$\Delta Q = Q_N - Q_E$	quality difference
y	object variables
z	random number with $z \sim \mathcal{N}(0, \sigma^2)$
z^*	random number with $z^* \sim \mathcal{N}(0, \sigma^{*2})$
c_k, d_k	parameters of the quality function
$\Pr[\Delta Q > 0]$	success probability
ϕ', $\phi'_{1,\lambda}$	progress rate (with/without selection)
$\phi_{1,\lambda}$	averaged progress rate
λ	number of offspring
u	random variable
$c_{1,\lambda}$	progress coefficients

Table 1: Nomenclature

ES Theory: Success probability

The quality function is given by:

$$Q = Q_0 + \sum_{k=1}^{n} c_k y_k \; - \sum_{k=1}^{n} d_k {y_k}^2; \;\; d_k > 0 \tag{11}$$

We assume that the parent individual is located at $y_k = 0$, $k = 1, \ldots, n$. Using ES type mutation: $y'_k = y_k + z_k$, where

$$\begin{aligned} z_k &\sim \mathcal{N}(0, \sigma^2) \\ w(z_k) &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{z_k^2}{2\sigma^2}\right) \end{aligned}$$

is normally distributed with variance σ^2, we can write

$$\begin{aligned} \Delta Q &= Q_N - Q_E \\ &= Q_0 + \sum_{k=1}^{n} c_k y'_k \; - \sum_{k=1}^{n} d_k {y'_k}^2 - Q_0 \\ &= \sum_{k=1}^{n} c_k z_k - \sum_{k=1}^{n} d_k {z_k}^2 \end{aligned} \tag{12}$$

Now, we use the following two relations in order to simplify (12):

1.

$$\mathrm{Var}\left[z^*\right] = \sum_{k=1}^{n} \mathrm{Var}\left[c_k \, z_k\right] = \sigma^2 \sum_{k=1}^{n} c_k^2$$

$$\Rightarrow \quad \sigma^{*2} = \sigma^2 \sum_{k=1}^{n} c_k^2$$

$$\Rightarrow \quad \sigma^{*} = \sigma \sqrt{\sum_{k=1}^{n} c_k^2} \tag{13}$$

2.

$$\mathrm{E}\left[\sum_{k=1}^{n} d_k z_k^2\right] = \sigma^2 \sum_{k=1}^{n} d_k \tag{14}$$

A sum of normally distributed random numbers results in a normally distributed random number with the standard deviation given by equation (13). Therefore, the first term of equation (12) is simply given by z^* with $z^* \sim \mathcal{N}(0, \sigma^{*2})$. The sum of n normally distributed random numbers with variance one, results for large n in the χ^2-distribution. Since the standard deviation only depends on $\sqrt{2n}$, we neglect the "randomness" of the second term and replace it by its average value given in equation (14). Thus, we write

$$\Delta Q \approx z^* - \sigma^2 \sum_{k=1}^{n} d_k; \quad z^* \sim \mathcal{N}(0, \sigma^{*2}) \tag{15}$$

The probability for success is given by:

$$\begin{aligned} \Pr\left[\Delta Q > 0\right] &= \Pr\left[z^* \geq \sigma^2 \sum_{k=1}^{n} d_k\right] \\ &= \frac{1}{\sqrt{2\pi}\sigma^*} \int_{\sigma^2 \sum d_k}^{\infty} \exp\left(-\frac{z^{*2}}{2\sigma^{*2}}\right) dz^* \end{aligned} \tag{16}$$

Now we substitute $p = \frac{z^*}{\sqrt{2}\sigma^*}$ and $dz^* = \sqrt{2}\sigma^* \, dp$:

$$\Pr\left[\Delta Q > 0\right] = \frac{1}{\sqrt{2\pi}\sigma^*} \sqrt{2}\sigma^* \int_{\frac{\sigma^2 \sum d_k}{\sigma^* \sqrt{2}}}^{\infty} e^{-p^2} \, dp \tag{17}$$

Using the error function $erf(x)$:

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-p^2} \, dp, \tag{18}$$

and $\sigma^* = \sigma\sqrt{\sum c_k^2}$ we can write

$$\Pr\left[\Delta Q > 0\right] = \frac{1}{\sqrt{\pi}} \left(\int_0^{\infty} e^{-p^2} \, dp - \int_0^{\frac{\sigma \sum d_k}{\sqrt{2 \sum c_k^2}}} e^{-p^2} \, dp \right)$$

$$= \frac{1}{2}\left(1 - \mathrm{erf}\left(\frac{\sigma \sum_{k=1}^{n} d_k}{\sqrt{2 \sum_{k=1}^{n} c_k^2}}\right)\right) \tag{19}$$

ES Theory: Progress rate

The progress rate ϕ' is defined as follows:

$$\phi' = \frac{\Delta Q}{\tan \alpha}, \tag{20}$$

where ΔQ is given by equation (14). Since $tan\, \alpha$ is the gradient at the parent position, i.e. at $y_k = 0;\ k = 1, \ldots, n$, we have

$$\tan \alpha \; = \; \left[\sum_{k=1}^{n} \left(\frac{\partial Q}{\partial y_k}\right)^2_{y_k = 0}\right]^{\frac{1}{2}} \; = \; \left[\sum_{k=1}^{n} c_k^2\right]^{\frac{1}{2}}. \tag{21}$$

