The Final Report

1 Introduction

Many online updating algorithms for paired comparison are useful when the numbers of teams/games are very large. For ranking sport teams, possibly the most prominent ranking system in use today is [1]. It has been used successfully by leagues organized around two-player games, such as world football league, the US Chess Federation (USCF), and a variety of others. [2] proposed the Glicko system, which incorporates the variability in parameter estimates. To begin, prior to a rating period, a player's skill (θ) is assumed to follow a Gaussian distribution which can be characterized by two numbers: the average skill of the player (μ) and the degree of uncertainty in the player's skill (σ). However, in video games a game often involves more than two players or teams. Recently Microsoft Research developed TrueSkill [3], a ranking system for Xbox Live. Similar to Glicko, TrueSkill is also a Bayesian ranking system using a Gaussian belief over a player's skill. In a two-team game the TrueSkill update rules are fairly simple, but for games with multiple teams and multiple players, the update rules are not possible to write down as they require an iterative procedure.

In this project, we describe a Bayesian approximation method to derive simple analytic update rules for online ranking of players from games with multiple teams and multiple players.

2 Approximation Techniques for Bayesian Inference

For Bayesians, both the observed data and the model parameters are considered random. Let D denote the observed data, and θ the unknown quantities of interest. The joint distribution of D and θ is

$$P(D, \theta) = P(D|\theta)P(\theta),$$

where $P(\theta)$ is the prior distribution and $P(D|\theta)$ the likelihood. The *posterior distribution* of θ given D is

$$P(\theta|D) = \frac{P(\theta, D)}{P(D)} = \frac{P(\theta, D)}{\int P(\theta, D)d\theta},$$

which is useful for estimation. The probability $P(D)$, called *evidence* or *marginal likelihood* of the data, is useful for model selection. Both $P(\theta|D)$ and $P(D)$ are major objects of Bayesian inference.

The integrations involved in Bayesian inference are usually intractable and approximations are often needed. The approximation techniques are divided into deterministic and nondeterministic methods. The nondeterministic method refers to the Monte Carlo integration such as Markov Chain Monte Carlo (MCMC) methods. However, when it comes to sequential updating with new data, the MCMC methods may not be computationally feasible, the reason being that it does not make use of the analytic from the previous data.

Popular deterministic approaches include Laplace method, variational Bayes, expectation propagation, among others. The Laplace method is a technique for approximating integrals:

$$\int e^{nf(\boldsymbol{x})} d\boldsymbol{x} \approx \left(\frac{2\pi}{n}\right)^{\frac{k}{2}} |-\nabla^2 f(\boldsymbol{x}_0)|^{-\frac{1}{2}} e^{nf(\boldsymbol{x}_0)},$$

where $\boldsymbol{x}$ is k-dimensional, n is a large number, $f : R^k \to R$ is twice differentiable with a unique global maximum at $\boldsymbol{x}_0$, and $|\cdot|$ is the determinant of a matrix. By writing $P(\theta, D) = \exp(\log P(\theta, D))$, one can approximate the integral $\int P(\theta, D) d\theta$.

The variational Bayes method is a family of techniques for approximating these intractable integrals. The idea is to construct a lower bound on the marginal likelihood and then try to optimize this bound. The Expectation Propagation algorithm [5] is an iterative approach to approximate posterior distributions. It tries to minimize Kullback-Leibler divergence between the true posterior and the approximated distribution. The TrueSkill system [3] is based on this algorithm.

Now we review an identity for integrals in Lemma 2.1 below, which forms the basis of our approximation method. Some definitions are needed. A function $f : R^k \to R$ is called almost differentiable if there exists a function $\nabla f : R^k \to R^k$ such that

$$f(\boldsymbol{z} + \boldsymbol{y}) - f(\boldsymbol{z}) = \int_0^1 \boldsymbol{y}^T \nabla f(\boldsymbol{z} + t\boldsymbol{y}) dt \tag{1}$$

for $\boldsymbol{z}, \boldsymbol{y} \in R^k$. Of course, a continuously differentiable function f is almost differentiable with ∇f equal to the gradient, and (1) is the indefinite integral in multi-dimensional case.

