Covariance Matrix Adaptation Evolution Strategy

CMA-ES for short

Made by:

- Daniel Holm Johansen
- Silas Larsen
- Villads Rendtorff Zimakoff

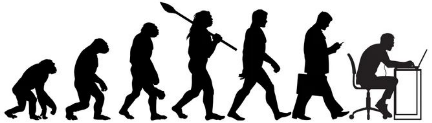
What is CMA-ES?

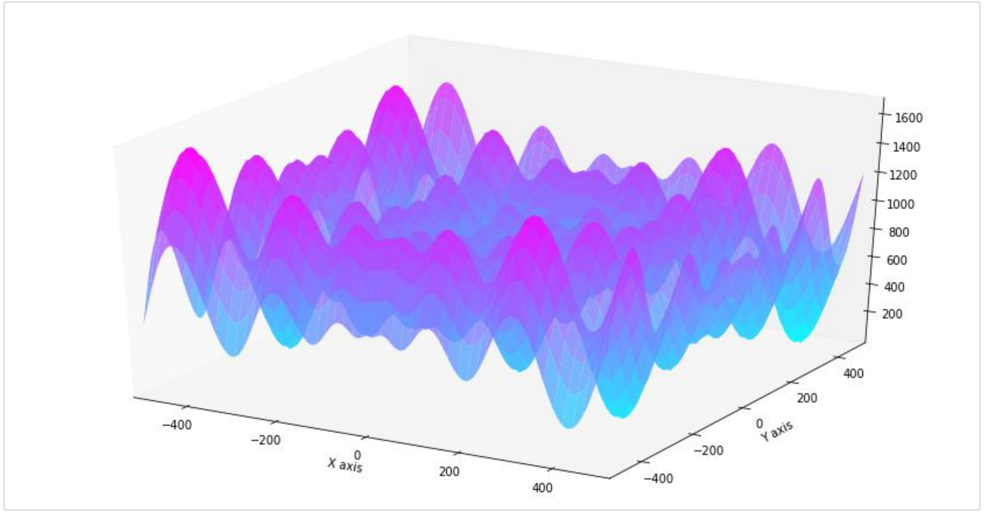
Biological inspiration:

- Mutation / Recombination
- Survival of the fittest
- Adaptation

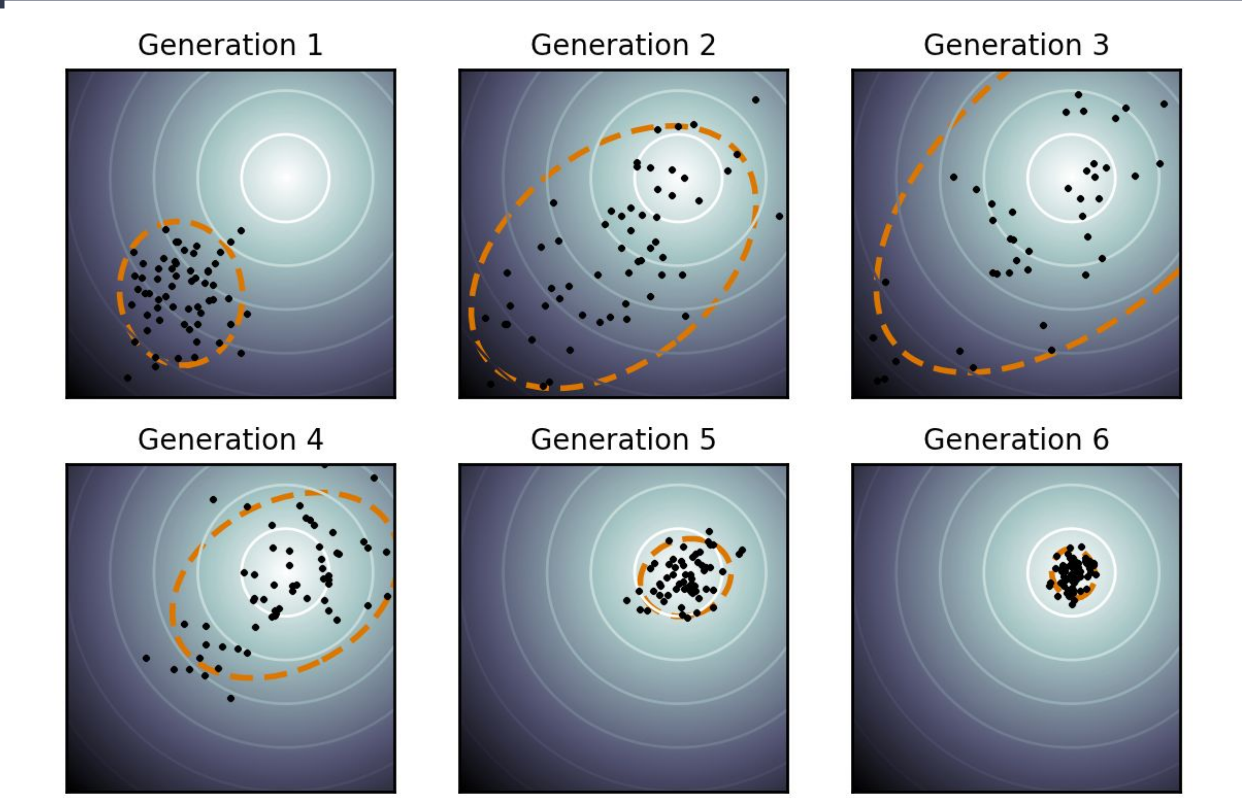
Applications:

