1. The SPSA "process"

We analyze the SPSA algorithm in 1 dimension as implemented in fishtest. w, d, l denote respectively the win, draw and loss ratio. We use the Bayes Elo model. I.e. we have for $\beta = \log(10)/400$

$$\begin{aligned} w &= \frac{1}{1+e^{-\beta(x-\delta)}} \\ l &= \frac{1}{1+e^{-\beta(-x-\delta)}} \end{aligned}$$

where δ is the drawelo parameter. We score an engine game by $\pm 1, 0, -1$ for respectively a win, draw, loss. We assume that x is small. We have for $E = e^{\beta\delta}$.

$$\begin{aligned} w &= \frac{1}{1+Ee^{-\beta x}} \\ &\cong \frac{1}{1+E(1-\beta x)} \\ &= \frac{1}{1+E}\left(1+\frac{\beta x}{1+E}\right) \\ &= \frac{1}{1+E} + \frac{\beta x E}{(1+E)^2} \end{aligned}$$

and similarly

$$l = \frac{1}{1+E} - \frac{\beta x E}{(1+E)^2}$$

Hence the expected score and its standard deviation are respectively given by

$$\begin{aligned} \mu &= w - l \\ &\cong \frac{2\beta x E}{(1+E)^2} \\ \sigma &= \sqrt{w + l - (w-l)^2} \\ &\cong \sqrt{\frac{2}{1+E}} \end{aligned}$$

Let p be an engine parameter and let $f(p)$ be the function that maps p to Bayes Elo. We play a match of $p+c$ against $p-c$ and replace p by[1]

$$\begin{cases} p + \frac{a}{c} & \text{in case of a win} \\ p & \text{in case of a draw} \\ p - \frac{a}{c} & \text{in case of a loss} \end{cases}$$

So we can think of the dynamics of p as a discrete version of the continuous process

$$dp = \frac{a}{c}(\mu dt + \sigma dW_t)$$

where (μ, σ) are the mean and standard deviation of a single game and W_t is a Wiener process. Using the approximation

$$x = f(p+c) - f(p-c) = 2cf'(p)$$

[1] We see that a is the learning rate. Using $a/2c$ would be somewhat more natural.

we find

$$dp = \frac{a}{c}\left(\frac{4\beta c f'(p)E}{(1+E)^2}dt + \sqrt{\frac{2}{1+E}}dW_t\right)$$

Hence

$$dp = \frac{4\beta a E}{(1+E)^2}f'(p)dt + \frac{a}{c}\sqrt{\frac{2}{1+E}}dW_t$$

which we will write as

$$dp = u_1 a f'(p)dt + u_2\frac{a}{c}dW_t$$

with

$$u_1 = \frac{4\beta E}{(1+E)^2}$$
$$u_2 = \sqrt{\frac{2}{1+E}}$$

Example 1.1. Assume $\delta = 220$ (a realistic value for self testing at short time controls). Then we find

$$u_1 = 0.003950$$
$$u_2 = 0.6631$$

In the special case that

$$f(p) = -\xi(p-p_0)^2 \tag{1}$$

$$dp = -2u_1 a\xi(p-p_0)dt + u_2 dW_t$$

If a, c are constant then this is an Ornstein-Uhlenbeck process

$$dp = -\theta(p-\mu)dt + \sigma dW_t$$

with

$$\mu = p_0$$
$$\theta = 2u_1 a\xi$$
$$\sigma = u_2\frac{a}{c}$$

Remark 1.2. The parameters in an Ornstein-Uhlenbeck process are actually just location and scaling parameters for p and t. Replacing p by $p-\mu$ we may assume $\mu = 0$. Putting $t = \theta^{-1}s$, $p = \sigma\theta^{-1/2}q$ and using that $W_{\theta s} = \theta^{1/2}W_s$ we end up with the equation

$$dq = -q ds + dW_s$$

We deduce that

$$\theta^{-1} = \frac{1}{2u_1\xi a}$$

acts like a natural time scale and that

$$\frac{\sigma}{\sqrt{\theta}} = \frac{u_2}{\sqrt{2u_1\xi}}\frac{\sqrt{a}}{c}$$

is a natural scale for the range of the parameter p in SPSA.

