A Note on the Boltzmann Distribution of the Linear Ordering Problem

Mikel de Velasco, Ieltzu Irazu, M Inés Fernandez

8 de octubre de 2015

Abstract

In this manuscript, we study the Boltzmann distribution associated to the linear ordering problem. Particularly, we demonstrate that this distribution is L-decomposable. Moreover, we investigate towards an efficient sampling method of the Boltzmann distribution.

1   The Linear Ordering Problem

As previously mentioned, in this manuscript we are particularly interested in the Boltzmann distribution associated to the Linear Ordering Problem (LOP) (Martí and Reinelt, 2011; Ceberio et al., 2014). Given a matrix $B = [b_{k,l}]_{n \times n}$ of numerical entries, the LOP consists of finding a simultaneous permutation σ of the rows and columns of B, such that the sum of the entries above the main diagonal is maximized (or equivalently, the sum of the entries below the main diagonal is minimized). The equation below formalizes the LOP function:

$$f(\sigma) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)} \tag{1}$$

where $\sigma(i)$ denotes the item of the row (and column) ranked at position i in the solution σ [1].

As a particular feature of the LOP, it is worth to remark that the contribution of an item $\sigma(i)$ to the fitness function f depends on the distribution of items in previous $\{1, \ldots, i-1\}$ and posterior $\{i+1, \ldots, n\}$ positions. However, the ordering of the items in each set does not influence the contribution of item $\sigma(i)$. This property of the LOP is of special relevance since the demonstration of the L-decomposability of the Boltzmann distribution is due to it.

2   The Boltzmann Distribution associated to the LOP

The *Boltzmann distribution* is an exponential probability distribution that assigns to a state (solution) a probability according to the energy (fitness) of the state. Formally, the Boltzmann distribution of a function $f(\sigma)$ is defined as

$$P(\sigma) = \frac{\exp(\beta f(\sigma))}{Z(\beta)} \tag{2}$$

where β denotes the Boltzmann constant, and $Z(\beta)$ is the partition function. When β equals 0, P becomes the uniform distribution. In contrast, when β is large, the probability concentrates around the optimum. In theory, if it were possible to sample efficiently the Boltzmann distribution for any β, optimisation would be an easy task (Muhlenbein et al., 2003).

The Boltzmann distribution associated to the LOP is given by

$$P(\sigma) = \frac{\exp(\beta \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)})}{Z(\beta)} \tag{3}$$

As far as we know, there is no close form expression for $Z(\beta)$, and thus, the only procedure is to calculate it by performing a sum over $\mathbb{S}_n$ as detailed in the following equation:

$$Z(\beta) = \sum_{\sigma \in S_n} \exp \left(\beta \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)} \right) \tag{4}$$

Note that the time complexity of the procedure above is $O(n!)$, and therefore it is only affordable for instances of $n < 15$.

---

[1] From now on, σ will denote any permutation in $\mathbb{S}_n$, and e will stand for the identity permutation $(1, 2, \ldots, n)$ of size n. In addition k and l will denote the items within a permutation σ, and i, j and z will be used to identify the positions of σ.

3 On the L-decomposability of the Boltzmann Distribution associated to the LOP

The *L-decomposability* is motivated by Luce's ranking postulate which states that the ranking process is assumed to be composed of a sequence of stages. The probability of the ensuing ranking is assumed to be the product of choice probabilities across the various stages. Formally, choice probabilities constitute a family of probability measures $\{P_B\}$ where B ranges over all subsets of the item labels $\{1, \ldots, n\}$. For each $i \in B$, $P_B(i)$ is the probability that item i is chosen as the most preferred item among those listed in B. Luce's ranking postulate states that for every permutation $\sigma = i_1 i_2 \ldots i_n$, the equation below is true.

$$P(\pi) = P_{\{i_1,\ldots,i_n\}}(i_1) \cdot P_{\{i_2,\ldots,i_n\}}(i_2) \cdots P_{\{i_{n-1},i_n\}}(i_{n-1}) \tag{5}$$

Any probability model $P(\sigma)$ which can be expressed in the form (5) for some collection of choice probabilities is said to be *L-decomposable* (Critchlow et al., 1991). This property implies that the choice probabilities at r-th stage depend only on the set of items remaining at that stage and not on the relative ordering of the $r-1$ items previously selected. Taking into account the comments on the contribution of the items in the LOP, we propose the following theorem.

