CS5020: Nonlinear Optimisation: Theory and Algorithms
Coding Exercise - 4 (5 Marks)

For the Rosenbrock function $f(x) = 100(x(2) - x(1)^2)^2 + (x(1) - 1)^2$, run the following algorithms for $t = 0, \ldots, T$ (choose different values of $T = 10, 100, 1000, 10000$ etc):

Gradient Descent

$$x_{t+1} = x_t - \alpha \nabla_x f(x)$$

Heavy Ball Method: Choose $\beta = 0.9$

$$x_{t+1} \quad = x_t + \beta(x_t - x_{t-1}) - \alpha \nabla f(x_t)$$

Nesterov's Accelerated Gradient Method

$$x_{t+1} \quad = x_t + \beta(x_t - x_{t-1}) - \alpha \nabla f\left(x_t + \beta(x_t - x_{t-1})\right)$$

Newton Method

$$x_{t+1} = x_t - \left(\nabla_x^2 f(x_t)\right)^{-1} \nabla_x f(x)$$

Quasi Newton Method

$$x_{t+1} = x_t - \alpha(\text{diagonal}\left(\nabla_x^2 f(x_t)\right)^{-1} \nabla_x f(x))$$

RMSProp: Choose $\beta = 0.9$

$$v_{t+1}(i) = \beta v_t(i) + (1 - \beta)\left[(\nabla_x f(x))(i)\right]^2$$
$$x_{t+1}(i) = x_t(i) - \frac{\alpha}{\sqrt{v_{t+1}(i)}}\left[\nabla_x f(x)\right](i)$$

Adam: Choose $\epsilon = 0.1$, choose $\beta_1 = \beta_2 = 0.9$.

$$m_{t+1} = \beta_1 m_t + (1 - \beta_1)\nabla_x f(x_t)$$
$$v_{t+1}(i) = \beta_2 v_t(i) + (1 - \beta_2)\left[(\nabla_x f(x))(i)\right]^2$$
$$m = \frac{m_{t+1}}{1 - \beta_1^{t+1}}$$
$$v(i) = \frac{v_{t+1}(i)}{1 - \beta_2^{t+1}}$$
$$x_{t+1}(i) = x_t(i) - \frac{\alpha m(i)}{\sqrt{v(i)} + \epsilon}$$

(a) Plot $||x_t - x_*||$ and $|f(x) - f(x_*)|$ as a function of t. (b) 3D-plot of the Rosenbrock function and in the same figure show $x_0, \ldots, x_T$.