z. Lu, D. Agarwal, and I. S. Dhillon, “A spatio-temporal approach to collaborative filtering,” *Proceedings of the third ACM conference on Recommender systems (RecSys ’09)*, New York, USA, pp.13-20, 2009

[13] B. Recht, “A Simpler Approach to Matrix Completion,” *Journal of Machine Learning Research*, pp.3413-3430, December 2011

[14] M. Weimer, A. Karatzoglou, and A. Smola, “Adaptive collaborative filtering,” *Proceedings of the 2008 ACM conference on Recommender systems (RecSys ’08)*, pp.275-282, 2008