Derivative Free Optimization and Evolutionary Strategies

Exercise 1: Coordinate Descent I
Minimize Ridge regression, i.e.,

$$\min_{\boldsymbol{\theta}} \frac{1}{2}\|\mathbf{X}\boldsymbol{\theta} - \mathbf{y}\|_2^2 + \frac{\lambda}{2}\|\boldsymbol{\theta}\|_2^2$$

for $\lambda \geq 0$ via coordinate descent under the assumption that $\mathbf{X}^\top \mathbf{X} = \mathbf{I}_d$.

Exercise 2: Coordinate Descent II
Consider the function

$$g : \mathbb{R}^2 \to \mathbb{R}, (x_1, x_2) \mapsto |x_1 - x_2| + 0.1(x_1 + x_2).$$

(a) Perform one round of coordinate descent starting from an arbitrary point (x_1, x_2). Show that after updating x_1 (while fixing x_2) and then updating x_2 (while fixing x_1) the algorithm arrives at a point where $x_1 = x_2$ and terminates. That is, show that coordinate descent will not move beyond the first iteration.

(b) Show that the global infimum of g is $-\infty$. Conclude that coordinate descent fails to find the true minimizer for this function.

Exercise 3: CMA-ES

Assume we have drawn the current population $\mathbf{x}_{1:\lambda}$ from the bivariate Gaussian distribution $\mathcal{N}(\mathbf{m}^{[0]}, \mathbf{C}^{[0]})$ with $\mathbf{m}^{[0]} = (1, 1)^\top, \mathbf{C}^{[0]} = \mathbf{I}$, such that

Id	x_1	x_2	Fitness value
1	1.14	0.24	0.67
2	1.54	-0.86	0.41
3	2.1	2.16	0.09
4	1.5	2.69	0.09
5	1.25	0.51	0.47
6	0.92	2.19	0.15

We want to do a simplified CMA-ES update step:

- Assume the parent number $\mu = 3$.
- Find $\mathbf{m}^{[1]}$ by updating $\mathbf{m}^{[0]}$ in the mean weighted[1] direction of $\mathbf{x}_{1:\mu}$ with stepsize 0.5.
- Compute $\mathbf{C}_\mu$, the (unweighted) sample covariance of $\mathbf{x}_{1:\mu}$ w.r.t. $\mathbf{m}^{[0]}$, and compute

$$\mathbf{C}^{[1]} = (1 - c) \cdot \mathbf{C}^{[0]} + c \cdot \mathbf{C}_\mu$$

  with $c = 0.1$.

[1]Simply scale the fitness values such that they sum up to one.