Submitted by: Ranji Raj

Assignment 18 (Evolution Strategies)

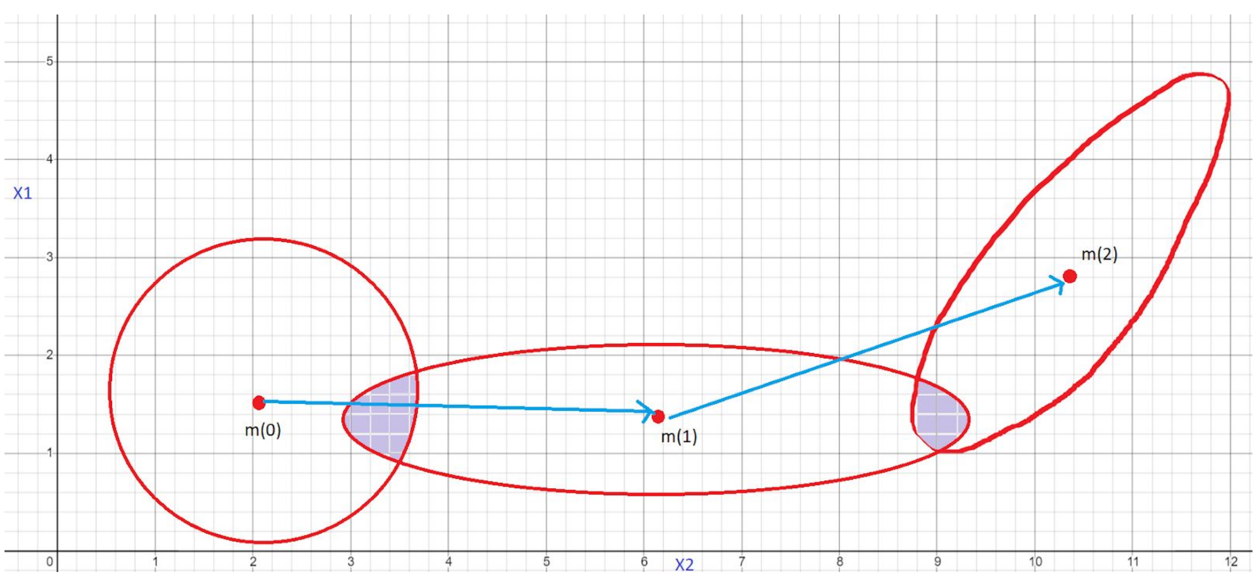
We will now have a look at Evolution Strategies. Make yourself familiar with the algorithm and answer the following questions.

- Explain how evolution strategies work. How are new solutions generated? How is the direction of exploration defined?

- Based on real-valued functions.
- Start with one solution and new solutions are generated by sampling.
- Sampling is done randomly or by a mutation process which focuses on more optimal solutions at a faster pace where it could be more promising.

Working:

- Initially start with only one solution *m(0)* and mutation operator *C(0)* called as the covariance matrix.
- Generate a set of the population by *m(t)* and *C(t)* and perform sampling to create *P(t)*.
- Based on the fitness of the populations (good or bad ones) *P(t)* and *m(t)* update the position for the next iteration.
- Also, update the *C(t)* based on the previous mutation operator and the population size.
- Iterate the procedure back-forth until the termination criteria are met.



Submitted by: Ranji Raj

- The following real-valued function is to be minimized using the covariance matrix adaptation evolution strategy (CMA-ES) algorithm:

$$f(x_1, x_2) = (x_1 - 4)^2 + (x_2 - 2)^2$$

The algorithms considers:

- $\lambda = 4$ and $\mu = 2$,
- the initial covariance matrix being the two-dimensional identity matrix,
- the initial mean vector $m_0 = (0, 3)$,
- the step size $\sigma = 1$,
- a learning rate of 0.2,
- the solution weights $w_1 = 0.7$ and $w_2 = 0.3$

Do not use a normal distribution to sample the offspring of the initial m. Instead, produce four children from the initial solution by going step size into the positive or negative direction of the x_1 or x_2 axis (to take out the randomness factor of the algorithm and make everybody´s solutions comparable).

