Learning Mallows Models with Pairwise Preferences

Tyler Lu TL@CS.TORONTO.EDU
Department of Computer Science, University of Toronto

Craig Boutilier CEBLY@CS.TORONTO.EDU
Department of Computer Science, University of Toronto

Abstract

Learning preference distributions is a key problem in many areas (e.g., recommender systems, IR, social choice). However, many existing methods require restrictive data models for evidence about user preferences. We relax these restrictions by considering as data arbitrary *pairwise comparisons*—the fundamental building blocks of ordinal rankings. We develop the first algorithms for learning Mallows models (and mixtures) with pairwise comparisons. At the heart is a new algorithm, the *generalized repeated insertion model (GRIM)*, for sampling from arbitrary ranking distributions. We develop approximate samplers that are exact for many important special cases—and have provable bounds with pairwise evidence—and derive algorithms for evaluating log-likelihood, learning Mallows mixtures, and non-parametric estimation. Experiments on large, real-world datasets show the effectiveness of our approach.

1. Introduction

With the abundance of preference data from search engines, review sites, etc., there is tremendous demand for learning detailed models of user preferences to support personalized recommendation, information retrieval, social choice, and other applications. Much work has focused on ordinal preference models and learning user or group *rankings* of items. We can distinguish two classes of models. First, we may wish to learn an underlying *objective* (or "correct") ranking from noisy data or noisy expressions of user preferences (e.g., as in web search, where user selection suggests relevance), a view adopted frequently in IR and "learning to rank" (Burges, 2010) and occasionally in social choice (Young, 1995). Second, we might assume that users have different *types* with inherently distinct preferences, and learn a population model that explains this diversity. Learning preference types (e.g., by segmenting or clustering the population) is key to effective personalization and preference elicitation; e.g., with a learned population preference distribution, choice data from a specific user allows inferences to be drawn about her preferences. We focus on the latter setting.

Considerable work in machine learning has exploited ranking models developed in the statistics and psychometrics literature, such as the Mallows model (Mallows, 1957), the Plackett-Luce model (Plackett, 1975; Luce, 1959), and others (Marden, 1995). However, most research to date provides methods for learning preference distributions using very restricted forms of evidence about individual user preferences, ranging from full rankings, to top-t/bottom-t items, to *partitioned preferences* (Lebanon & Mao, 2008). Missing from this list are arbitrary *pairwise comparisons* of the form "a is preferred to b." Such pairwise preferences form the building blocks of almost all reasonable evidence about preferences, and subsumes the most general evidential models proposed in the literature. Furthermore, preferences in this form naturally arise in active elicitation of user preferences and choice contexts (e.g., web search, product comparison, advertisement clicks), where a user selects one alternative over others (Louviere et al., 2000).

While learning with pairwise preferences is clearly of great importance, most believe that this problem is impractically difficult; so, for instance, the Mallows model is often shunned in favor of more inference-friendly models (e.g., Plackett-Luce, which accommodates more general, but still restrictive, preferences (Cheng et al., 2010; Guiver & Snelson, 2009)). To date, no methods have been proposed for learning from arbitrary paired preferences in any of the commonly used

Appearing in *Proceedings of the 28th International Conference on Machine Learning*, Bellevue, WA, USA, 2011.

ranking models in ML. We tackle this problem directly by developing techniques for learning Mallows models, and mixtures thereof, from pairwise preference data.

Our core contribution is the *generalized repeated insertion model (GRIM)*, a new method for sampling from *arbitrary ranking distributions*—including conditional Mallows—that generalizes the repeated insertion method for unconditional sampling of Mallows models (Doignon et al., 2004). Though we prove this problem is #P-hard in general, we derive another method, AMP, which efficiently and approximately samples from the conditional Mallows distribution. Moreover, we show that AMP is exact for important classes of evidence (including partitioned preferences), and that empirically it provides very close approximations given pairwise evidence. We use this sampler as the core of a Monte Carlo EM algorithm to learn Mallows mixtures as well as evaluating log likelihood. We also extend the non-parametric framework of Lebanon & Mao (2008) to handle unrestricted ordinal preference data. Experiments show our algorithms can effectively learn Mallows mixtures, with very reasonable running time, on datasets (e.g., Movielens) with hundreds of items and thousands of users.

2. Preliminaries

We assume a set of *items* $A = \{a_1, \ldots, a_m\}$ and n agents, or *users*, $N = \{1, \ldots, n\}$. Each agent ℓ has *preferences* over the set of items represented by a total ordering or *ranking* $\succ_\ell$ over A. We write $x \succ_\ell y$ to mean ℓ prefers x to y. Rankings can be represented as permutations of A. For any positive integer b, let $[b] = \{1, \ldots, b\}$. A bijection $\sigma : A \to [m]$ represents a ranking by mapping each item into its *rank*. Thus, for $i \in [m]$, $\sigma^{-1}(i)$ is the item with rank i. We write $\sigma = \sigma_1\sigma_2 \cdots \sigma_m$ for a ranking with i-th ranked item σ_i, and $\succ_\sigma$ for the induced preference relation. For any $X \subseteq A$, let $\sigma|_X$ denote the restriction of σ to items in X. Let $\mathbf{1}[\cdot]$ be the indicator function.

Generally, we do not have access to the complete preferences of agents, but only partial information about their rankings (e.g., based on choice behavior, query responses, etc.). We assume this data has a very general form: for each agent ℓ we have a set of *revealed pairwise preference comparisons* over A, or simply *preferences*: $v_\ell = \{x_1^\ell \succ_\ell y_1^\ell, \ldots, x_{k_\ell}^\ell \succ_\ell y_{k_\ell}^\ell\}$. Let $\mathsf{tc}(v_\ell)$ denote the transitive closure of v_ℓ. Since preferences are strict, $\mathsf{tc}(v_\ell)$ is a strict partial order on A. We assume each v_ℓ is *consistent*, i.e., $\mathsf{tc}(v_\ell)$ contains no cycles.[1] Preferences v_ℓ are *complete* iff $\mathsf{tc}(v)$ is a total order on A. Let $\Omega(v)$ be the *linear extensions* of v, i.e., the set of rankings consistent with v; $\Omega = \Omega(\emptyset)$ is the set of all $m!$ complete preferences. A collection $V = (v_1, \ldots, v_n)$ is a *(partial) preference profile*: this comprises our observed data.

