The AGA Ratings System

Philip Waldron

June 23, 2010

1 Basic description

The AGA ratings system uses a Bayesian approach where a player's rating rating is updated based on additional information that is available in the form of new game results. For players that are new to the system, an initial guess is made by the tournament director based on an initial assessment of the player's strength (see section 4.1).

The rating system is scaled so that it corresponds to the traditional kyu/dan ranks. Ratings in the range $(3.0, 4.0)$, for example, correspond to that of a 3-dan, while those in the range $(-4.0, -3.0)$ correspond to that of a 3-kyu. There are no ratings between -1.0 and 1.0; there is a ratings difference of 0.02 between -1.01 and 1.01.

2 Mathematical description

Player's prior ratings are modelled as a normally distributed parameter with mean μ (usually equal to the input rating from the previous ratings run) and standard deviation σ, which is updated based on new information (i.e. new games). The probability that a player has a rating of r is given by

$$P_p(r_i) = \frac{1}{\sqrt{2\pi}} \exp(-z^2/2), \tag{1}$$

where $z = \frac{r-\mu}{\sigma}$.

Game results are modelled with three parameters: d, b and σ_{px}. The handicap equivalent, d, is given by

$$d = \begin{cases} 0.580 - 0.0757k & \text{if } h\text{=0 or 1} \\ h - 0.0757k & \text{otherwise} \end{cases} \tag{2}$$

Handicap	b
2	1.13672
3	1.18795
4	1.22841
5	1.27457
6	1.31978
7	1.35881
8	1.39782
9	1.43614

Table 1: b parameters for σ_{px} calculations for 2-9 stone handicaps.

where h and k are the handicap and komi conditions of the game. A second parameter, σ_{px}, is also required:

$$\sigma_{\text{px}} = \begin{cases} 1.0649 - 0.0021976k + 0.00014984k^2 & \text{if } h\text{=0 or 1} \\ -0.0035169k + b & \text{otherwise} \end{cases} \tag{3}$$

where b is found in Table 1 and k is once again the komi.

The probability of White winning a particular game, j, is estimated as

$$\begin{aligned} P_g(\text{White wins}) &= \frac{1}{2} + \frac{1}{2}\text{erf}\left(\frac{\text{RD}_j}{\sigma_{\text{px}}\sqrt{2}}\right) && (4) \\ &= \frac{1}{2} - \frac{1}{2}\text{erf}\left(\frac{-\text{RD}_j}{\sigma_{\text{px}}\sqrt{2}}\right) && (5) \\ &= \frac{1}{2}\text{erfc}\left(\frac{-\text{RD}_j}{\sigma_{\text{px}}\sqrt{2}}\right) && (6) \end{aligned}$$

where $\text{erf}(x)$ denotes the error function and

$$\text{RD}_j = r_{\text{white}} - r_{\text{black}} - d_j. \tag{7}$$

The probability of Black winning the same game is

$$\begin{aligned} P_g(\text{Black wins}) &= 1 - P(\text{White wins}) && (8) \\ &= \frac{1}{2}\text{erfc}\left(\frac{\text{RD}_j}{\sigma_{\text{px}}\sqrt{2}}\right). && (9) \end{aligned}$$

For a single game between players 1 and 2, the total likelihood function is given by the product of the player's individual rating probability multiplied by their game result:

$$L = P_p(r_1) \cdot P_p(r_2) \cdot P_g(\text{game result}|r_1, r_2). \tag{10}$$

Generalizing to multiple players, the likelihood function becomes

$$L = \prod_i P_p(r_i) \prod_j P_g(\text{outcome of game } j). \tag{11}$$

AGA ratings are calculated by finding the set of ratings, $\{r_i\}$, that maximize the likelihood given by equation (11). Conceptually this is straightforward; the difficulty lies in the implementation details.

3 Implementation

The nature of computers as finite-precision system presents a few challenges in implementing the ratings algorithm as stated. Multiplication of probabilities will lead to catastrophic round-off errors for calculations of even moderate size. It is convenient to recast the likelihood function of equation (11) by taking its logarithm.

$$\log(L) = \sum_i \log\left(P_p(r_i)\right) + \sum_j \log\left(P_g(\text{outcome of game } j)\right). \tag{12}$$

The maximum likelihood still occurs for the same set of $\{r_i\}$, but addition is easier than multiplication.

