The Mid Report

1 Introduction

Many have proposed online updating algorithms for paired comparison experiments. These online algorithms are especially useful when the objects to be compared and the number of games are very large. For ranking of many sports, possibly the most prominent ranking system in use today is ELO. The ELO ranking system was originally invented by Arpad Elo (1903-1992), a Hungarian-born American physics professor, as an improved chess rating system and now it has been used successfully by a variety of leagues organized around two-player games, such as world football league, the US Chess Federation or the World Chess Federation, and a variety of others. Another updating algorithm is the Glicko system [?], developed by Mark E. Glickman, chairman of the US Chess Federation (USCF) ratings committee. To the best of our knowledge, the Glicko system is the first Bayesian online rating systen. It improves over Elo by incorporating the variability (called rating deviation) in skill parameter. For example, a player with a rating of 1000 and a rating deviation of 50 will have a real strength between 900 and 1100 with a chance of 95Also the rating change after a game depends on the rating deviation. Glicko uses the Bradley-Terry model [?] and is designed to update players' skills after a rating period. Though the ELO and Glicko ranking system have been successfully, they are designed for two-player games. In video games many of these leagues have game modes with more than two players per match. To support such games, recently Microsoft Research developed the TrueSkill ranking system [?] for Xbox. Like Glicko, the TrueSkill ranking system is a Bayesian rating system and characterizes skill also by two numbers: the average skill of the gamer (μ) and the degree of uncertainty in the gamer's skill (σ). The TrueSkill differs from Glicko in some ways. First of all, Glicko uses the Bradley-Terry model (which indeed can be derived by assuming that the performance follows a logistic distribution), while the TrueSkill uses normal model (derived by assuming that the performance follows a normal distribution). Secondly, TrueSkill uses a *draw margin* that depends on the game mode, as some games may easily result in draws but some may not. Thirdly, the Glicko system considers only two-player game. In the case of multiple players, the TrueSkill ranking system suggests to iterate two team update equations between all teams on neighboring ranks, that is, the 1st versus the 2nd team, the 2nd team versus the 3rd team and so on. This involves certain numerical Bayesian approximation methods.

Basically, the update rules come from first assuming the strength of each player follows some prior distribution, and then deriving some analytic approximations to the posterior mean and variance of this strength parameter after observing one or more game outputs. One reason for using the posterior mean is that it minimizes the risk with square loss; that is, it minimizes $Q(\delta) \equiv E((\theta - \delta)^2|\text{data})$.

2 Literature reivew

In multi-player games or even in teams, it is generally believed that the logistic model would make it computationally more expensive. In this project, we propose to use an approximate Bayesian method together with a generalization of the Bradley-Terry model [?] to obtain simple update formulas in cases where there are $k \geq 3$ players.

We first review the modeling of ranked data, and the relation between the logistic distributions and the Bradley-Terry model. Let R represents the random ordering of k objects, where each object has a continuous but unobserved random variable X_i $(i = 1, ..., k)$ associated with it. The observed ordering that player i_1 comes in first, player i_2 comes in second and so on is then determined by the X_i's:

$$R = (i_1, ..., i_k) \text{ if and only if } X_{i_1} > X_{i_2} > \cdots > X_{i_k}. \tag{1}$$

Thurstone [?] invented this model and proposed using the normal distribution. Other researchers continued to look at cases in which X_i's are not normal. For example, Luce [?] and Yellott [?] used Gumbel distribution. In fact, if $X_i \sim N(\theta_i, \sigma^2)$, then

$$P(X_1 > X_2) = \Phi(\frac{\theta_1 - \theta_2}{\sqrt{2}\sigma}); \tag{2}$$

if X_i follows Gumbel distribution with location parameter θ_i and scale parameter δ (i.e. $X_i \sim F_i(x) = \exp(-e^{-(x-\theta_i)/\delta})$), then $X_1 - X_2$ follows a logistic distribution and

$$P(X_1 > X_2) = \frac{e^{\theta_1/\delta}}{e^{\theta_1/\delta} + e^{\theta_2/\delta}}. \tag{3}$$

Taking $e^{\theta_i/\delta} = v_i$, the model (3) is exactly the Bradley-Terry model [?].