Theorem 1. *The Boltzmann distribution associated to the linear ordering problem is L-decomposable.*

Proof. In order to verify that a probability distribution $P(\sigma)$ is L-decomposable, it is a necessary and sufficient condition that for each $r = 3, \ldots, n$ and for each $i_1, \ldots, i_r$, the conditional probability

$$P(\sigma(r) = i_r | \sigma(1) = i_1, \ldots, \sigma(r-1) = i_{r-1}) \tag{6}$$

is a symmetric function of $i_1, \ldots, i_{r-1}$ (Critchlow et al., 1991). Therefore, without loss of generality, by proving the equality below

$$P(\sigma(r) = i_r | \sigma(1) = i_1, \ldots, \sigma(r-1) = i_{r-1}) = P(\sigma(r) = i_r | \sigma(1) = i_{r-1}, \ldots, \sigma(r-1) = i_1) \tag{7}$$

we demonstrate that $P(\sigma)$ is L-decomposable.

From (7), by applying the Bayes' theorem, we obtain

$$\frac{P(\sigma(1) = i_1, \ldots, \sigma(r-1) = i_{r-1}, \sigma(r) = i_r)}{P(\sigma(1) = i_1, \ldots, \sigma(r-1) = i_{r-1})} = \frac{P(\sigma(1) = i_{r-1}, \ldots, \sigma(r-1) = i_1, \sigma(r) = i_r)}{P(\sigma(1) = i_{r-1}, \ldots, \sigma(r-1) = i_1)} \tag{8}$$

by replacing the Boltzmann distribution of the LOP (3), we rewrite (8) as follows

$$\frac{\sum_{\sigma \in C_r} \exp(\beta \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)})}{\sum_{\sigma \in C_{r-1}} \exp(\beta \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)})} = \frac{\sum_{\sigma \in D_r} \exp(\beta \sum_{i=1}^{n-1} \sum_{y=i+1}^{n} b_{\sigma(i),\sigma(j)})}{\sum_{\sigma \in D_{r-1}} \exp(\beta \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)})} \tag{9}$$

where C_r and C_{r-1} denote the set of permutations that agree with $i_1, \ldots, i_r$ and $i_1, \ldots, i_{r-1}$ orderings in the first r and $r-1$ positions, and D_r and D_{r-1} denote the set of permutations that agree with $i_{r-1}, \ldots, i_1, i_r$ and $i_{r-1}, \ldots, i_1$ orderings in the first r and $r-1$ positions. Note that partition functions $Z(\beta)$ are omitted by simple division.

By decomposing the fitness function of the left side in (9), we obtain

$$\frac{\sum_{\sigma \in C_r} \exp\left(\sum_{i=1}^{r-1} \sum_{j=i+1}^{r} b_{\sigma(i),\sigma(j)} + \sum_{i=1}^{r} \sum_{j=r+1}^{n} b_{\sigma(i),\sigma(j)} + \sum_{i=r+1}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)}\right)}{\sum_{\sigma \in C_{r-1}} \exp\left(\sum_{i=1}^{r-2} \sum_{j=i+1}^{r-1} b_{\sigma(i),\sigma(j)} + \sum_{i=1}^{r-1} \sum_{j=r}^{n} b_{\sigma(i),\sigma(j)} + \sum_{i=r}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)}\right)} \tag{10}$$

and extracting the common factor from the dividend and divisor that corresponds to the computation of the $i_1, \ldots, i_{r-1}$ items

$$\frac{\sum_{\sigma \in C_r} \left[\exp\left(\sum_{i=1}^{r-1} \sum_{j=i+1}^{r} b_{\sigma(i),\sigma(j)}\right) \times \exp\left(\sum_{i=1}^{r} \sum_{j=r+1}^{n} b_{\sigma(i),\sigma(j)} + \sum_{i=r+1}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)}\right)\right]}{\sum_{\sigma \in C_{r-1}} \left[\exp\left(\sum_{i=1}^{r-2} \sum_{j=i+1}^{r-1} b_{\sigma(i),\sigma(j)}\right) \times \exp\left(\sum_{i=1}^{r-1} \sum_{j=r}^{n} b_{\sigma(i),\sigma(j)} + \sum_{i=r}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)}\right)\right]} =$$