Numerical calculations involving games between players with wildly different ratings is also troublesome. Calculating $\log(\text{erfc}(x))$ using intermediate function calls yields underflow errors for even moderately negative values of x. The GNU Scientific Library (GSL) offers a library function to calculate $\log(\text{erfc}(x))$ directly.

Problems in scientific computation tend to focus on minimization rather than maximization, so it it more useful to convert the maximum likelihood calculation into one of minimizing the negative of the likelihood function. A ratings run is performed by passing the negative of equation (12) to the minimization algorithms of the GSL. Minimization is usually done with the Fletcher-Reeves conjugate gradient method; the simpler simplex method is

used as a backup. Because conjugate-gradient methods require information about the gradient of equation (12), the relevant expressions are included here:

For the probability of a player's rating:

$$P_p(r_i) = \frac{1}{\sqrt{2\pi}} \exp(-z^2/2) \tag{13}$$

so

$$\begin{aligned} \log\left(P_p(r_i)\right) &= \log\left(\frac{1}{\sqrt{2\pi}}\right) + \log\left(\exp(-z^2/2)\right) && (14) \\ &= -\frac{z^2}{2} - \log(\sqrt{2\pi}) && (15) \end{aligned}$$

and

$$\frac{\partial}{\partial r_i} \log\left(P_p(r_i)\right) = -z = -\frac{r_i - \mu_i}{\sigma_i^2}. \tag{16}$$

For the game probabilities

$$\begin{aligned} P_g(\text{White wins}) &= \frac{1}{2}\text{erfc}\left(\frac{-\text{RD}}{\sigma_{\text{px}}\sqrt{2}}\right) && (17) \\ P_g(\text{Black wins}) &= \frac{1}{2}\text{erfc}\left(\frac{\text{RD}}{\sigma_{\text{px}}\sqrt{2}}\right) && (18) \end{aligned}$$

we have

$$\begin{aligned} \log\left(P_g(\text{White wins})\right) &= \log\left(\text{erfc}\left(\frac{-\text{RD}}{\sigma_{\text{px}}\sqrt{2}}\right)\right) - \log(2) && (19) \\ \log\left(P_g(\text{Black wins})\right) &= \log\left(\text{erfc}\left(\frac{\text{RD}}{\sigma_{\text{px}}\sqrt{2}}\right)\right) - \log(2). && (20) \end{aligned}$$

Therefore if White wins a game

$$\begin{aligned} \frac{\partial \log\left(P_g(\text{White wins})\right)}{\partial r_{\text{White}}} &= \frac{1}{\sigma_{\text{px}}}\sqrt{\frac{2}{\pi}}\frac{1}{\text{erfc}\left(\frac{-\text{RD}}{\sigma_{\text{px}}\sqrt{2}}\right)} \exp\left(-\frac{\text{RD}^2}{2\sigma_{\text{px}}^2}\right) && (21) \\ \frac{\partial \log\left(P_g(\text{White wins})\right)}{\partial r_{\text{Black}}} &= -\frac{1}{\sigma_{\text{px}}}\sqrt{\frac{2}{\pi}}\frac{1}{\text{erfc}\left(\frac{-\text{RD}}{\sigma_{\text{px}}\sqrt{2}}\right)} \exp\left(-\frac{\text{RD}^2}{2\sigma_{\text{px}}^2}\right) && (22) \end{aligned}$$

If Black wins a game

$$\frac{\partial \log\left(P_g(\text{Black wins})\right)}{\partial r_{\text{White}}} = -\frac{1}{\sigma_{\text{px}}}\sqrt{\frac{2}{\pi}}\frac{1}{\operatorname{erfc}\left(\frac{\text{RD}}{\sigma_{\text{px}}\sqrt{2}}\right)}\exp\left(-\frac{\text{RD}^2}{2\sigma_{\text{px}}^2}\right) \quad (23)$$

$$\frac{\partial \log\left(P_g(\text{Black wins})\right)}{\partial r_{\text{Black}}} = \frac{1}{\sigma_{\text{px}}}\sqrt{\frac{2}{\pi}}\frac{1}{\operatorname{erfc}\left(\frac{\text{RD}}{\sigma_{\text{px}}\sqrt{2}}\right)}\exp\left(-\frac{\text{RD}^2}{2\sigma_{\text{px}}^2}\right) \quad (24)$$