Given $\sigma = \sigma_1\sigma_2 \cdots \sigma_m$ and preference v, define:

$$d(v, \sigma) = \sum_{i<j} \mathbf{1}[\sigma_j \succ \sigma_i \in \mathsf{tc}(v)]. \tag{1}$$

This measures dissimilarity between a preference set and a ranking using number of pairwise disagreements (i.e., those pairs in v that are misordered relative to σ). If v is a complete ranking σ', then $d(\sigma', \sigma)$ is the classic Kendall-tau metric on rankings.

Arbitrary sets v of paired comparisons model a wide range of realistic revealed preferences. Full rankings (Murphy & Martin, 2003) require $m-1$ paired comparisons ($a \succ b \succ c \ldots$); top-t preferences (Busse et al., 2007) need $m-1$ pairs ($t-1$ pairs to order the top t items, $m-t$ pairs to set the tth item above the remaining $m-t$); rankings of subsets $X \subseteq A$ (Guiver & Snelson, 2009; Cheng et al., 2010) are also representable. We also consider the following rich class:

Definition 1 (Lebanon & Mao 2008)**.** *A preference set v is a* partitioned preference *if A can be partitioned into subsets $A_1, \ldots, A_q$ s.t.: (a) for all $i < j \leq q$, if $x \in A_i$ and $y \in A_j$ then $x \succ_{\mathsf{tc}(v)} y$; and (b) for each $i \leq q$, items in A_i are incomparable under $\mathsf{tc}(v)$.*

Partitioned preferences are very general, subsuming the special cases above. However, they cannot represent many naturally occurring revealed preferences, including something as simple as a single paired comparison: $v = \{a \succ b\}$.

There are many distributional models of rankings—Marden (1995) provides a good overview. The two most popular in the learning community are the Mallows (1957) model and the Plackett-Luce model (Plackett, 1975; Luce, 1959). We focus on Mallows in this work, though we believe our methods can be extended to other models. The *Mallows ϕ-model* is parameterized by a modal or *reference ranking* σ and a *dispersion parameter* $\phi \in (0, 1]$. Let r be a ranking, then the Mallows model specifies:

$$P(r) = P(r \,|\, \sigma, \phi) = \frac{1}{Z}\phi^{d(r,\sigma)} , \tag{2}$$

where $Z = \sum_{r' \in \Omega} \phi^{d(r',\sigma)}$ and can be shown to equal $1\cdot(1+\phi)\cdot(1+\phi+\phi^2)\cdots(1+\cdots+\phi^{m-1})$. When $\phi = 1$ we obtain the uniform distribution, and as $\phi \to 0$ we get a distribution that concentrates all mass on σ. Sometimes the model is written as $P(r|\sigma, \lambda) = \frac{1}{Z}e^{-\lambda d(r,\sigma)}$,

[1] Many of the concepts developed in this paper can be applied to models where revealed preferences are noisy; we leave this to future research.

where $\lambda = -\ln \phi \geq 0$. To overcome the unimodal nature of Mallows models, mixture models have been proposed. A mixture with K components requires reference rankings $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_K)$, dispersion parameters $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_K)$, and mixing coefficients $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_K)$. EM for mixtures have been studied (Murphy & Martin, 2003; Busse et al., 2007), as well as inference in a Dirichlet process mixture context (Meila & Chen, 2010) (both limited to top-t data).

The *repeated insertion model (RIM)*, introduced by Doignon et al. (2004), is a generative process that gives rise to a family of distributions over rankings and provides a practical way to sample rankings from a Mallows model. Assume a reference ranking $\sigma = \sigma_1\sigma_2 \cdots \sigma_m$, and *insertion probabilities* p_{ij} for each $i \leq m, j \leq i$. RIM generates a new *output ranking* using the following process, proceeding in m steps. At Step 1, σ_1 is added to the output ranking. At Step 2, σ_2 is inserted above σ_1 with probability $p_{2,1}$ and inserted below with probability $p_{2,2} = 1 - p_{2,1}$. More generally, at the ith step, the output ranking will be an ordering of $\sigma_1, \ldots, \sigma_{i-1}$ and σ_i will be inserted at rank $j \leq i$ with probability p_{ij}. Critically, the insertion probabilities are *independent of the ordering of the previously inserted items.*

We can sample from a Mallows distribution using RIM with appropriate insertion probabilities.

Definition 2. *Let $\sigma = \sigma_1 \cdots \sigma_m$ be a reference ranking. Let an* insertion vector *be any positive integer vector $(j_1, \ldots, j_m)$ satisfying $j_i \leq i, \forall i \leq m$; and let I be the set of such insertion vectors. A* repeated insertion function *$\Phi_\sigma : I \to \Omega$ maps an insertion vector $(j_1, \ldots, j_m)$ into a ranking $\Phi_\sigma(j_1, \ldots, j_m)$ by placing each σ_i, in turn, into rank j_i, for all $i \leq m$.*

The definition is best illustrated with example. Consider insertion vector $(1, 1, 2, 3)$ and $\sigma = abcd$. Then $\Phi_\sigma(1, 1, 2, 3) = bcda$ because: we first insert a into rank 1; we then insert b into rank 1, shifting a down to get partial ranking ba; we then insert c into rank 2, leaving b but moving a down, giving bca; finally, we insert d at rank 3, giving $bcda$. Given reference ranking σ, there is a one-to-one correspondence between rankings and insertion vectors. Hence, sampling by RIM can be described as: draw an insertion vector $(j_1, \ldots, j_m) \in I$ at random, where each $j_i \leq i$ is drawn independently with probability p_{ij_i}—note that $\sum_{j=1}^{i} p_{ij_i} = 1$, for all i—and return ranking $\Phi_\sigma(j_1, \ldots, j_m)$.

