Hessian Matrix

In mathematics, the Hessian matrix or Hessian is a square matrix of second-order partial derivatives of a scalar-valued function, or scalar field. It describes the local curvature of a function of many variables.

In Simple words — The Hessian matrix is a mathematical tool used to calculate the curvature of a function at a certain point in space.

The Hessian matrix plays an important role in many machine learning algorithms, which involve optimizing a given function. The Hessian is nothing more than the gradient of the gradient, a matrix of second partial derivatives.

The Hessian matrix was developed in the 19th century by the German mathematician **Ludwig Otto Hesse** and later named after him. Hesse originally used the term “functional determinants”.

Several Applications

1. Second-derivative Test: For Convex function, the eigen-values of the Hessian matrix defines it local/global optima.

2. In Optimization: Used in large-scale Optimization.

3. To find out the Inflection Point.

4. To find out the Critical Point based on the nature of gradient.

The Hessian matrix will always be a square matrix whose dimension will be equal to the number of variables of the function. For example, if the function has 2 variables, the Hessian matrix will be a 2×2 - dimension matrix.

$$f : R^2 \to R$$

$$H_{(f(x,y))} = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}$$

If the function has n variables, the Hessian matrix will be a $n \times n$ - dimension matrix.

$$\mathbf{H}_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\,\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\,\partial x_n} \\ \frac{\partial^2 f}{\partial x_2\,\partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\,\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n\,\partial x_1} & \frac{\partial^2 f}{\partial x_n\,\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Problem: Compute the Hessian of $f(x, y) = x^3 - 2xy - y^6$ at the point (1, 2).

Solution:

$$H_{(f(x,y))} = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}$$

We need all the second partial derivatives of f, so let's first compute both partial derivatives:

$$f_x(x,y) = \frac{\partial}{\partial x}(x^3 - 2xy - y^6) = 3x^2 - 2y$$

$$f_y(x,y) = \frac{\partial}{\partial y}(x^3 - 2xy - y^6) = -2x - 6y^5$$

With these, we compute all four second partial derivatives:

$$f_{xx}(x,y) = \frac{\partial}{\partial x}(3x^2 - 2y) = 6x$$

$$f_{xy}(x,y) = \frac{\partial}{\partial y}(3x^2 - 2y) = -2$$

$$f_{yx}(x,y) = \frac{\partial}{\partial x}(-2x - 6y^5) = -2$$

$$f_{yy}(x,y) = \frac{\partial}{\partial y}(-2x - 6y^5) = -30y^4$$

The Hessian matrix in this case is a   matrix with these functions as entries:

$$\mathbf{H}f(x,y) = \begin{bmatrix} f_{xx}(x,y) & f_{yx}(x,y) \\ f_{xy}(x,y) & f_{yy}(x,y) \end{bmatrix} = \begin{bmatrix} 6x & -2 \\ -2 & -30y^4 \end{bmatrix}$$

We were asked to evaluate this at the point (1, 2), so we plug in these values:

$$\mathbf{H}f(1,2) = \begin{bmatrix} 6(1) & -2 \\ -2 & -30(2)^4 \end{bmatrix} = \begin{bmatrix} 6 & -2 \\ -2 & -480 \end{bmatrix}$$