Another popular model for comparisons involving $k > 2$ players per match is a generalization of the Bradley-Terry model, termed the Plackett-Luce model by

Marden [?]. This model, motivated by a k-horse race, has the form

$$P(R = r; v) = \prod_{i=1}^{k} \frac{v_{r_i}}{v_{r_i} + v_{r_{i+1}} + \cdots + v_{r_k}}, \tag{4}$$

where v_i's are strength parameters. In the particular case of triple comparison, (4) reduces to the Pendergrass-Bradley model [?]. Recently, Hunter [?] fits the model (4) to estimate drivers' strengths based on 36 automobile races for the 2002 NASCAR season in the United States. Model (4) is closely related to Luce *choice* axiom [?]. Letting $A \subset B \subset \{1, ..., k\}$, the axiom states that if individual i has a positive probability of beating individual j, the model must satisfy

$$P_B(i \text{ wins}) = P_A(i \text{ wins}) P_B(\text{anything from } A \text{ wins}). \tag{5}$$

Luce [?] showed that (5) is equivalent to the following

$$P_B(i \text{ wins}) = \frac{v_i}{\sum_{j \in B} v_j}. \tag{6}$$

It is not hard to show that (4) is equivalent to (6). For a detailed account, see Marden [?] and Hunter [?].

Next we consider an Bayesian approximation method. Throughout let Φ_p and ϕ_p denote the cumulative distribution function (cdf) and probability density function (pdf) of a standard p-variate normal distribution, and abbreviate them as Φ and ϕ when $p = 1$. Write $\Phi_p h = \int h d\Phi_p$, where h is a function from $\Re^p$ to $\Re$. Let Z be a random variable with pdf $\zeta(z) = C\phi_p(z)f(z)$, where $C = (\Phi_p f)^{-1}$ is the normalizing constant. Let E denote the expectation with respect to Z. The equations (7) and (8) below play important roles in the present paper.

$$EZ = E(\frac{\nabla f(Z)}{f(Z)}), \tag{7}$$

$$E(Z_i Z_j) = \delta_{ij} + E[\frac{\nabla^2 f(Z)}{f(Z)}]_{ij}, \tag{8}$$

where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise.

3 Current progresses of the project

We use a simple two-player game to illustrate the application of above equations. Assume that players 1 and 2 have prior strengths $\theta_i \sim N(\mu_i, \sigma_i^2)$ and that player 1 beats player 2 in a game. Let Z_1 and θ_1 is related by

$$Z_i = \frac{\theta_i - \mu_i}{\sigma_i}. \tag{9}$$

Then, the posterior distribution of Z is

$$\zeta(z|1 \text{ beats } 2) \propto \phi_2(z)\Phi(\frac{\theta_1 - \theta_2 - \epsilon}{c});$$

and hence by (7) we have

$$\begin{aligned} E(Z_1|1 \text{ beats } 2) &= E(\frac{\partial f(Z_1)/\partial Z_1}{f(Z_1)}|1 \text{ beats } 2) \\ &= \frac{\sigma_1}{c}E\big(V(\frac{\theta_1 - \theta_2 - \epsilon}{c})|1 \text{ beats } 2\big). \end{aligned}$$

We propose to approximate the above expectation by replacing θ_i's with μ_i's; that is,

$$E(Z_1|1 \text{ beats } 2) \approx \frac{\sigma_1}{c}V(\frac{\mu_1 - \mu_2 - \epsilon}{c});$$

and then the update rule for θ_1 is obtained using the relation between θ_i and Z_i in (9). This leads to the TrueSkill's formula for a two-player game.

