Workshop 7 – Deep Learning – Stochastic Gradient Descent

Advanced Analytics and Applications [AAA]

Information Systems for Sustainable Society (is3) | WiSo Faculty | Univ.-Prof. Dr. Wolfgang Ketter | 6/3/22

Calculation

Programming

Question 1: Gradient Descent

1. $\nabla(x^2 + y^2 + x) = \begin{pmatrix} 2x + 1 \\ 2y \end{pmatrix}$ Partial Derivative

2. $\nabla(x^2 + y^2 + x) = \begin{pmatrix} 2x + 1 \\ 2y \end{pmatrix} =^{!} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ First Order Cond.

3. $\begin{pmatrix} x* \\ y* \end{pmatrix} = \begin{pmatrix} -1/2 \\ 0 \end{pmatrix}$ Solve for extrema

Task 1): Analytically calculate the minima of the following function. You don't need to verify whether the critical point is a minima using the hessian matrix.

$$x^2 + y^2 + x$$

Question 1: Gradient Descent

1. $$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

2. $$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} - 0.3 * \begin{pmatrix} 2 * (-1) + 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -0.7 \\ 0 \end{pmatrix}$$

3. $$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} -0.7 \\ 0 \end{pmatrix} - 0.3 * \begin{pmatrix} 2 * (-0.7) + 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -0.58 \\ 0 \end{pmatrix}$$

Task 2): Now, for the same function, calculate the minima using the gradient descent algorithm. Select a learning rate of 0.3, (x0, y0) = (-1, 0) as the starting point, and terminate after 2 iterations.

$$x^2 + y^2 + x$$

$$\nabla(x^2 + y^2 + x) = \begin{pmatrix} 2x + 1 \\ 2y \end{pmatrix}$$

Question 1: Gradient Descent

1. $$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

2. $$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} - 1 * \begin{pmatrix} 2 * (-1) + 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

3. $$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - 1 * \begin{pmatrix} 2 * (0) + 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

Task 3): What happens when we set the learning rate to 1?

$$x^2 + y^2 + x$$

$$\nabla(x^2 + y^2 + x) = \begin{pmatrix} 2x + 1 \\ 2y \end{pmatrix}$$

Calculation

Programming

Contact

For general questions and enquiries on **research**, **teaching**, **job openings** and new **projects** refer to our website at www.is3.uni-koeln.de

For specific enquiries regarding this course contact us by sending an email to the **IS3 teaching** address at is3-teaching@wiso.uni-koeln.de