- Numerical optimization of complicated functions
- Can be used similar to MLE's



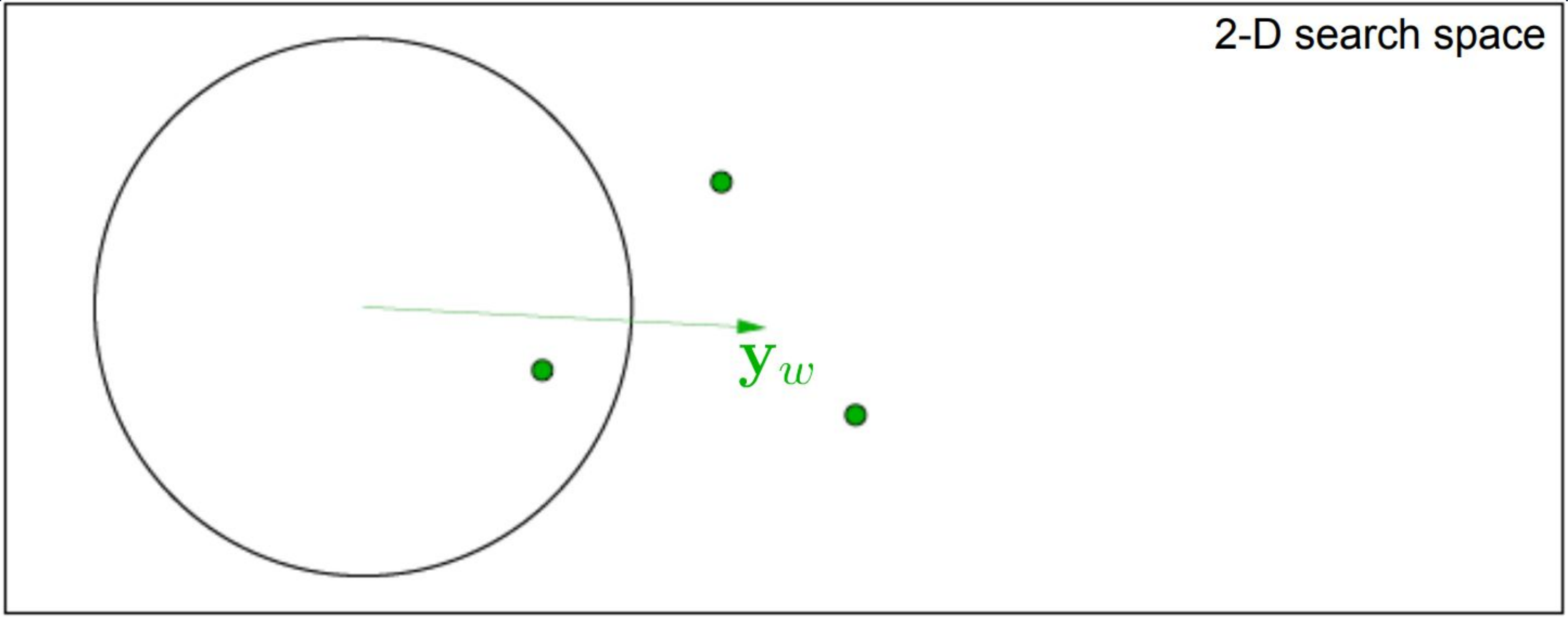


Algorithm general idea



Source: “CMA-ES” by Wikipedia

Moving the mean



Source: Hansen N. “The CMA-ES (and its application to space flight trajectory optimization)”, INRIA 2011.

Moving the mean

$$\boldsymbol{m}^{(g+1)} = \sum_{i=1}^{\mu} w_i \, \boldsymbol{x}_{i:\lambda}^{(g+1)}$$

$$\sum_{i=1}^{\mu} w_i = 1, \qquad w_1 \geq w_2 \geq \cdots \geq w_\mu > 0$$

Source: Hansen N. “The CMA Evolution Strategy: A Tutorial”, INRIA 2023

Moving the mean

$$\boldsymbol{m}^{(g+1)} = \boldsymbol{m}^{(g)} + c_{\mathrm{m}} \sum_{i=1}^{\mu} w_i \left(\boldsymbol{x}_{i:\lambda}^{(g+1)} - \boldsymbol{m}^{(g)} \right)$$

Source: Hansen N. “The CMA Evolution Strategy: A Tutorial”, INRIA 2023