2. Solving a case with varying hyperparameters

Now we consider the case where f is quadratic as in (1) and $a = a(t)$, $c = c(t)$ are time dependent via

$$a(t) = \frac{a_0}{(A+t)^\alpha}$$
$$c(t) = \frac{c_0}{t^\gamma}$$

It is clear that without loss of generality we may assume $p_0 = 0$. We have to slove

$$dp = -2\xi u_1 \frac{a_0}{(A+t)^\alpha} p dt + u_2 \frac{a_0}{c_0} \frac{t^\gamma}{(A+t)^\alpha} dW_t$$

We consider first the unperturbed system

$$dp = -2\xi u_1 \frac{a_0}{(A+t)^\alpha} p dt$$

It can be rewritten as

$$d \log p = -\frac{2\xi u_1 a_0}{1-\alpha} d(A+t)^{1-\alpha}$$

so that we get

$$p = C \exp\left(-\frac{2\xi u_1 a_0}{1-\alpha}(A+t)^{1-\alpha}\right)$$

We now find a particular solution of the perturbed system using variation of constants. We get

$$\exp\left(-\frac{2\xi u_1 a_0}{1-\alpha}(A+t)^{1-\alpha}\right) dC = u_2 \frac{a_0}{c_0} \frac{t^\gamma}{(A+t)^\alpha} dW_t$$

so that we may as a particular solution we may take

$$u_2 \frac{a_0}{c_0} \exp\left(-\frac{2\xi u_1 a_0}{1-\alpha}(A+t)^{1-\alpha}\right) \int_0^t \exp\left(\frac{2\xi u_1 a_0}{1-\alpha}(A+s)^{1-\alpha}\right) \frac{s^\gamma}{(A+s)^\alpha} dW_s$$

and the full solution is given by

$$\exp\left(-\frac{2\xi u_1 a_0}{1-\alpha}(A+t)^{1-\alpha}\right)\left(C + u_2 \frac{a_0}{c_0} \int_0^t \exp\left(\frac{2\xi u_1 a_0}{1-\alpha}(A+s)^{1-\alpha}\right) \frac{s^\gamma}{(A+s)^\alpha} dW_s\right)$$

If follows that a time t we have $p \sim N(\mu_t, \sigma_t^2)$ with

$$\mu_t = C \exp\left(-\frac{2\xi u_1 a_0}{1-\alpha}(A+t)^{1-\alpha}\right)$$

$$\sigma_t = u_2 \frac{a_0}{c_0} \exp\left(-\frac{2\xi u_1 a_0}{1-\alpha}(A+t)^{1-\alpha}\right) \sqrt{\int_0^t \exp\left(\frac{4\xi u_1 a_0}{1-\alpha}(A+s)^{1-\alpha}\right) \frac{s^{2\gamma}}{(A+s)^{2\alpha}} ds}$$

Let us now investigate the asymptotic behaviour of the intergral under the square root sign. Put

$$r = (A+s)^{1-\alpha}$$

so that

$$\begin{aligned} dr &= (1-\alpha)(A+s)^{-\alpha} ds \\ &= (1-\alpha) r^{-\alpha/(1-\alpha)} ds \end{aligned}$$

and hence

$$ds = \frac{1}{1-\alpha} r^{\alpha/(1-\alpha)} dr$$

We obtain

$$\int_0^t \exp\left(\frac{4\xi u_1 a_0}{1-\alpha}(A+s)^{1-\alpha}\right) \frac{s^{2\gamma}}{(A+s)^{2\alpha}} ds = \frac{1}{1-\alpha} \int_{A^{1-\alpha}}^{(A+t)^{1-\alpha}} \exp\left(\frac{4\xi u_1 a_0}{1-\alpha} r\right) \frac{(r^{1/(1-\alpha)} - A)^{2\gamma}}{r^{\alpha/(1-\alpha)}} dr$$

So it seems that we have to understand the asymptotic behaviour of

$$\int_{K_0}^{K} e^{r-K} r^{-\psi} dr = \int_0^{K-K_0} \frac{e^{-r}}{(K-r)^{\psi}} dr$$

for $K \to \infty$, where $\psi = (\alpha - 2\gamma)/(1-\alpha)$ which in the case of fishtest ($\alpha = 0.602$, $\gamma = 0.101$ is slightly bigger 1).