$$\frac{\sum_{\sigma \in C_r} \left[\exp\left(\sum_{i=1}^{r-2} \sum_{j=i+1}^{r-1} b_{\sigma(i),\sigma(j)} + \sum_{i=1}^{r-1} b_{\sigma(i),\sigma(r)}\right) \times \exp\left(\sum_{i=1}^{r} \sum_{j=r+1}^{n} b_{\sigma(i),\sigma(j)} + \sum_{i=r+1}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)}\right)\right]}{\sum_{\sigma \in C_{r-1}} \left[\exp\left(\sum_{i=1}^{r-2} \sum_{j=i+1}^{r-1} b_{\sigma(i),\sigma(j)}\right) \times \exp\left(\sum_{i=1}^{r-1} \sum_{j=r}^{n} b_{\sigma(i),\sigma(j)} + \sum_{i=r}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)}\right)\right]}$$

$$= \frac{\sum_{\sigma \in C_r} \left[\exp\left(\sum_{i=1}^{r-1} b_{\sigma(i),\sigma(r)}\right) \times \exp\left(\sum_{i=1}^{r} \sum_{j=r+1}^{n} b_{\sigma(i),\sigma(j)} + \sum_{i=r+1}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)}\right)\right]}{\sum_{\sigma \in C_{r-1}} \exp\left(\sum_{i=1}^{r-1} \sum_{j=r}^{n} b_{\sigma(i),\sigma(j)} + \sum_{i=r}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)}\right)} \tag{11}$$

Following a similar procedure, we rewrite the right side of (9),

$$\frac{\sum_{\sigma \in D_r} \left[\exp\left(\sum_{i=1}^{r-2} \sum_{j=i+1}^{r-1} b_{\sigma(i),\sigma(j)} + \sum_{i=1}^{r-1} b_{\sigma(i),\sigma(r)}\right) \times \exp\left(\sum_{i=1}^{r} \sum_{j=r+1}^{n} b_{\sigma(i),\sigma(j)} + \sum_{i=r+1}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)}\right)\right]}{\sum_{\sigma \in D_{r-1}} \left[\exp\left(\sum_{i=1}^{r-2} \sum_{j=i+1}^{r-1} b_{\sigma(i),\sigma(j)}\right) \times \exp\left(\sum_{i=1}^{r-1} \sum_{j=r}^{n} b_{\sigma(i),\sigma(j)} + \sum_{i=r}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)}\right)\right]}$$

and, as done previously, by extracting the common factor, we obtain

$$\frac{\sum_{\sigma \in D_r} \left[\exp\left(\sum_{i=1}^{r-1} b_{\sigma(i),\sigma(r)} \right) \times \exp\left(\sum_{i=1}^{r} \sum_{j=r+1}^{n} b_{\sigma(i),\sigma(j)} + \sum_{i=r+1}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)} \right) \right]}{\sum_{\sigma \in D_{r-1}} \exp\left(\sum_{i=1}^{r-1} \sum_{j=r}^{n} b_{\sigma(i),\sigma(j)} + \sum_{i=r}^{n-1} \sum_{j=i+1}^{n} b_{\sigma(i),\sigma(j)} \right)} \quad (12)$$

Since (11) and (12) are equal, we proved that Eq. 7 is true, and therefore, the Boltzmann distribution of the LOP is *L-decomposable.* □

Acknowledgements

Thank you every one!

References

Josu Ceberio, Alexander Mendiburu, and Jose A. Lozano. The Linear Ordering Problem Revisited. *European Journal of Operational Research*, 241(3):686–696, 2014.

Douglas E Critchlow, Michael A Fligner, and Joseph S Verducci. Probability models on rankings. *Journal of Mathematical Psychology*, 35(3):294 – 318, 1991. ISSN 0022-2496. doi: 10.1016/0022-2496(91)90050-4.

Rafael Martí and Gerhard Reinelt. *The linear ordering problem: exact and heuristic methods in combinatorial optimization*, volume 175. Springer, 2011.

Heinz Muhlenbein, Thilo Mahnig, and Fhg Ais. Evolutionary algorithms and the boltzmann distribution. In *Foundations of Genetic Algorithms 7*, pages 525–556. Morgan Kaufmann Publishers, 2003.