CONDITIONAL LOGISTIC REGRESSION FOR BJJ/MMA

NOAH SIMON

1. Conditional Logit

Suppose competitors are numbered $j = 1, \ldots, m$; and we observe n bouts between pairs of competitors. Consider modeling

$$p\left(\text{winner} = j \middle| \text{ competitors are } j, k\right) \equiv p_{j,k}(j) = \frac{e^{\eta_j}}{e^{\eta_j} + e^{\eta_k}}$$

In this case our likelihood is

$$L\left(\eta\right) = \prod_{i=1}^{n} \left(\frac{e^{\eta_{j(i,w)}}}{e^{\eta_{j(i,w)}} + e^{\eta_{j(i,l)}}}\right)$$

where the indices of the two competitors in match i are $j(i,w)$ the winner, and $j(i,l)$ the loser (in this pair-of-competitor setup there is an equivalent logistic regression). The log-likelihood is now

$$\ell\left(\eta\right) = \sum_{i=1}^{n} \left[\eta_{j(i,w)} - \log\left(e^{\eta_{j(i,w)}} + e^{\eta_{j(i,l)}}\right)\right]$$

We regularize this and solve

$$\hat{\eta} \leftarrow \operatorname{argmin} -\ell(\eta) + \lambda \|\eta\|_2^2.$$

Regularization turns out to be *very* important here, as otherwise we end up with complete separation for some small subset of competitors with almost no matches. Calculating the gradient gives

$$\frac{\partial}{\partial \eta_k} \ell\left(\eta\right) = \sum_{i \in Win_k} \left[1 - \left(\frac{e^{\eta_{j(i,w)}}}{e^{\eta_{j(i,w)}} + e^{\eta_{j(i,l)}}}\right)\right] + \sum_{i \in Lose_k} \left[-\left(\frac{e^{\eta_{j(i,l)}}}{e^{\eta_{j(i,w)}} + e^{\eta_{j(i,l)}}}\right)\right]$$

This can be calculated efficiently by

---

(1) Set $g_j \leftarrow 0$ for $j = 1, \ldots, m$ (where m is total number of teams)

(2) Calculate

$$p_i \leftarrow p\left(\text{winner} = j(i,w) \middle| \text{ competitors are } j(i,w), j(i,l)\right) = \frac{e^{\eta_{j(i,w)}}}{e^{\eta_{j(i,w)}} + e^{\eta_{j(i,l)}}}$$

for each $i = 1, \ldots, n$

(3) for $i = 1, \ldots, n$ update $g_{j(i,w)}$ and $g_{j(i,l)}$ by

$$g_{j(i,w)} += 1 - p_i$$
$$g_{j(i,l)} += p_i - 1$$

and we can use gradient descent to find the minimizer. This idea can also be extended in neat ways to modeling win-probabilities via low rank matrix completion.