Given $h : R^k \to R$, let $h_0 = \int h(\boldsymbol{z})d\Phi_k(\boldsymbol{z})$ be a constant, $h_k(\boldsymbol{z}) = h(\boldsymbol{z})$,

$$h_j(z_1, \ldots, z_j) = \int_{R^{k-j}} h(z_1, \ldots, z_j, \boldsymbol{w})d\Phi_{k-j}(\boldsymbol{w}), \text{ and} \tag{2}$$

$$g_j(z_1, \ldots, z_k) = e^{z_j^2/2} \int_{z_j}^{\infty} [h_j(z_1, \ldots, z_{j-1}, w) - h_{j-1}(z_1, \ldots, z_{j-1})]e^{-w^2/2}dw, \tag{3}$$

for $-\infty < z_1, \ldots, z_k < \infty$ and $j = 1, \ldots, k$. Then let

$$Uh = [g_1, \ldots, g_k]^T \qquad \text{and} \qquad Vh = \frac{U^2h + (U^2h)^T}{2}, \tag{4}$$

where U^2h is the $k \times k$ matrix whose jth column is Ug_j and g_j is as in (3).

Let Γ be a measure of the form:

$$d\Gamma(\boldsymbol{z}) = f(\boldsymbol{z})\phi_k(\boldsymbol{z})d\boldsymbol{z}, \tag{5}$$

where f is a real-valued function (not necessarily non-negative) defined on R^k.

Lemma 2.1 *(W-Stein's Identity) Suppose that* $d\Gamma$ *is defined as in (5), where* f *is almost differentiable. Let* h *be a real-valued function defined on* R^k*. Then,*

$$\int h(\boldsymbol{z})d\Gamma(\boldsymbol{z}) = \int f(\boldsymbol{z})d\Phi_k(\boldsymbol{z}) \cdot \int h(\boldsymbol{z})d\Phi_k(\boldsymbol{z}) + \int (Uh(\boldsymbol{z}))^T \nabla f(\boldsymbol{z})d\Phi_k(\boldsymbol{z}), \tag{6}$$

provided all the integrals are finite.

Lemma 2.1 was given by [8]. The idea of this identity originated from Stein's lemma [6], but the latter considers the expectation with respect to a normal distribution (i.e. the integral $\int h(\boldsymbol{z})d\Phi_k(\boldsymbol{z})$), while the former studies the integration with respect to a "nearly normal distribution" Γ in the sense of (5). Stein's lemma is famous and of interest because of its applications to James-Stein estimator [4] and empirical Bayes methods.

Essentially the proof is based on exchanging the order of integration (Fibini theorem), and it is the very idea for proving Stein's lemma. Due to this reason, Woodroofe termed (6) a version of Stein's identity. However, to distinguish it from Stein's lemma, here we refer to it as W-Stein's identity.

Now we assume that $\partial f/\partial z_j,\ j = 1, \ldots, k$ are almost differentiable. Then, by writing

$$(Uh(\boldsymbol{z}))^T \nabla f(\boldsymbol{z}) = \sum_{i=1}^{k} g_i(\boldsymbol{z}) \frac{\partial f(\boldsymbol{z})}{\partial z_i}$$

and applying (6) with h and f replacing by g_i and $\partial f/\partial z_i$, we obtain

$$\int g_i \frac{\partial f}{\partial z_i} d\Phi_k(\boldsymbol{z}) = \Phi_k(g_i) \int \frac{\partial f}{\partial z_i} d\Phi_k(\boldsymbol{z}) + \int (U(g_i))^T \nabla \left(\frac{\partial f}{\partial z_i} \right) d\Phi_k(\boldsymbol{z}), \quad (7)$$

provided all the integrals are finite. Note that $\Phi_k(g_i)$ in the above equation is a constant defined as

$$\Phi_k(g_i) = \int g_i(\boldsymbol{z}) \phi_k(\boldsymbol{z}) d\boldsymbol{z}.$$

By summing up both sides of (7) over $i = 1, \ldots, k$, we can rewrite (6) as

$$\begin{aligned} \int h(\boldsymbol{z}) f(\boldsymbol{z}) d\Phi_k(\boldsymbol{z}) &= \int f(\boldsymbol{z}) d\Phi_k(\boldsymbol{z}) \cdot \int h(\boldsymbol{z}) d\Phi_k(\boldsymbol{z}) + (\Phi_k U h)^T \int \nabla f(\boldsymbol{z}) d\Phi_k(\boldsymbol{z}) \\ &\quad + \int \operatorname{tr} \left[(V h(\boldsymbol{z})) \nabla^2 f(\boldsymbol{z}) \right] d\Phi_k(\boldsymbol{z}); \end{aligned} \quad (8)$$

see Proposition 2 of [9] and Lemma 1 of [7]. Here $\Phi_k U h = (\Phi_k(g_1), ..., \Phi_k(g_k))^T$, "tr" denotes the trace of a matrix, and $\nabla^2 f$ the Hessian matrix of f.

