

Exercises

§3.5: Gradient Descent

Python: There is an optional starter script that supports Python-based problem(s) on this sheet. Use it as scaffolding: read the comments, complete the lines marked **TODO**, and compare the output with the mathematics requested in the problem(s).

Problem 1 Let $g(w) = (w - 5)^2$. Starting from $w_0 = 1$ with $\alpha = 0.1$, find w_1 and w_2 . Compute the cost at w_0 , w_1 , and w_2 .

Problem 2 Let $g(w_1, w_2) = w_1^2 + 2w_2^2$. Starting from $\mathbf{w}_0 = (2, -1)$ with $\alpha = 0.1$, find $\mathbf{w}_1$.

Problem 3 Suppose $\nabla g(\mathbf{w}) = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$ and $\alpha = 0.05$. Find the gradient-descent step $\Delta \mathbf{w}$. Using the local linear approximation, estimate the change in cost.

Problem 4 For $g(w_1, w_2) = (w_1 - 2)^2 + 4(w_2 + 1)^2$, start at $(0, 0)$ with $\alpha = 0.05$. Compute the first two gradient-descent iterates by hand.

Problem 5 Modify the Python gradient-descent loop from class to use $\alpha = 0.05$ and 30 iterations. Print the final parameter vector and final cost.

Problem 6 Why does standard gradient descent usually use $-\nabla g$ rather than always normalizing to $-\frac{\nabla g}{\|\nabla g\|}$?