What is the value of the updated m and what does the covariance matrix look like after one iteration of the algorithm? Which are the directions of the ellipsiod axes? (Hint: Calculate the Eigenvectors of the Covariance Matrix to get information about the directions.)

We have σ =1 (positive direction) σ = -1 (negative direction)

$f(x_1,x_2) = (x_1- 4)^2 + (x_2- 2)^2$

We use $x_i = m + \sigma$

For σ = 1, we have $x_1 = 0 + 1 = 1$ & $x_2 = 3 + 1 = 4$

$f(1,4) = (1 - 4)^2 + (4 - 2)^2 = 13$

For σ = -1, we have $x_1 = 0 - 1 = -1$ & $x_2 = 3 - 1 = 2$

$f(-1,2) = (-1 - 4)^2 + (2 - 2)^2 = 25$

Ranking: 13 < 25

We then update **m** as $m = \sum_{i=1}^{\mu} w_i \, x_{i:\lambda} \quad , \quad \sum_{i=1}^{\mu} w_i = 1$

$w_1 = 0.7$, $W_2 = 0.3$

With w_1, m = 0.7*25 + 0.7*13 = 26.6

With W_2, m = 0.3*25 + 0.3*13 = 11.4

So m_1= (26.6, 11.4)

For the Covariance matrix, C

$C \leftarrow (1 - c_{cov})C + c_{cov}\mu_w y_w y_w^T$

$C = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$\mu_w = \frac{1}{0.3^2+0.7^2} = 1.7241 \geq 1$

We substitute the values,

$C \leftarrow (1 - 0.2) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + (0.2)*(1.72)y_w y_w^T$

$C \leftarrow \begin{bmatrix} 0.8 & 0 \\ 0 & 0.8 \end{bmatrix} + (0.344)y_w y_w^T$

$y_w^1 = 0.7 * 4 + 0.7 * 2 = 4.2$

$y_w^2 = 0.3 * 4 + 0.3 * 2 = 1.8$

$y_w = \begin{bmatrix} 4.2 \\ 1.8 \end{bmatrix}$, $y_w^T = \begin{bmatrix} 4.2 & 1.8 \end{bmatrix}$

$y_w y_w^T = \begin{bmatrix} 17.64 & 7.56 \\ 7.56 & 3.24 \end{bmatrix} \Rightarrow$ Rank 1 matrix , $0.344 * \begin{bmatrix} 17.64 & 7.56 \\ 7.56 & 3.24 \end{bmatrix} = \begin{bmatrix} 6.06 & 2.6 \\ 2.6 & 1.11 \end{bmatrix}$

$$C \leftarrow \begin{bmatrix} 6.86 & 2.6 \\ 2.6 & 1.91 \end{bmatrix}$$

To get directions of the ellipsoid axes we calculate eigenvalues & eigenvectors,

We write the characteristic equation (λI - A)X = 0

$$\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} 6.86 & 2.6 \\ 2.6 & 1.91 \end{bmatrix} = 0$$

$$\Rightarrow \begin{bmatrix} (\lambda - 6.86) & -2.6 \\ -2.6 & (\lambda - 1.91) \end{bmatrix} = 0$$

$\Rightarrow (\lambda - 6.86)(\lambda - 1.91) - 6.76 = 0$

$\Rightarrow \lambda^2 - 8.77\lambda + 6.34 = 0$

$\Rightarrow \lambda_1 = 7.97,\ \lambda_2 = 0.79$

When λ = 7.97,

$$\begin{bmatrix} 1.11 & -2.6 \\ -2.6 & 6.06 \end{bmatrix} = x_1$$

When λ = 0.8,

$$\begin{bmatrix} -6.06 & -2.6 \\ -2.6 & -1.11 \end{bmatrix} = x_2$$