Understanding the Covariance

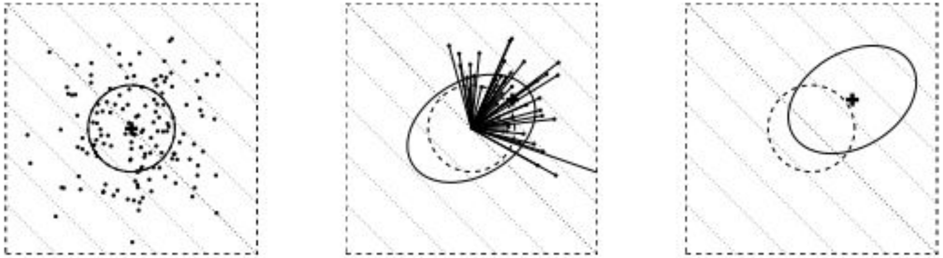
$$\boldsymbol{C}^{(g+1)} = (1 \underbrace{- c_1 - c_\mu \sum w_j}_{\text{can be close or equal to } 0}) \boldsymbol{C}^{(g)} + c_1 \underbrace{\boldsymbol{p}_c^{(g+1)} \boldsymbol{p}_c^{(g+1)^{\mathsf{T}}}}_{\text{rank-one update}} + c_\mu \underbrace{\sum_{i=1}^{\lambda} w_i \, \boldsymbol{y}_{i:\lambda}^{(g+1)} \left(\boldsymbol{y}_{i:\lambda}^{(g+1)}\right)^{\mathsf{T}}}_{\text{rank-}\mu\text{ update}} \tag{30}$$

Source: Hansen N. "The CMA Evolution Strategy: A Tutorial", INRIA 2023

Covariance:

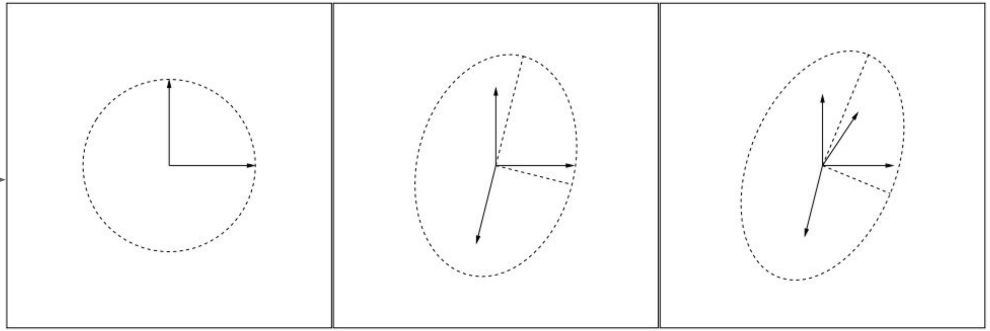
1. Alters the form of the distribution (rank-μ update)
2. Aligns the distribution to the best points (rank-one update)
3. Applies the new information based on c_1 and c_μ