Therefore, the progress rate is given by

$$\phi' = \frac{z^*}{\left[\sum_{k=1}^{n} c_k^2\right]^{\frac{1}{2}}} \; - \; \sigma^2 \frac{\sum_{k=1}^{n} d_k}{\left[\sum_{k=1}^{n} c_k^2\right]^{\frac{1}{2}}}. \tag{22}$$

Using the following equation

$$\mathrm{Var}\left[z^*\right] \; = \; \sigma^2 \left(\sum_{k=1}^{n} c_k^2\right) \; = \; \sum_{k=1}^{n} c_k^2 \, \mathrm{Var}\left[z\right] = \mathrm{Var}\left[\left(\sum_{k=1}^{n} c_k^2\right)^{\frac{1}{2}} z\right], \tag{23}$$

we can simplify equation (22):

$$\phi' = z \; - \; \sigma^2 \frac{\sum_{k=1}^{n} d_k}{\left[\sum_{k=1}^{n} c_k^2\right]^{\frac{1}{2}}}. \tag{24}$$

Since the random variable z is symmetric around $\mathrm{E}\left[z\right] = 0$, we see from equation (24) that a positive progress rate can only be achieved with selection. If we assume $(1, \lambda)$ selection, we have to derive the probability density of a new random variable u, which is the largest out of λ random numbers z_i, $i = 1, \ldots, \lambda$ and $z_i = \mathcal{N}(0, \sigma^2)$. Therefore, after selection we have to replace equation (24) by

$$\phi'_{1,\lambda} = u \; - \; \sigma^2 \frac{\sum_{k=1}^{n} d_k}{\left[\sum_{k=1}^{n} c_k^2\right]^{\frac{1}{2}}}. \tag{25}$$

If we denote the probability density of u by $w_\lambda(u)$, the expectation value of $\phi'_{1,\lambda}$ is given by:

$$\phi_{1,\lambda} = \mathrm{E}\left[\phi'_{1,\lambda}\right] = \int_{-\infty}^{\infty} u\, w_\lambda(u)\, du \;-\; \sigma^2 \frac{\sum_{k=1}^n d_k}{\left[\sum_{k=1}^n c_k^2\right]^{\frac{1}{2}}} \tag{26}$$

$$= \sigma\, c_{1,\lambda} \;-\; \sigma^2 \frac{\sum_{k=1}^n d_k}{\left[\sum_{k=1}^n c_k^2\right]^{\frac{1}{2}}}, \tag{27}$$

$$w_\lambda = \frac{\lambda}{\sqrt{2\pi}\sigma} \exp\left(-\frac{u^2}{2\sigma^2}\right) \left\{\frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{u}{\sqrt{2}\sigma}\right)\right]\right\}^{\lambda-1} \tag{28}$$

where we have introduced the progress rate coefficients $c_{1,\lambda}$:

$$c_{1,\lambda} = \frac{\sqrt{2}\lambda}{\sqrt{\pi}2^{\lambda-1}} \int_{-\infty}^{\infty} z e^{-z^2} \left[1 + \mathrm{erf}(z)\right]^{\lambda-1} dz \tag{29}$$

Note, that the formula for the progress rate, equation (27), remains the same even for the more general (μ, λ) selection, only the progress rate coefficients change and are replaced by $c_{\mu,\lambda}$.

ES theory: $1/5$-rule

With the results from the last section, we are now able to determine the step-size that leads to the maximal progress rate:

$$\frac{\partial \phi_{1,\lambda}}{\partial \sigma} = 0 \quad \Rightarrow \quad c_{1,\lambda} = 2\,\sigma \frac{\sum_{k=1}^n d_k}{\left[\sum_{k=1}^n c_k^2\right]^{\frac{1}{2}}}$$

$$\Rightarrow \quad \sigma_{opt} = \frac{c_{1,\lambda}}{2} \frac{\left[\sum_{k=1}^n c_k^2\right]^{\frac{1}{2}}}{\sum_{k=1}^n d_k}. \tag{30}$$

This corresponds to a progress rate of

$$\phi_{1,\lambda,opt} = \frac{1}{4}\, c_{1,\lambda}^2 \,\frac{\left[\sum_{k=1}^n c_k^2\right]^{\frac{1}{2}}}{\sum_{k=1}^n d_k} \tag{31}$$

At the same time the sucess probability for σ_{opt} is given by

$$\Pr\left[\Delta Q > 0\right] = \frac{1}{2}\left(1 - \mathrm{erf}\left(\frac{\sigma_{opt} \sum_{k=1}^n d_k}{\sqrt{2\sum_{k=1}^n c_k^2}}\right)\right)$$

$$= \frac{1}{2}\left(1 - \mathrm{erf}\left(\frac{c_{1,\lambda}}{2\sqrt{2}}\right)\right).$$