Theorem 3 (Doignon et al. 2004)**.** *By setting $p_{ij} = \phi^{i-j}/(1+\phi+\cdots+\phi^{i-1})$ for $j \leq i \leq m$, RIM induces the same distribution over rankings as the Mallows model.*

Thus RIM offers a simple, useful way to sample rankings from the Mallows distribution.[2]

3. Generalized Repeated Insertion

While RIM provides a powerful tool for sampling from Mallows models (and by extension, Mallows mixtures), it samples unconditionally, without (direct) conditioning on evidence. We now proceed to generalize RIM by permitting conditioning at each insertion step. Our *generalized repeated insertion model (GRIM)* can sample from *arbitrary* rank distributions.

3.1. Sampling from Arbitrary Distributions

Rather than focus on conditional Mallows distribution given evidence about agent preferences, we present GRIM abstractly as a means of sampling from *any* distribution over rankings. We rely on the simple insight that the chain rule allows us to represent any distribution over rankings in a concise way, as long as we admit dependencies in our insertion probabilities: specifically, the insertion probabilities for any item σ_i in the reference ranking must be conditioned on the ordering of the previously inserted items $(\sigma_1, \ldots, \sigma_{i-1})$.

Let Q be any distribution over rankings and σ an (arbitrary) reference ranking. Recall that we can (uniquely) represent any ranking $r \in \Omega$ using σ and an insertion vector $\mathbf{j}^r = (j_1^r, \ldots, j_m^r) \in I$, where $r = \Phi_\sigma(\mathbf{j}^r)$. Thus Q can be represented by a distribution Q' over I: $Q'(\mathbf{j}^r) = Q(r)$. Similarly, for $k < m$, any partial ranking $r[k] = (r_1, \ldots, r_k)$ of the items $\{\sigma_1, \ldots, \sigma_k\}$, can be represented by a partial insertion vector $\mathbf{j}[k] = (j_1^r, \ldots, j_k^r)$. Letting $Q(r[k]) = \sum\{Q(r) : r_1 \succ r_2 \succ \cdots \succ r_k\}$, and $Q'(\mathbf{j}[k]) = \sum\{Q'(\mathbf{j}') : \mathbf{j}'[k] = \mathbf{j}[k]\}$, we have $Q'(\mathbf{j}[k]) = Q(r[k])$. Define *conditional insertion probabilities*

$$p_{ij \mid \mathbf{j}[i-1]} = Q'(j_i = j \mid \mathbf{j}[i-1]). \qquad (3)$$

This denotes the probability with which the ith item σ_i in the reference ranking is inserted at position $j \leq i$, conditioned on the specific insertions $(r_1, \ldots, r_{i-1})$ of all previous items. By the chain rule, we have

$$Q'(\mathbf{j}) = Q'(j_m|\mathbf{j}[m-1])Q'(j_{m-1}|\mathbf{j}[m-2]) \cdots Q'(\mathbf{j}[1]).$$

Suppose we run RIM with conditional insertion probabilities $p_{ij|\mathbf{j}[i-1]}$ defined above; that is, we draw random insertion vectors $\mathbf{j}$ by sampling j_1 through j_m, in turn, but each conditioned on the previously sampled components. The chain rule ensures that the re-

[2] RIM can also be used to sample from variants of the Mallows model, e.g., those using *weighted Kendall-tau* distance; we give details in a longer version of the paper.

sulting insertion vector is sampled from the distribution Q'. Hence the induced distribution over rankings $r = \Phi_\sigma(\mathbf{j})$ is Q. We call the aforementioned procedure the *generalized repeated insertion model (GRIM)*. Based on the arguments above, we have:

Theorem 4. *Let Q be a ranking distribution and σ a reference ranking. For any $r \in \Omega$, with insertion vector $\mathbf{j}^r$ (i.e., $r = \Phi_\sigma(\mathbf{j}^r)$), GRIM, using the insertion probabilities in Eq. 3, generates insertion vector $\mathbf{j}^r$ with probability $Q'(\mathbf{j}^r) = Q(r)$.*

Example 1. We illustrate GRIM using a simple example, sampling from a (conditional) Mallows model over $A = \{a, b, c\}$, with dispersion ϕ, given evidence $v = \{a \succ c\}$. The table illustrates the process:

Insert a, b		Insert c given ab		Insert c given ba	
r	Insrt. Prob.	r	Insrt. Prob.	r	Insrt. Prob.
a	$P(j_a = 1) = 1$	cab	$P(j_c = 1) = 0$	cba	$P(j_c = 1) = 0$
ab	$P(j_b = 1) = \frac{1}{1+\phi}$	acb	$P(j_c = 2) = \frac{\phi}{1+\phi}$	bca	$P(j_c = 2) = 0$
ba	$P(j_b = 2) = \frac{\phi}{1+\phi}$	abc	$P(j_c = 3) = \frac{1}{1+\phi}$	bac	$P(j_c = 3) = 1$

The resulting ranking distribution Q is given by the product of the conditional insertion probabilities: $Q(abc) = 1/(1+\phi)^2$; $Q(acb) = \phi/(1+\phi)^2$; and $Q(bac) = \phi/(1+\phi)$. As required, $Q(r) = 0$ iff r is inconsistent with evidence v.

3.2. Sampling a Mallows Posterior

While GRIM allows sampling from arbitrary distributions over rankings, as presented above it is largely a theoretical device, since it requires inference to compute the required conditional probabilities. To sample from a Mallows posterior, given arbitrary pairwise comparisons v, we show how to compute the required terms. The Mallows posterior is given by:

$$P_v(r) = P(r \mid v) = \frac{\phi^{d(r,\sigma)}}{\sum_{r' \in \Omega(v)} \phi^{d(r',\sigma)}} \mathbf{1}[r \in \Omega(v)], \qquad (4)$$

which requires summing over an intractable number of rankings to compute the normalization constant.