Let $\boldsymbol{Z} = [Z_1, \ldots, Z_k]^T$ be a k-dimensional random vector with the probability density

$$C \phi_k(\boldsymbol{z}) f(\boldsymbol{z}), \quad (9)$$

where

$$C = \left(\int \phi_k(\boldsymbol{z}) f(\boldsymbol{z}) d\boldsymbol{z} \right)^{-1}$$

is the normalizing constant. Lemma 2.1 can be applied to obtain expectations of functions of $\boldsymbol{Z}$ in the following corollary.

Corollary 2.1 *Suppose that $\boldsymbol{Z}$ has probability density (9). Then,*

$$\int f d\Phi_k = C^{-1} \text{ and } Eh(\boldsymbol{Z}) = \int h(\boldsymbol{z}) d\Phi_k(\boldsymbol{z}) + E\left[(Uh(\boldsymbol{Z}))^T \frac{\nabla f(\boldsymbol{Z})}{f(\boldsymbol{Z})} \right]. \quad (10)$$

Further, (8) *and* (10) *imply*

$$Eh(\boldsymbol{Z}) = \int h(\boldsymbol{z}) d\Phi_k(\boldsymbol{z}) + (\Phi_k U h)^T E\left[\frac{\nabla f(\boldsymbol{Z})}{f(\boldsymbol{Z})} \right] + E\left[\operatorname{tr} \left(V h(\boldsymbol{Z}) \frac{\nabla^2 f(\boldsymbol{Z})}{f(\boldsymbol{Z})} \right) \right]. \quad (11)$$

In particular, if $h(\boldsymbol{z}) = z_i$, then by (4) it follows $Uh(\boldsymbol{z}) = \boldsymbol{e}_i$ (a function from R^k to R^k); and if $h(\boldsymbol{z}) = z_i z_j$ and $i < j$, then $Uh(\boldsymbol{z}) = z_i \boldsymbol{e}_j$, where $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_k\}$ denote the standard basis for R^k.

With these special h functions, (10) and (11) become

$$E[\boldsymbol{Z}] = E\left[\frac{\nabla f(\boldsymbol{Z})}{f(\boldsymbol{Z})}\right], \tag{12}$$

$$E[Z_i Z_q] = \delta_{iq} + E\left[\frac{\nabla^2 f(\boldsymbol{Z})}{f(\boldsymbol{Z})}\right]_{iq}, \quad i, q = 1, \ldots, k, \tag{13}$$

where $\delta_{iq} = 1$ if $i = q$ and 0 otherwise, and $[\cdot]_{iq}$ indicates the (i, q) component of a matrix.

In the current context of online ranking, since the skill θ is assumed to follow a Gaussian distribution, the update procedure is mainly for the mean and the variance. Therefore, (12) and (13) will be useful. The detailed approximation procedure is in the next section.

3 Expectation Approximation

Let θ_i be the strength parameter of team i whose ability is to be estimated. The Bayesian framework starts by assuming that θ_i has a prior distribution $N(\mu_i, \sigma_i^2)$ with μ_i and σ_i^2 known. next models the game outcome by some probability models, and then updates the skill (the posterior mean and variance of θ_i) at the end of the game. These revised mean and variance are considered as prior information for the next game.

To see how Eqs. (12) and (13) can be applied to online skill updates, first suppose that team i has a strength parameter θ_i and assume that the prior distribution of θ_i is $N(\mu_i, \sigma_i^2)$. Upon the completion of a game, their skills are characterized by the posterior mean and variance of $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_k]^T$. Let D denote the result of a game and $\boldsymbol{Z} = [Z_1, \ldots, Z_k]^T$ with

$$Z_i = \frac{\theta_i - \mu_i}{\sigma_i}, i = 1, \ldots, k, \tag{14}$$

where k is the number of teams. The posterior density of $\boldsymbol{Z}$ given the game outcome D is

$$P(\boldsymbol{z}|D) = C\phi_k(\boldsymbol{z})f(\boldsymbol{z}), \tag{15}$$

where $f(\boldsymbol{z})$ is the probability of game outcome $P(D|\boldsymbol{z})$. Thus, $P(\boldsymbol{z}|D)$ is of the form (9). Subsequently we omit D in all derivations.

