# 1 Optimization Problems

## 1.1 Evolution Strategy

An evolution strategy is a technique that is primarily used for optimization problems. As the name suggests, this technique tries to find the optimal solution for a problem through the methods of evolution: Mutation and selection. Several implementations exist that differ in what they mutate and what they select. But the general process is, it generates a set of candidate solution which it analyzes on the basis of a fitness or an objective function. The proposed solutions that yield the best fitness values are then used to generate the next generation of candidate solutions. This process only ceases until a predefined criteria has been met.

## 1.2 The CMA-ES

One kind of evolution strategies proposes new candidate solutions by randomly sampling from a multivariate normal distributions with $\mu$ and a fixed $\Sigma$. Each generation the current mean is updated based on the best candidates from the current generation. But because the $\Sigma$ is fixed and with that the search radius, one shortcoming is that when $\Sigma$ is inadequately chosen, the search can be rather slow for $\Sigma$ is too small or even worse, the search gets stuck in a local optimum.

The **C**ovariance **M**atrix **A**daption **E**volution **S**trategy (CMA-ES) is a special kind of an evolution strategy that overcomes this issue, since it not only updates the mean $\mu$ every generation but also the covariance matrix $\Sigma$. As is illustrated in Figure 1, this modification allows for a large search radius in the beginning and thus a fast convergence to the optimum and a smaller search radius towards the end for finetuning the found optimum.

The general procedure of the CMA-ES can be presented as follows:

- Create multivariate normal distribution $X \sim \mathcal{N}(\mu, \mathbf{\Sigma})$ (The initial values are usually $\mu_0 = 0$ and $\mathbf{\Sigma}_0 = I$)
- Sample N points from $X$, such that $Y = (y_1, ..., y_N)$ with $y_i \in X \ \forall \ i = 1...N$
- Evaluate all samples from $Y$ with a previously defined fitness function $f$, such that $F = (f(y_1), ..., f(y_N)) \ \forall \ y_i \in Y$
- From $F$ choose the M samples with the best fitness value (i.e. the highest or the lowest) and calculate the new mean $\mu$ and the new covariance matrix $\mathbf{\Sigma}$

This procedure is repeated until a termination criteria has been met, for example a certain amount of generations have passed or a certain threshold was surpassed.

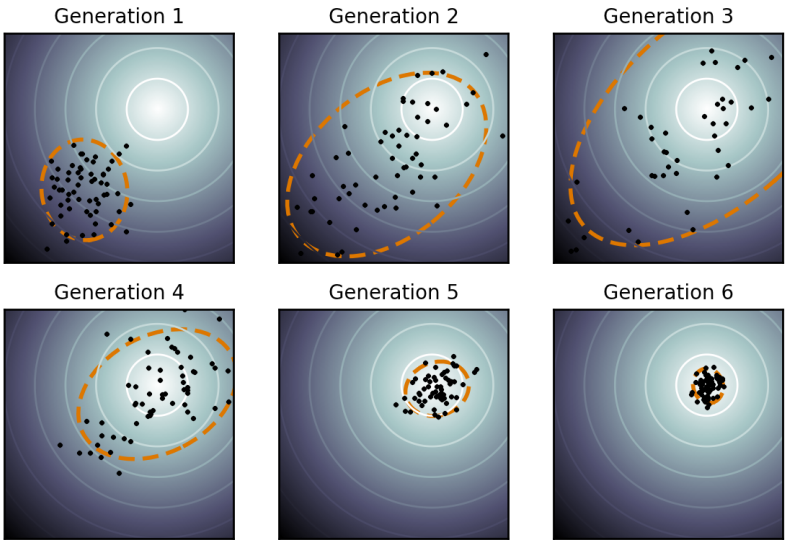


Figure 1: Simple Autoencoder: Platzhalter