Exercise Sheet 5

Optimization and Integration

1. Newton's Method in $\mathbb{R}^2$

Consider the function

$$f(x, y) = x^2 + xy + y^2 - 4x - 2y.$$

(i) Compute the gradient $\nabla f(x, y)$ and the Hessian $H(x, y) = \nabla^2 f(x, y)$.

(ii) Starting from $(x_0, y_0) = (0, 0)$, perform one Newton step and state the resulting position.

(iii) How does Newton's method behave for this function? Are there any instabilities? How many steps will it require to converge?

2. Gradient Descent Stability

Consider the function

$$f(x, y) = \frac{1}{2}(x^2 + 4y^2).$$

(i) Write out the gradient descent iteration with step size η.

(ii) Express the updates for x and y separately.

(iii) For what range of η does gradient descent converge to the minimum?

(iv) Besides faster convergence, state an advantage of second-order methods (e.g., Newton's method) over first-order methods (e.g., gradient descent).

(v) Gradient descent with momentum introduces a velocity $\mathbf{v}_k$ that accumulates gradient information:

$$\mathbf{v}_{k+1} = \beta \mathbf{v}_k + \eta \nabla f(\mathbf{x}_k), \qquad \mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{v}_{k+1}.$$

If the gradient were constant, what effective step size does momentum achieve in the long run? How much larger is it than η when $\beta = 0.999$?

3. Numerical Integration

Let $f(x) = 6x(1 - x)$ and consider the integral

$$I = \int_0^1 f(x)\, dx.$$

(i) Compute I analytically.

(ii) Approximate I using the trapezoid rule with $N = 3$ quadrature points, and compute the absolute error.

(iii) Approximate I using Monte Carlo integration with $N = 3$ samples $X_i \sim \text{Uniform}[0, 1]$. What is the standard deviation of this estimator? You may use $\int_0^1 f(x)^2\, dx = \frac{6}{5}$.

(iv) State two characteristic differences between quadrature and Monte Carlo integration when integrating over the d-dimensional unit cube $[0, 1]^d$.