We could use RIM for rejection sampling: sample unconditional insertion ranks, and reject a ranking at any stage if it is inconsistent with v. However, this is impractical because of the high probability of rejection. Instead we use GRIM. The main obstacle is computing the insertion probability of a specific item given the insertion positions of previous items in Eq. 3 when Q' (more precisely, the corresponding Q) is the Mallows posterior. Indeed, this is #P-hard even with a uniform distribution over $\Omega(v)$:

Proposition 5. *Given v, a reference ordering σ, a partial ranking $r_1 \cdots r_{i-1}$ over $\{\sigma_1, \ldots \sigma_{i-1}\}$, and $j \in \{1, \ldots, i\}$, computing the probability of inserting σ_i at rank j w.r.t. the uniform Mallows posterior P (i.e., computing $P(r) \propto \mathbf{1}[r \in \Omega(v)]$) is #P-hard.*

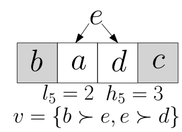

Fig. 1: Valid insertion ranks for e are $\{l_5, \ldots, h_5\} = \{2, 3\}$ given previous insertions and constraints v.

This suggests it is hard to sample exactly, and that computing the normalization constant in a Mallows posterior is difficult. Nevertheless we develop an approximate sampler AMP that is very efficient to compute. While it can perform poorly in the worst-case, we will see that, empirically, it produces excellent posterior approximations.[3]

AMP uses the same intuitions as illustrated in Example 1, where we use the (unconditional) insertion probabilities used by RIM, but subject to constraints imposed by v. At each step, the item being inserted can only be placed in positions that do not contradict $\mathsf{tc}(v)$. We can show that the valid insertion positions for any item, given v, form a contiguous "region" of the ranking (see Fig. 1 for an illustration).

Proposition 6. *Let insertion of $\sigma_1, \ldots, \sigma_{i-1}$ give a ranking $r_1 \cdots r_{i-1}$ consistent with $\mathsf{tc}(v)$. Let $L_i = \{i' \le i | r_{i'} \succ_{\mathsf{tc}(v)} \sigma_i\}$ and $H_i = \{i' \le i | r_{i'} \prec_{\mathsf{tc}(v)} \sigma_i\}$. Then inserting σ_i at rank j is consistent with $\mathsf{tc}(v)$ area* iff *$j \in \{l_i, l_i + 1, \ldots, h_i - 1, h_i\}$, where*

$$l_i = \begin{cases} 1 & \textit{if } L_i = \emptyset \\ \text{argmax}\, L_i + 1 & \textit{otherwise} \end{cases} \qquad (5)$$

$$h_i = \begin{cases} i & \textit{if } H_i = \emptyset \\ \text{argmin}\, H_i & \textit{otherwise} \end{cases} \qquad (6)$$

Prop. 6 immediately suggests a modification of the GRIM algorithm, AMP, for approximate sampling of the Mallows posterior: First initialize ranking r with σ_1 at rank 1. Then for $i = 2 \ldots m$, compute l_i, h_i and insert σ_i at rank $j \in \{l_i, \ldots, h_i\}$ with probability proportional to ϕ^{i-j}.

AMP induces a sampling distribution $\hat{P}_v$ that does not match the posterior P_v exactly: indeed the KL-divergence between the two can be severe, as the following example shows. Let $A = \{a_1, \ldots a_m\}$ and $v = a_2 \succ a_3 \succ \cdots \succ a_m$. Let P be the uniform Mallows prior ($\phi = 1$) with $\sigma = a_1 \cdots a_m$. There are m rankings in $\Omega(v)$, one r_i for each placement of a_1. The true Mallows posterior P_v is uniform over $\Omega(v)$. But AMP induces an approximation with $\hat{P}_v(r_i) = 2^{-i}$ for $i \le m-1$ and $\hat{P}_v(r_m) = 2^{-m-1}$. The KL-divergence of P_v and $\hat{P}_v$ is $(m-1)/2 + (1 - 2/m)\log_2 m - (1 + 1/m)$.

[3]We can also bound approximation quality theoretically. Further results, and proofs of all results, can be found in a longer version of the paper.

While AMP can perform poorly in the worst-case, it does very well in practice (see Sec. 5). We can also prove interesting properties, and provide theoretical guarantees of exact sampling in important special cases. First, it isn't hard to show that AMP will always produce a ranking (insertion positions always exist given any consistent v). Furthermore:

Proposition 7. *The support of distribution $\hat{P}_v$ induced by* AMP *is $\Omega(v)$ (i.e., identical to that of the Mallows posterior, Eq. 4).*

Proposition 8. *For any $r \in \Omega(v)$,* AMP *outputs r with probability:*

$$\hat{P}_v(r) = \frac{\phi^{d(r,\sigma)}}{\prod_{i=1}^{m}(\phi^{i-h_i} + \phi^{i-h_i+1} + \cdots + \phi^{i-l_i})}. \quad (7)$$

Using this result we can show that if v lies in the class of partitioned preferences, AMP's induced distribution is exactly the Mallows posterior:

Proposition 9. *If v is partitioned, the distribution $\hat{P}_v$ induced by* AMP *is the Mallows posterior P_v.*

While AMP may have (theoretically) poor worst-case performance, we can develop a statistically sound sampler MMP by using AMP to propose new rankings for the Metropolis algorithm. With Eq. 7, we can derive the acceptance ratio for Metropolis. At step $t+1$ of Metropolis, let $r^{(t)}$ be the previous sampled ranking. Ranking r, proposed by AMP independently of $r^{(t)}$, will be accepted as $r^{(t+1)}$ with probability

$$\min\left(1, \prod_{i=1}^{m} \begin{cases} \frac{h_i - l_i + 1}{h_i^t - l_i^t + 1} & \text{if } \phi = 1 \\ \frac{\phi^{h_i^t - h_i}(1-\phi^{h_i - l_i + 1})}{1-\phi^{h_i^t - l_i^t + 1}} & \text{otherwise} \end{cases}\right), \quad (8)$$

where the l_i's and h_i's are as in Eq. 5 and 6, respectively (defined w.r.t. r; and l_i^t and h_i^t are defined similarly, but w.r.t $r^{(t)}$). Prop. 7 helps show:

Theorem 10. *The Markov chain as defined in* MMP *is ergodic on the class of states $\Omega(v)$.*

3.3. Sampling a Mallows Mixture Posterior

Extending the GRIM, AMP and MMP algorithms to sampling from a mixture of Mallows models is straightforward. There is relatively little work on probabilistic models of partial rankings, and to the best of our knowledge, no proposed generative models for arbitrary sets of consistent paired comparisons. We first describe such a model before extending our algorithms to sample from a mixture of Mallows models.

