

Multivariate Optimization 2

Exercise 1: Gradient Descent

A radial basis function (RBF) network has been fitted to a unknown blackbox function resulting in a model

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \mathbf{x} \mapsto \sum_{i=1}^2 w_i \cdot \rho(\|\mathbf{x} - \mathbf{c}_i\|_{S_i})$$

with $\mathbf{c}_1 = (-1.1, 1.1)^\top$, $\mathbf{c}_2 = (0.7, -0.7)^\top$, quartic (biweight) kernel function

$$\rho : \mathbb{R} \rightarrow \mathbb{R}, u \mapsto \begin{cases} (1 - u^2)^2 & |u| < 1 \\ 0 & \text{otherwise} \end{cases}, w_1 = 1, w_2 = -1 \text{ and Mahalanobis distance } \|\cdot\|_{S_i} \text{ with covariance matrices}$$

$$S_1 = \mathbf{I} \text{ and } S_2 = \begin{pmatrix} 1.1 & -0.9 \\ -0.9 & 1.1 \end{pmatrix}.$$

The Mahalanobis distance is given by $\|\mathbf{x} - \mathbf{c}\|_S = (\mathbf{x} - \mathbf{c})^\top S^{-1}(\mathbf{x} - \mathbf{c})$.

(Note: We chose the kernel function and the distance measure for educational purposes; often, a Gaussian kernel and the Euclidean distance are used in practice.)

- (a) Plot f in the range $[-2, 2] \times [-2, 2]$
- (b) Show that $\cap_{i=1}^2 \{\mathbf{x} \in \mathbb{R}^2 \mid \rho(\|\mathbf{x} - \mathbf{c}_i\|_{S_i}) \neq 0\} = \emptyset$.
- (c) Find the global minimum of f analytically.
Hint: b)
- (d) Write an R script which computes two gradient descent steps starting at $x^{[0]} = (-0.45, 0.5)^\top$ with step size $\alpha = 0.15$. What do you observe?
- (e) Perform analytically two gradient descent steps starting at $x^{[0]} = (-0.45, 0.5)^\top$ with step size $\alpha = 0.15$.
- (f) Write an R script which finds the global minimum with the settings in e) but with momentum. (Set $\nu^{[0]} = (0.4, -0.4)^\top$, $\varphi = 0.5$ and stop after 15 iterations.)