The estimated uncertainty in each player's rating is determined from the Fischer information matrix:

$$\mathcal{I} = -\begin{pmatrix} \frac{\partial^2 \log(L)}{\partial r_1^2} & \frac{\partial^2 \log(L)}{\partial r_1 \partial r_2} & \cdots & \frac{\partial^2 \log(L)}{\partial r_1 \partial r_n} \\ \frac{\partial^2 \log(L)}{\partial r_1 \partial r_2} & \frac{\partial^2 \log(L)}{\partial r_2^2} & \cdots & \frac{\partial^2 \log(L)}{\partial r_2 \partial r_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \log(L)}{\partial r_1 \partial r_n} & \frac{\partial^2 \log(L)}{\partial r_2 \partial r_n} & \cdots & \frac{\partial^2 \log(L)}{\partial r_n^2} \end{pmatrix}, \quad (25)$$

where the functions are evaluated at the maximum likelihood point given by the minimization algorithm. Fast matrix inversion routines are available in the GSL. The variance of the rating calculated for player i is estimated as

$$\sigma_i = \mathcal{I}_{ii}^{-1}. \quad (26)$$

The second derivatives present in equation 25 come from two sources. The contribution of the player's rating probability is given by

$$\frac{\partial^2 \log(P_p)}{\partial r_i \partial r_j} = -\delta_{ij}\sigma, \quad (27)$$

where δ_{ij} is the Kronecker delta function.

The contribution from game probability depends on the winner of the game. If White wins

$$\begin{aligned} \frac{\partial^2 \log\left(P_g\right)}{\partial r_{\text{White}}^2} &= -\sqrt{\frac{2}{\pi}}\frac{\text{RD}}{\sigma_{px}^3}\frac{\exp\left(-\frac{(\text{RD})^2}{2\sigma_{px}^2}\right)}{\operatorname{erfc}\left(-\frac{\text{RD}}{\sqrt{2}\sigma_{px}}\right)} - \frac{2}{\pi}\frac{1}{\sigma_{px}^2}\frac{\exp^2\left(-\frac{(\text{RD})^2}{2\sigma_{px}^2}\right)}{\operatorname{erfc}^2\left(-\frac{\text{RD}}{\sqrt{2}\sigma_{px}}\right)} \quad (28) \\ &= \frac{\partial^2 \log\left(P_g\right)}{\partial r_{\text{Black}}^2} \\ &= -\frac{\partial^2 \log\left(P_g\right)}{\partial r_{\text{Black}} \partial r_{White}} \\ &= -\frac{\partial^2 \log\left(P_g\right)}{\partial r_{\text{Black}} \partial r_{White}} \end{aligned}$$

If Black wins

$$\begin{aligned}
\frac{\partial^2 \log(P_g)}{\partial r^2_{\text{White}}} &= \sqrt{\frac{2}{\pi}} \frac{\text{RD}}{\sigma^3_{px}} \frac{\exp\left(-\frac{(\text{RD})^2}{2\sigma^2_{px}}\right)}{\text{erfc}\left(-\frac{\text{RD}}{\sqrt{2}\sigma_{px}}\right)} - \frac{2}{\pi}\frac{1}{\sigma^2_{px}} \frac{\exp^2\left(-\frac{(\text{RD})^2}{2\sigma^2_{px}}\right)}{\text{erfc}^2\left(-\frac{\text{RD}}{\sqrt{2}\sigma_{px}}\right)} \qquad (29) \\
&= \frac{\partial^2 \log(P_g)}{\partial r^2_{\text{Black}}} \\
&= -\frac{\partial^2 \log(P_g)}{\partial r_{\text{Black}} \partial r_{White}} \\
&= -\frac{\partial^2 \log(P_g)}{\partial r_{\text{Black}} \partial r_{White}}.
\end{aligned}$$

4 Other details

4.1 Initial player seeding

Players entering the rating system for the first time are given an initial rating and sigma seed. The seed rating is taken to be in the middle of the rating range corresponding to a player's entry rank. A new 3-dan, for example, is given an initial rating of 3.5.

The initial σ depends on the seed rating. Players with seed ratings greater than 7.5 are assigned $\sigma = 1$, while those with ratings less than -50.5 are assigned $\sigma = 6$. For other players, the initial sigma is plotted in Figure 1. A spline interpolation is used instead of a lookup table.

After the initial seeding, output from one rating run is used as the input for the next ratings run.

4.2 Self-promotions

Self promotion is a necessary part of the AGA ratings. It allows the system to adapt to rapidly improving players and compensates for inactive players returning to rated play after having improved markedly in their absence. We break the self-promotion question into two cases. The first is the case of a player who claims a three or more stones improvement from his last calculated rating and wins at least one playing at the new rating. We assume that such a promotion is largely substantiated by actual play and simply reseed the player at the claimed rating.