We assume each agent has a latent preference ranking r, drawn from a Mallows mixture with parameters $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_K)$, $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_K)$, and $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_K)$. We use a component indicator vector $\mathbf{z} = (z_1, \ldots, z_K) \in \{0,1\}^K$, drawn from a multinomial with proportions $\boldsymbol{\pi}$, which specifies the mixture component from which an agent's ranking is drawn: if $z_k = 1$, r is sampled from the Mallows model with parameters σ_k, ϕ_k. Our observed data is a preference profile $V = (v_1, \ldots, v_n)$. Let $\mathbf{Z} = (\mathbf{z}_1, \ldots, \mathbf{z}_n)$ denote the latent indicators for each agent. To generate ℓ's preferences v_ℓ, we use a simple distribution, parameterized by $\alpha \in [0,1]$, that reflects a missing completely at random assumption.[4] We define $P(v|r,\alpha) = \alpha^{|v|}(1-\alpha)^{\binom{m}{2}-|v|}$ if $r \in \Omega(v)$; and $P(v|r,\alpha) = 0$ otherwise. One can view this as a process in which an α-coin is flipped for each pair of items to decide whether that pairwise comparison in r is revealed by v. Taken together, we have the joint distribution:
$P(v, r, \mathbf{z}|\boldsymbol{\pi}, \boldsymbol{\sigma}, \boldsymbol{\phi}, \alpha) = P(v|r,\alpha)P(r|\mathbf{z}, \boldsymbol{\sigma}, \boldsymbol{\phi})P(\mathbf{z}|\boldsymbol{\pi})$.

Now consider sampling from the mixture posterior, $P(r, \mathbf{z}|v, \boldsymbol{\pi}, \boldsymbol{\sigma}, \boldsymbol{\phi}) \propto P(v|r,\alpha)P(r|\mathbf{z}, \boldsymbol{\sigma}, \boldsymbol{\phi})P(\mathbf{z}|\boldsymbol{\pi})$. We use Gibbs sampling to alternate between r and $\mathbf{z}$, since the posterior does not factor in a way that permits us to draw samples exactly by sampling one variable, then conditionally sampling another. We initialize with some $\mathbf{z}^{(0)}$ and $r^{(0)}$, then repeatedly sample the conditionals of $\mathbf{z}$ given r and r given $\mathbf{z}$. For the t-th sample, $\mathbf{z}^{(t)}$ is drawn from a multinomial with K outcomes: $P(\mathbf{z} : z_k = 1|r^{(t-1)}) \propto \phi_k^{d(r^{(t-1)},\sigma_k)}\pi_k$. Then sample $r^{(t)}$ given $\mathbf{z}^t$, $P(r|\mathbf{z}^{(t)}, v) \propto P(v|r)P(r|\mathbf{z}^{(t)})P(\mathbf{z}^{(t)}) \propto \phi_k^{d(r,\sigma_k)}\mathbf{1}[r \in \Omega(v)]$, if $z_k^{(t)} = 1$. This is, of course, Mallows posterior sampling, so we use AMP or MMP.

4. EM for Learning Mallows Mixtures

Armed with the sampling algorithms derived from GRIM, we now turn to maximum likelihood learning of the parameters $\boldsymbol{\pi}$, $\boldsymbol{\sigma}$, and $\boldsymbol{\phi}$ of a Mallows mixture using EM. Before detailing our EM algorithm, we first consider the evaluation of log likelihood, which is used to select K or test convergence.

Evaluating Log Likelihood. Log likelihood $\mathcal{L}_\alpha(\boldsymbol{\pi}, \boldsymbol{\sigma}, \boldsymbol{\phi}|V)$ in our model can be written:

$$\sum_{\ell \in N} \ln \left[\sum_{k=1}^{K} \sum_{r_\ell \in \Omega(v_\ell)} \frac{\pi_k \phi_k^{d(r_\ell,\sigma_k)}}{Z_k}\right] + \ln \alpha^{|v_\ell|}(1-\alpha)^{\binom{m}{2}-|v_\ell|},$$

where Z_k is the Mallows normalization constant. It is easy to derive the maximum likelihood estimate for α: $\alpha^* = \sum_{\ell \in N} 2|v_\ell|/(nm(m-1))$. So we ignore this additive constant, and focus on the first term in the sum, denoted $\mathcal{L}(\boldsymbol{\pi}, \boldsymbol{\sigma}, \boldsymbol{\phi}|V)$. Unfortunately, evaluating this term is provably hard:

[4] This won't be realistic in all settings, but serves as a useful starting point.

Theorem 11. *Given profile $V = (v_1, \ldots, v_n)$, computing the log likelihood $\mathcal{L}(\boldsymbol{\pi}, \boldsymbol{\sigma}, \boldsymbol{\phi}|V)$ is #P-hard.*

As a result we consider approximations. We might rewrite $\mathcal{L}(\boldsymbol{\pi}, \boldsymbol{\sigma}, \boldsymbol{\phi}|V)$ as $\sum_{\ell \in N} \ln \left[\sum_{k=1}^{K} \pi_k \, \mathbb{E}_{P(r|\sigma_k, \phi_k)} \mathbf{1}[r \in \Omega(v)] \right]$, and estimate the inner expectations by sampling from the Mallows model $P(r|\sigma_k, \phi_k)$. However, this can require exponential sample complexity in the worst case (e.g., if $K = 1$ and v is far from σ, i.e., $d(v, \sigma)$ is large, then a sample of exponential size is expected to ensure v is in the sample). But we can rewrite the summation inside the log as $\sum_{\ell \in N} \ln \left[\sum_{k=1}^{K} \frac{\pi_k}{Z_k} \sum_{r \in \Omega(v_\ell)} \phi_k^{d(r, \sigma_k)} \right]$, and evaluate $\sum_{r \in \Omega(v_\ell)} \phi_k^{d(r, \sigma_k)}$ via importance sampling: we generate samples using AMP, then empirically approximate

$$\sum_{r \in \Omega(v_\ell)} \phi_k^{d(r,\sigma_k)} = \mathop{\mathbb{E}}_{r \sim \hat{P}_{v_\ell}} \left[\frac{\phi^{d(r, \sigma_k)}}{\hat{P}_{v_\ell}(r|\sigma_k, \phi_k)} \right]. \tag{9}$$