Aligning and deforming



Rank-μ update
(Deforming)

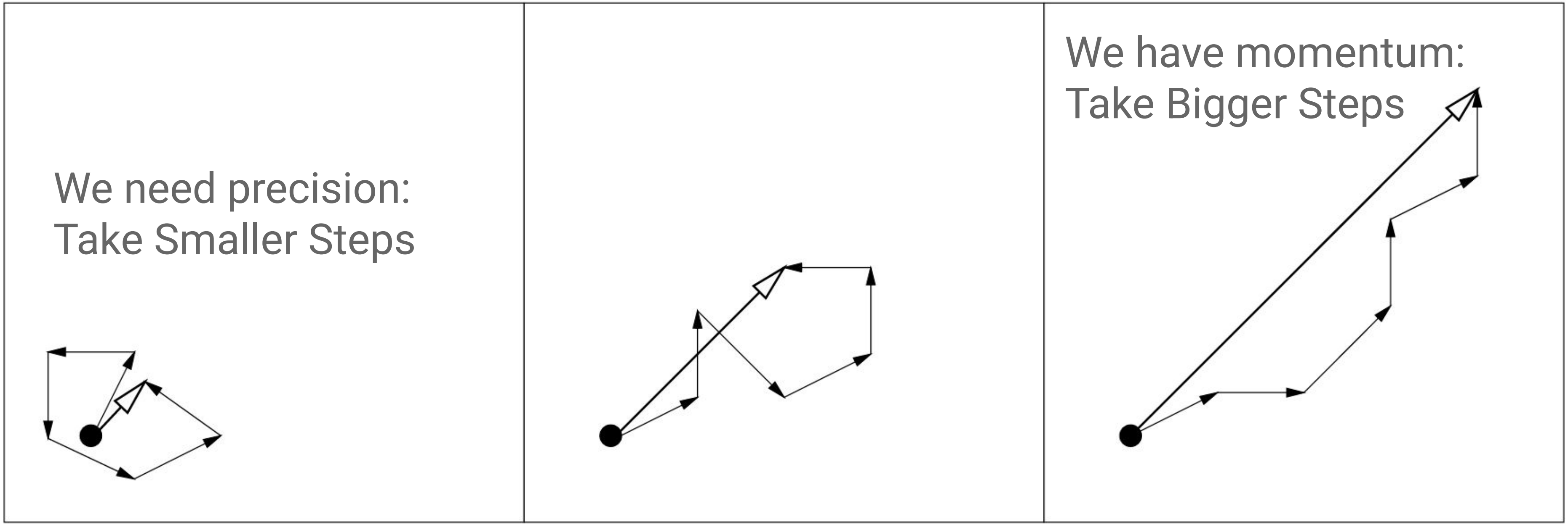
Rank-one update
(aligning)



$\mathcal{N}(\mathbf{0}, \boldsymbol{C}^{(0)})$ $\mathcal{N}(\mathbf{0}, \boldsymbol{C}^{(1)})$ $\mathcal{N}(\mathbf{0}, \boldsymbol{C}^{(2)})$

Controlling the step size

Evolution Path $\boldsymbol{p}_\sigma$:



Source: Hansen N. “The CMA Evolution Strategy: A Tutorial”, INRIA 2023

Controlling the step size

$$\sigma^{(g+1)} = \sigma^{(g)} \exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{\|\boldsymbol{p}_\sigma^{(g+1)}\|}{\mathsf{E}\|\mathcal{N}(\mathbf{0},\mathbf{I})\|} - 1\right)\right)$$

Source: Hansen N. “The CMA Evolution Strategy: A Tutorial”, INRIA 2023

Conclusion

CMA-ES Is Not Perfect But:

- Powerful in bumpy landscapes, good at avoiding local minima
- Can be tuned for specific problems
- Efficient in higher dimensions:
    - CMA-ES: N(dim) = lambda
    - Raster Scan: N(dim) ~ x^{dim}