Next, Eqs. (12), (13) and the relation between Z_i and θ_i in (14) give that

$$\begin{aligned}\mu_i^{\text{new}} =& E[\theta_i] = \mu_i + \sigma_i E[Z_i] \\ =& \mu_i + \sigma_i E\left[\frac{\partial f(\boldsymbol{Z})/\partial Z_i}{f(\boldsymbol{Z})}\right]\end{aligned} \tag{16}$$

and

$$\begin{aligned}(\sigma_i^{\text{new}})^2 =& \text{Var}[\theta_i] = \sigma_i^2 \text{Var}[Z_i] \\ =& \sigma_i^2 \left(E[Z_i^2] - E[Z_i]^2\right) \\ =& \sigma_i^2 \left(1 + E\left[\frac{\nabla^2 f(\boldsymbol{Z})}{f(\boldsymbol{Z})}\right]_{ii} - E\left[\frac{\partial f(\boldsymbol{Z})/\partial Z_i}{f(\boldsymbol{Z})}\right]^2\right).\end{aligned} \tag{17}$$

Similarly, we can write the last two terms on the right side of (17) as

$$\sigma_i^2 \left(E\left[\frac{\nabla^2 f(\boldsymbol{Z})}{f(\boldsymbol{Z})}\right]_{ii} - E\left[\frac{\partial f(\boldsymbol{Z})/\partial Z_i}{f(\boldsymbol{Z})}\right]^2\right) = E\left[\frac{\partial^2 \log f(\boldsymbol{Z})}{\partial \theta_i^2}\right],$$

which is the average of the rate of change of $\partial(\log f)/\partial\theta_i$ with respect to θ_i.

We propose approximating expectations in (16) and (17):

$$\mu_i \leftarrow \mu_i + \Omega_i, \tag{18}$$

$$\sigma_i^2 \leftarrow \sigma_i^2 \max(1 - \Delta_i, \kappa), \tag{19}$$

where

$$\Omega_i = \sigma_i \left.\frac{\partial f(\boldsymbol{z})/\partial z_i}{f(\boldsymbol{z})}\right|_{\boldsymbol{z}=\boldsymbol{0}} \tag{20}$$

and

$$\begin{aligned}\Delta_i =& -\left.\frac{\partial^2 f(\boldsymbol{z})/\partial^2 z_i}{f(\boldsymbol{z})}\right|_{\boldsymbol{z}=\boldsymbol{0}} + \left(\left.\frac{\partial f(\boldsymbol{z})/\partial z_i}{f(\boldsymbol{z})}\right|_{\boldsymbol{z}=\boldsymbol{0}}\right)^2 \\ =& -\left.\frac{\partial}{\partial z_i}\left(\frac{\partial f(\boldsymbol{z})/\partial z_i}{f(\boldsymbol{z})}\right)\right|_{\boldsymbol{z}=\boldsymbol{0}}.\end{aligned} \tag{21}$$

We set $\boldsymbol{z} = \boldsymbol{0}$ so that $\boldsymbol{\theta}$ is replaced by $\boldsymbol{\mu}$. Such a substitution is reasonable as the posterior density of $\boldsymbol{\theta}$ is likely to be concentrated on $\boldsymbol{\mu}$. Then the right-hand sides of (18)-(19) are functions of $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$, so we can use the current values to obtain new estimates. Due to the approximation (21), $1 - \Delta_i$ may be negative. Hence in (19) we set a small positive lower bound κ to avoid a negative σ_i^2.

4 Remarks

We derive analytic update rules. Unlike the TrueSkill system, our rules do not need numerical integrations and are very easy to interpret and implement. Further experiments on game data show that our accuracy is competitive with state of the art systems such as TrueSkill, but the running time is much shorter.

References

[1] A. E. Elo. *The Rating of Chessplayers, Past and Present.* Arco Pub., New York, 2nd edition, 1986.

[2] M. Glickman. Parameter estimation in large dynamic paired comparison experiments. *Applied Statistics*, 48(3):377–394, 1999.

[3] R. Herbrich, T. Minka, and T. Graepel. TrueSkill™: A bayesian skill rating system. In B. Schölkopf, J. Platt, and T. Hoffman, editors, *Advances in Neural Information Processing Systems 19*, pages 569–576. MIT Press, Cambridge, MA, 2007.

[4] W. James and C. Stein. Estimation with quadratic loss. In *Proc. Fourth Berkeley Symp. Math. Statist. Prob.*, volume 1, pages 361–379, 1961.

[5] T. Minka. *A family of algorithms for approximate Bayesian inference.* PhD thesis, MIT, 2001.

[6] C. Stein. Estimation of the mean of a multivariate normal distribution. *Ann. Statist.*, 9:1135–1151, 1981.

[7] R. C. Weng and M. Woodroofe. Integrable expansions for posterior distributions for multiparameter exponential families with applications to sequential confidence levels. *Statistica Sinica*, 10:693–713, 2000.

[8] M. Woodroofe. Very weak expansions for sequentially designed experiments: linear models. *The Annals of Statistics*, 17:1087–1102, 1989.

[9] M. Woodroofe and D. S. Coad. Corrected confidence sets for sequentially designed experiments. *Statistica Sinica*, 7:53–74, 1997.