We generate samples $r_{\ell k}^{(1)}, \ldots, r_{\ell k}^{(T)}$ with $\mathsf{AMP}(v_\ell, \sigma_k, \phi_k)$ for $\ell \leq n$ and $k \leq K$, then substitute $\hat{P}_v$ from Eq. 7 into Eq. 9. With some algebraic manipulation we obtain the estimate:

$$\sum_{\ell \in N} \ln \left[\frac{1}{T} \sum_{k=1}^{K} \sum_{t=1}^{T} \pi_k \right.$$
$$\left. \begin{cases} \frac{1}{m!} \prod_{i=1}^{m} (h_i^{(\ell k t)} - l_i^{(\ell k t)} + 1) & \text{if } \phi_k = 1, \\ \phi_k^{\sum_{i=1}^{m} i - h_i^{(\ell k t)}} \prod_{i=1}^{m} \frac{1 - \phi_k^{h_i^{(\ell k t)} - l_i^{(\ell k t)} + 1}}{1 - \phi_k^i} & \text{otherwise,} \end{cases} \right]$$

where $h_i^{(\ell k t)}$ and $l_i^{(\ell k t)}$ are defined in Eqs. 5 and 6 (defined w.r.t. $r_{\ell k}^{(t)}$, σ_k, v_ℓ).

EM for Mallows Mixtures Learning a Mallows mixture is challenging, since even evaluating log likelihood is #P-hard. But we exploit our posterior sampling methods to render EM tractable. We apply the EM approach of Neal & Hinton (1999) as follows (recall we needn't consider α): We initialize our parameters with values $\boldsymbol{\pi}^{\text{old}}$, $\boldsymbol{\sigma}^{\text{old}}$, and $\boldsymbol{\phi}^{\text{old}}$. For the E-step, instead of working directly with the intractable posterior $P(\mathbf{z}_\ell, r_\ell | v_\ell, \boldsymbol{\pi}^{\text{old}}, \boldsymbol{\sigma}^{\text{old}}, \boldsymbol{\phi}^{\text{old}})$, we use GRIM-based Gibbs sampling (see Sec. 3.3), to obtain samples $(\mathbf{z}_\ell^{(t)}, r_\ell^{(t)})_{t=1}^{T}$, $\ell \in N$. In the M-step, we find a (local) maximum, $\boldsymbol{\pi}^{\text{new}}, \boldsymbol{\sigma}^{\text{new}}, \boldsymbol{\phi}^{\text{new}}$, of the empirical expectation:

$$\mathop{\text{argmax}}_{\boldsymbol{\pi}, \boldsymbol{\sigma}, \boldsymbol{\phi}} \sum_{\ell \in N} \frac{1}{T} \sum_{t=1}^{T} \ln P(v_\ell, r_\ell^{(t)}, \mathbf{z}_\ell^{(t)} | \boldsymbol{\pi}, \boldsymbol{\sigma}, \boldsymbol{\phi}). \tag{10}$$

If we were to fully maximize each parameter in the order $(\boldsymbol{\pi}, \boldsymbol{\sigma}, \boldsymbol{\phi})$ we would obtain a global maximum. Of course, exact optimization is intractable, so we approximate the components of the M-step. Abusing notation, let indicator vector $\mathbf{z}_\ell^{(t)}$ denote the mixture component to which the t-th sample of ℓ belongs. We partition all agents' samples into such classes: let $S_k = (\rho_{k1}, \ldots, \rho_{kj_k})$ be the sub-sample of rankings $r_\ell^{(t)}$ that belong in the k-th component, i.e., where $\mathbf{z}_\ell^{(t)} = k$. Note that $j_1 + \cdots + j_K = nT$. We can rewrite the M-step objective as: $\frac{1}{T} \sum_{k=1}^{K} \sum_{i=1}^{j_k} \ln P(v_{\ell(k,i)} | \rho_{ki}) P(\rho_{ki} | \sigma_k, \phi_k) P(k | \pi_k)$, where $\ell(k, i)$ is the agent for sample $\rho_{k,i}$. We ignore $\ln P(v_{\ell(k,i)} | \rho_{ki})$, which only impacts α; and we know $\rho_{ki} \in \Omega(v_{\ell(k,i)})$. Thus, we rewrite the objective as:

$$\sum_{k=1}^{K} \sum_{i=1}^{j_k} \ln \pi_k + d(\rho_{ki}, \sigma_k) \ln \phi_k - \sum_{w=1}^{m} \ln \frac{1 - \phi_k^w}{1 - \phi_k}. \tag{11}$$

Optimizing $\boldsymbol{\pi}$. Applying Lagrange multipliers yields:

$$\pi_k = j_k / (nT), \; \forall k \leq K. \tag{12}$$

Optimizing $\boldsymbol{\sigma}$. The only term involving $\boldsymbol{\sigma}$ in Eq. 11 is $\sum_{k=1}^{K} \sum_{i=1}^{j_k} d(\rho_{ki}, \sigma_k) \ln \phi_k$. Since $\ln \phi_k$ is a negative scaling factor, and we can optimize the σ_k independently, we obtain:

$$\sigma_k^* = \mathop{\text{argmin}}_{\sigma_k} \sum_{i=1}^{j_k} d(\rho_{ki}, \sigma_k). \tag{13}$$

Optimizing σ_k requires computing *Kemeny consensus* of the rankings in S_k, an NP-hard problem. Drawing on the notion of *local Kemenization* (Dwork et al., 2001), we instead compute a locally optimal σ_k, where swapping two adjacent items in σ_k cannot reduce the sum of distances in the Kemeny objective (details are included in a longer version of the paper).

Optimizing $\boldsymbol{\phi}$. When optimizing $\boldsymbol{\phi}$ in Eq. 11, the objective decomposes into a sum that permits independent optimization of each ϕ_k. Exact optimization of ϕ_k is difficult; however, we can use gradient ascent with $\frac{\partial \text{ (Eq. 11)}}{\partial \phi_k} = \frac{d(S_k, \sigma_k)}{\phi_k} - j_k \sum_{i=1}^{m} \frac{[(i-1)\phi_k - i]\phi_k^{i-1} + 1}{(1 - \phi_k^i)(1 - \phi_k)}$, where $d(S_k, \sigma_k) = \sum_{i=1}^{j_k} d(\rho_{ki}, \sigma_k)$.