For the Plackett-Luce mode, we note that one feature of (4) is the lack of a corresponding variance parameter. Below we show how to incorporate a variance parameter in the Plackett-Luce model. The motivation is from the relation between the normal model (2) and the Bradley-Terry model (3). To begin, we reparametrize v_i in (3) as $e^{\theta_i/\tau}$ so that (3) can be written as

$$\frac{e^{(\theta_1-\theta_2)/\tau}}{1 + e^{(\theta_1-\theta_2)/\tau}}.$$

Since

$$F(x) = \frac{e^{x/\tau}}{(1 + e^{x/\tau})}$$

is the cumulative distribution function of a logistic distribution with mean 0 and variance $(\tau\pi/\sqrt{3})^2$, it can be approximated by the cumulative distribution function of a normal distribution with the same mean and variance. So,

$$\begin{aligned} (3) &= \frac{e^{(\theta_1-\theta_2)/\tau}}{1 + e^{(\theta_1-\theta_2)/\tau}} \\ &\approx \int_{-\infty}^{\theta_1} \frac{1}{\sqrt{2\pi}(\tau\pi)/\sqrt{3}} e^{-(t-\theta_2)^2/(2(\tau\pi/\sqrt{3})^2)} dt \\ &= \Phi\left(\frac{\theta_1 - \theta_2}{\tau\pi/\sqrt{3}}\right). \end{aligned} \tag{10}$$

By comparing (10) with (2), it suggests to take $\tau^2 \propto (\sigma_i^2 + \sigma_j^2)$; that is, incorporating the variance into the Plackett-Luce model by taking $v_i = e^{\theta_i/\tau}$ with $\tau^2 = \sigma_i^2 + \sigma_j^2$. When there are k players, we consider the model (4) with $v_i = e^{\theta_i/\tau}$, where $\tau^2 = \sum_{i=1}^k \sigma_i^2$.

Now we consider a Bayesian model in which a normal prior $N(\mu_i, \sigma_i^2)$ is imposed on the ith player's current skill parameter $\theta_i^{(t)}$, where μ_i and σ_i are known. To take into account the fact that players' skills may change over time, we follow similar ideas in Glicko and TrueSkill by assuming that

$$\theta_i^{(t+1)} = \theta_i^{(t)} + e_i^{(t)},$$

where $e_i^{(t)}$ is the amount player abilities change from game t to game $t+1$. We assume that $e_i^{(t)}$ follows $N(0, \beta^2)$ distribution; and hence the prior distribution on $\theta_i^{(t+1)}$ is $N(\mu_i, \sigma_i^2 + \beta^2)$. Together with the discussion in the previous paragraph, the likelihood of $\theta^{(t+1)}$ for game $t+1$ is

$$\begin{aligned} &P(R = r; \theta^{(t+1)}) \\ =&\prod_{i=1}^k \frac{e^{\theta_{r_i}^{(t+1)}/\tau}}{e^{\theta_{r_i}^{(t+1)}/\tau} + e^{\theta_{r_{i+1}}^{(t+1)}/\tau} + \cdots + e^{\theta_{r_k}^{(t+1)}/\tau}}, \end{aligned} \tag{11}$$

where

$$\tau^2 = \sum_{i=1}^k (\sigma_i^2 + \beta^2). \tag{12}$$

Now we can use Plackett-Luce model (11) and equations (7) and (8) to derive update rules. Suppose that k teams play against each other, where the ith team has n_i players. Assume that the jth player in the ith team has strength θ_{ij}, the prior distribution of θ_{ij} is $N(\mu_{ij}, \sigma_{ij}^2)$, and a team's skill is the sum of skills of its members. Let $r = (r(1), \ldots, r(k))$ denote the rank of the k teams, where ties are possible. Let

$$A_i = \{x : r(x) = i\} \text{ and } C_i = \{x : r(x) \geq i\}.$$

Define

$$p_{i,C_j} = \frac{e^{\theta_i/\tau}}{\sum_{s \in C_j} e^{\theta_s/\tau}}, \tag{13}$$

the probability that i is the winner among players in C_j, where we set $\tau^2 = \sum_i \sum_j (\sigma_{ij}^2 + \beta^2)$ following (12); define

$$\hat{p}_{i,C_j} = \frac{e^{\mu_i/\tau}}{\sum_{s \in C_j} e^{\mu_s/\tau}}, \tag{14}$$

an estimate of p_{i,C_j}. Now for $a = 1, \ldots, k$ we update the team skills by

$$\mu_a \leftarrow \mu_a + \Omega_a \tag{15}$$

$$\sigma_a^2 \leftarrow \sigma_a^2(1 - \Delta_a), \tag{16}$$

where Ω_a and Δ_a are to be determined later.