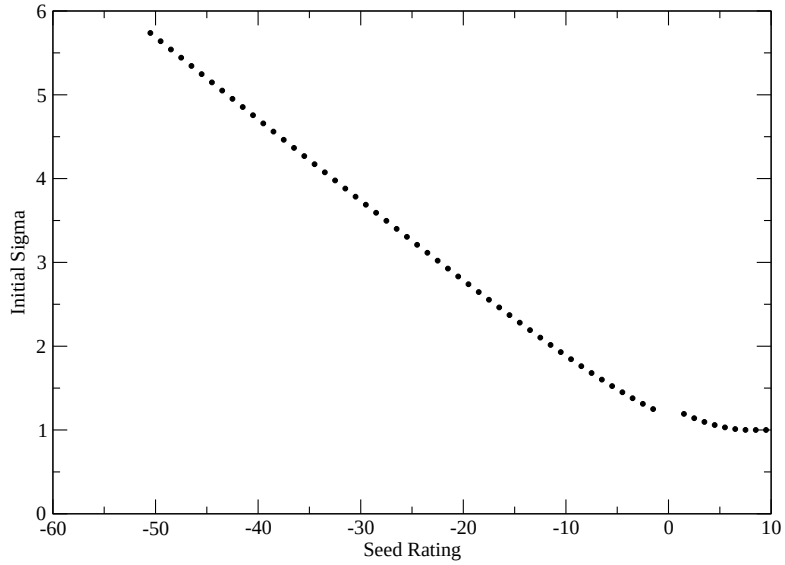

Figure 1: Assigned sigma parameter as a function of seed rating for a new player entering the system.

For promotions of fewer than three stones, the new seed rating, μ is adjusted according to

$$\mu = \mu_0 + \begin{cases} 0 & \text{if promotion is less than 1.0} \\ 0.024746 + 0.32127\Delta r & \text{otherwise} \end{cases} \tag{30}$$

where μ_0 is the result of the previous rating run and Δr is the amount of the self-promotion. The new σ is given by

$$\sigma = \sqrt{\sigma_0^2 + 0.256\Delta r^{1.9475}}, \tag{31}$$

where σ_0 is the value of σ from the previous ratings run and Δr is the amount of the self-promotion.

4.3 Out-of-date ratings

The reliability of ratings decreases over time. All ratings are devalued somewhat compared to more recent calculations by increasing the σ parameter for older ratings according to the expression

$$\sigma(t) = \sqrt{\sigma_0^2 + \alpha^2 t^2}, \tag{32}$$

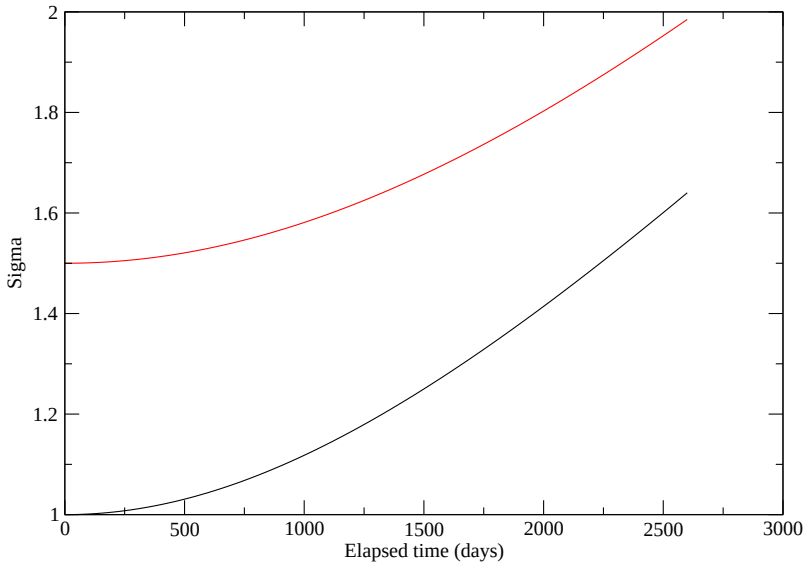

Figure 2: Evolution of sigma with time.

where σ_0 is the value of σ after the last ratings run and $\alpha = 0.0005\,\mathrm{day}^{-1}$. Figure 2 gives the time evolution of $\sigma = 1.0$ and $\sigma = 1.5$. Equation 32 applies to all ratings, although the effect will be larger for older values.