Complexity of EM. One iteration of the E-step takes $O(nT_P T_{Gibbs}(T_{Metro} m^2 + Km \log m))$ time where T_{Metro} is number of Metropolis steps, T_{Gibbs} the number of Gibbs steps, and T_P is the posterior sample size for each v_ℓ. The M-step takes time $O(Km^2)$. Space complexity is $O(Km^2)$, dominated by the K *tournament graphs* used to compute Kemeny consensus.

Application to Non-Parametric Estimation Lebanon & Mao (2008) propose non-parametric estimators for Mallows models when observations form partitioned preferences. Indeed, they offer closed-form solutions by exploiting the existence of a closed-form for Mallows normalization with partitioned preferences. Unfortunately, with general *pairwise comparisons*, this normalization is intractable unless #P= P. But we can use AMP for approximate marginalization to support non-parametric estimation with general preference data. Define a joint distribution over $\Omega(v_\ell) \times \Omega$ by $q_\ell(s,r) = \frac{\phi^{d(r,s)}}{|\Omega(v_\ell)|Z_\phi}$, where Z_ϕ is the Mallows normalization constant. This corresponds to drawing a ranking s uniformly from $\Omega(v_\ell)$, then drawing r from a Mallows distribution with reference ranking s and dispersion ϕ. We extend the non-parametric estimator to paired comparisons using $p(v) = \frac{1}{n}\sum_{\ell \in N} q_\ell(s \in \Omega(v_\ell), r \in \Omega(v)) = \frac{1}{n}\sum_{\ell \in N, s \in \Omega(v_\ell), r \in \Omega(v)} \frac{\phi^{d(r,s)}}{|\Omega(v_\ell)|Z_\phi}$. We can approximate p using importance sampling: choose $\sigma \in \Omega(v_\ell)$ and sample rankings $s_\ell^{(1)}, \ldots, s_\ell^{(T)}$ from $\mathsf{AMP}(v_\ell, \sigma, \phi = 1)$, obtaining (see a longer version of the paper for derivations): $\hat{p}(v) = \frac{1}{nZ_\phi}\sum_{\ell \in N} \frac{\sum_{t=1}^{T} w_{\ell t} \sum_{r \in \Omega(v)} \phi^{d(r, s_\ell^{(t)})}}{\sum_{t=1}^{T} w_{\ell t}}$, where $w_{\ell t} = 1/\hat{P}_{v_\ell}(s_\ell^{(t)})$ is computed using Eq. 7. Evaluating $\sum_{r \in \Omega(v)} \phi^{d(r, s_\ell^{(t)})}$ is also intractable, but can be approximated using Eq. 9.

5. Experiments

We conducted experiments to measure the quality of the AMP algorithm both in isolation and in the context of log likelihood evaluation and EM. Full details of any experiments summarized below can be found in a longer version of the paper.

Sampling Quality. We first assess how well AMP approximates the true Mallows posterior P_v. We vary parameters m, ϕ and α, and fix a canonical reference ranking $\sigma = (1, 2, \cdots m)$. For each parameter setting, we generate 20 preferences v using our mixture model, and evaluated the KL-divergence of $\hat{P}_v$ and P_v (normalized by the entropy of P_v). In summary, our results show that AMP approximates the posterior very well, with average normalized KL error ranging from 1–5%, across the parameter ranges tested.

Log Likelihood and EM on Synthetic Data. We defer details to a longer version of the paper, but we note that our sampling methods provide excellent approximations of the log likelihood, and EM successfully reconstructs artificially generated mixtures, using pairwise preferences as data.

Sushi. The Sushi dataset consists of 5000 full rankings over 10 varieties of sushi indicating sushi preferences (Kamishima et al., 2005). We used 3500 preferences for training and 1500 for validation. We ran EM experiments by generating revealed paired comparisons for training with various probabilities α. To mitigate issues with local maxima, we ran EM ten times (more than is necessary) for each instance. Fig. 2 shows that that, even without full preferences, EM learns well even with only 30-50% of all paired comparisons, though it degrades significantly at 20%, in part because only 10 items are ranked (still performance at 20% is good when $K = 1, 2$). With $K = 6$ components, a good fit is found when training on full preferences: Fig. 2 shows the learned clusters (all with reasonably low dispersion), illustrating interesting patterns (e.g., fatty tuna is strongly preferred by all but one group; a strong correlation exists across groups in preference/dispreference for salmon roe and sea urchin, which are "atypical fish"; and cucumber roll is consistently dispreferred).

Movielens. We applied our EM algorithm to a subset of the Movielens dataset (see `www.grouplens.org`) to find "preference types" across users. We used the 200 (out of roughly 3900) most frequently rated movies, and used the ratings of the 5980 users (out of roughly 6000) who rated at least one of these. Integer ratings from 1 to 5 were converted to pairwise preferences in the obvious way (for ties, no preference was added to v). 3986 preferences were used for training and 1994 for validation. We ran EM with number of components $K = 1, \ldots, 20$; for each K we ran EM 20 times to mitigate the impact of local maxima (a lot more than necessary). For each K, we evaluated average log likelihood of the best run on the validation set to select K. Log likelihoods were approximated using our Monte Carlo estimates (with $K \cdot T = 120$). The C++ implementation of our algorithms gave EM wall clock times of 15–20 minutes (Intel Xeon dual-core, 3GHz), certainly practical for a data set of this size. Log likelihood results are shown in Fig. 2 as a function of the number of mixture components. This suggests that the best component sizes are $K = 10$ and $K = 5$ on the validation set. (The longer version of the paper details the top 20 movies in each component.)

