CS395T: Learning Theory; Fall 2011

Lecture 04 ; Date: 09/07/11

Lecturer: Ambuj Tiwari ; Scribe: Sharayu Moharir

Keywords: Probability Distributions, Rankings

Note: Scribed notes have only been lightly proofread.

Reference: Analyzing and Modeling Rank Data (Marden)

Let there be m objects $\{O_1, O_2, ..O_m\}$.

For instance, let $m = 3$. A valid ordering is

$$w = \begin{pmatrix} O_2 \\ O_3 \\ O_1 \end{pmatrix}.$$

An equivalent representation is the ranking vector

$$r = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix},$$

where the i^{th} entry in r denotes the rank of the i^{th} object.

Some popular models of probability distributions over rankings or permutations are:

1. Thurstonian Model

   Generate $z_1, z_2, ...z_m \in R$ according to some distribution.
   Output

$$w = \begin{pmatrix} O_{i_1} \\ O_{i_2} \\ . \\ . \\ O_{i_m} \end{pmatrix}$$

iff $z_{i_1} < z_{i_2} < ... < z_{i_m}$.
One specific example: Let $(z_1, z_2, ...z_m)$ be a multivariate normal.

2. Plackett-Luce Model

   For any subset $O' \subseteq O$ if there is a competition only among $o \in O'$, then

   $$P(o_i \in O' \quad is \quad first) = \frac{v(o_i)}{\sum_{o \in O'} v(o)}.$$

   Therefore,

   $$P\left(w = \begin{pmatrix} O_{i_1} \\ O_{i_2} \\ . \\ . \\ O_{i_m} \end{pmatrix}\right) = \frac{v(o_{i_1})}{\sum_{j=1}^{m} v(o_{i_j})} . \frac{v(o_{i_2})}{\sum_{j=2}^{m} v(o_{i_j})} .... \frac{v(o_{i_{m-1}})}{\sum_{j=m-1}^{m} v(o_{i_j})}.$$

3. Mallow's Model

   Given a reference ordering w^{ref},

   $$P(w) \propto exp\left(-\theta_i d\left(w, w^{ref}\right)\right),$$

   where d is some distance metric like L1 distance or Kendall tau distance which is defined as the minimum number of swaps required to convert w^{ref} to w. This model can be generalized to have multiple reference orderings namely $w^1, w^2, ..w^k$ such that

   $$P(w) \propto exp\left(-\sum_{i=1}^{k} \theta_i d(w, w^i)\right).$$

4. Babington Smith Model

   This Model is based on $\binom{m}{2}$ parameters.

   $$T_{O_i, O_j}(w) = 1(O_i < O_j).$$

   $$P(w) \propto exp\left(-\sum_{i<j} \theta_{i,j} T_{O_i, O_j}(w)\right).$$

5. (This one doesn't have a name)

   $$T_{O_i, j}(w) = 1(O_i \quad \textit{in position} \quad j).$$

   $$P(w) \propto exp\left(-\sum_{i} \theta_{O_i, j} j\right).$$

6. Fligner-Verducci Model

This model proceeds in $m-1$ steps. Given a starting configuration, in the i^{th} step, an integer n_i between 0 to $m-i+1$ is chosen according to some distribution. The element at the $(i+n_i)^{th}$ position is inserted at the i^{th} position and the algorithm proceeds to the next step.