6. Concluding Remarks

We have developed a set of algorithms to support the efficient and effective learning of ranking or preference distributions when observed data comprise a set of unrestricted pairwise comparisons of items. Given the fundamental nature of pairwise comparisons in revealed preference, our methods extend the reach of rank learning in a vital way. Our main technical contribution, the GRIM algorithm, allows sampling of ar-

$\pi_0 = 0.17$, $\phi_0 = 0.66$	$\pi_1 = 0.15$, $\phi_1 = 0.74$	$\pi_2 = 0.17$, $\phi_2 = 0.61$
fatty tuna	shrimp	sea urchin
salmon roe	sea eel	fatty tuna
tuna	squid	sea eel
sea eel	egg	salmon roe
tuna roll	fatty tuna	shrimp
shrimp	tuna	tuna
egg	tuna roll	squid
squid	cucumber roll	tuna roll
cucumber roll	salmon roe	egg
sea urchin	sea urchin	cucumber roll

$\pi_3 = 0.18$, $\phi_3 = 0.64$	$\pi_4 = 0.16$, $\phi_4 = 0.61$	$\pi_5 = 0.18$, $\phi_5 = 0.62$
fatty tuna	fatty tuna	fatty tuna
tuna	sea urchin	sea urchin
shrimp	tuna	salmon roe
tuna roll	salmon roe	shrimp
squid	sea eel	tuna
sea eel	tuna roll	squid
egg	shrimp	tuna roll
cucumber roll	squid	sea eel
salmon roe	egg	egg
sea urchin	cucumber roll	cucumber roll

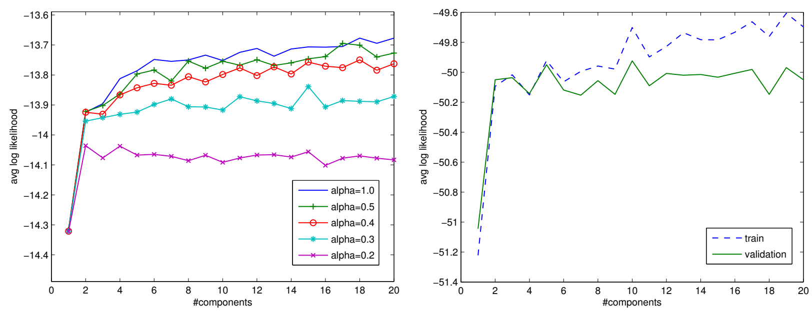

Fig. 2: The table shows the learned clusters for sushi with $\alpha = 1$. Left plot shows sushi avg. validation log likelihoods on various learned models (w.r.t. K) on various α. Right plot is for Movielens showing log likelihoods of various learned models (w.r.t. K).

bitrary distributions, including Mallows models conditioned on pairwise data. It supports a tractable approximation to the #P-hard problem of log likelihood evaluation of Mallows mixtures; and it forms the heart of an EM algorithm that was shown to be quite effective in our experiments. GRIM can also be used for non-parametric estimation.

We are pursuing a number of interesting directions, including various extensions and applications of the model developed here. Extensions include exploring other probabilistic models of incomplete preferences that employ different distributions over rankings such as Plackett-Luce or weighted Mallows, or that account for noisy comparison data from users. In another vein, we are interested in exploiting learned preference models of the type developed here for decision-theoretic tasks in social choice or personalized recommendation. Learned preferences can be leveraged in both active *preference elicitation* (e.g., in social choice or group decision making (Lu & Boutilier, 2011)), or in passive (purely observational) settings. It would also be interesting to apply GRIM to other posterior distributions such as energy models, and to compare it to different MCMC techniques like chain flipping.

References

Burges, C. From ranknet to lambdarank to lambdamart: An overview. TR-2010-82, Microsoft Research, 2010.

Busse, L. M., Orbanz, P. and Buhmann, J. M. Cluster analysis of heterogeneous rank data. *24th Intl. Conf. on Machine Learning (ICML-07)*, pp.113–120, 2007.

Cheng, W., Dembczynski, K., and Hüllermeier, E. Label ranking methods based on the Plackett-Luce model. *27th Intl. Conf. on Machine Learning*, pp.215–222, 2010.

Doignon, J., Pekec, A., and Regenwetter, M.. The repeated insertion model for rankings: Missing link between two subset choice models. *Psychometrika*, 69(1):33–54, 2004.

Dwork, C., Kumar, R., Naor, M., and Sivakumar, D. Rank aggregation methods for the web. *10th Intl. World Wide Web Conf. (WWW-01)*, pp.613–622, 2001.

Guiver, J. and Snelson, E.. Bayesian inference for Plackett-Luce ranking models. *26th Intl. Conf. on Machine Learning (ICML-09)*, pp.377–384, 2009.

Kamishima, T., Kazawa, H., and Akaho, S.. Supervised ordering: an empirical survey. *IEEE Intl. Conf. on Data Mining (ICDM-05)*, pp.673-676, 2005.

Lebanon, G. and Mao, Y.. Non-parametric modeling of partially ranked data. *J. Machine Learning Research*, 9: 2401–2429, 2008.

Louviere, J., Hensher, D., and Swait, J. *Stated Choice Methods: Analysis and Application.* Cambridge, 2000.

Lu, T. and Boutilier, C. Robust Approximation and Incremental Elicitation in Voting Protocols *22nd Intl. Joint Conf. on AI (IJCAI-11)*, to appear, 2011.

Luce, R. D. *Individual choice behavior: A theoretical analysis.* Wiley, 1959.

Mallows, C. L. Non-null ranking models. *Biometrika*, 44: 114–130, 1957.

Marden, J. I. *Analyzing and modeling rank data.* Chapman and Hall, 1995.

Meila, M. and Chen, H. Dirichlet process mixtures of generalized mallows models. *26th Conf. on Uncertainty in AI (UAI-10)*, pp. 358–367, 2010.

Murphy, T. B. and Martin, D. Mixtures of distance-based models for ranking data. *Computational Statistics and Data Analysis*, 41:645–655, 2003.

Neal, R. and Hinton, G. A view of the EM algorithm that justifies incremental, sparse, and other variants. In Jordan, M. (ed.), *Learning in Graphical Models*, pp.355–368. MIT Press, Cambridge, MA, 1999.

Plackett, R. The analysis of permutations. *Applied Statistics*, 24:193–202, 1975.

Young, P. Optimal voting rules. *J. Economic Perspectives*, 9:51–64, 1995.