University of Anbar
College of Science
Physics Department

The Hessian Matrix

As we have seen, a function $f(x, y)$ of two variables has four different partial derivatives:

$$f_{xx}(x, y), \qquad f_{xy}(x, y), \qquad f_{yx}(x, y), \qquad f_{yy}(x, y).$$

Of course, $f_{xy}(x, y)$ and $f_{yx}(x, y)$ are always equal, so perhaps they shouldn't count as different.

It is convenient to gather all four of these into a single matrix.

The Hessian of $f(x, y)$

The **Hessian matrix** for a twice differentiable function $f(x, y)$ is the matrix

$$Hf = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

Note that the four entries of the Hessian matrix are actually functions of x and y. Thus the Hessian is itself a function

$$Hf(x, y) = \begin{bmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{yx}(x, y) & f_{yy}(x, y) \end{bmatrix}$$

The Hessian Hf is the first example we have seen of a **matrix-valued function**, i.e. a function whose output is a matrix.

Specifically, Hf is a function that takes x and y as input and outputs a 2×2 matrix.

EXAMPLE 1

Compute the Hessian of the function $f(x, y) = x^4 y^2$.

SOLUTION We must compute all of the second partial derivatives of f. The first partial derivatives are

$$f_x(x, y) = 4x^3y^2 \quad \text{and} \quad f_y(x, y) = 2x^4y,$$

so the second partial derivatives are

$$f_{xx}(x, y) = 12x^2y^2, \quad f_{xy}(x, y) = 8x^3y, \quad f_{yx}(x, y) = 8x^3y, \quad f_{yy}(x, y) = 2x^4.$$

Thus

$$Hf(x, y) = \begin{bmatrix} 12x^2y^2 & 8x^3y \\ 8x^3y & 2x^4 \end{bmatrix}.$$

The Hessian generalizes easily to functions of three variables.

The Hessian of $f(x, y, z)$

The **Hessian matrix** for a twice differentiable function $f(x, y, z)$ is the matrix

$$Hf = \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix}$$

Here we have simply placed each derivative in the correct location. For example, $f_{xx}(x, y, z) = 6xz$, so this should be the upper-left entry of the Hessian matrix.

EXAMPLE 2

Compute $Hf(1, 2, 3)$ if $f(x, y, z) = x^3 z + yz^2$.

SOLUTION The first partial derivatives are

$$f_x(x, y, z) = 3x^2 z, \qquad f_y(x, y, z) = z^2, \qquad f_z(x, y, z) = x^3 + 2yz.$$

Thus

$$Hf(x, y, z) = \begin{bmatrix} 6xz & 0 & 3x^2 \\ 0 & 0 & 2z \\ 3x^2 & 2z & 2y \end{bmatrix}.$$

Substituting in $x = 1$, $y = 2$, and $z = 3$ gives

$$Hf(1, 2, 3) = \begin{bmatrix} 18 & 0 & 3 \\ 0 & 0 & 6 \\ 3 & 6 & 4 \end{bmatrix}$$

The Hessian can be thought of as an analog of the gradient vector for second derivatives. In the same way that the gradient ∇f combines all of the first partial derivatives of f into a single vector, the Hessian Hf combines all of the second partial derivatives of f into a single matrix.

Equivalently, a square matrix A is **symmetric** if

$$A = A^T,$$

where A^T denotes the transpose of A.

Note that the Hessian is always a **symmetric matrix**, meaning that the entries of the Hessian are symmetric across its main diagonal. For example, in the Hessian of a two-variable function $f(x, y)$, the two off-diagonal entries are always equal:

$$\begin{bmatrix} f_{xx} & \underline{f_{xy}} \\ \underline{f_{yx}} & f_{yy} \end{bmatrix}$$

In the case of a three-variable function $f(x, y, z)$, there are three pairs of identical entries in the Hessian matrix:

Each red entry of this matrix is equal to the corresponding blue entry.

$$\begin{bmatrix} f_{xx} & \underline{f_{xy}} & \underline{f_{xz}} \\ \underline{f_{yx}} & f_{yy} & \underline{f_{yz}} \\ \underline{f_{zx}} & \underline{f_{zy}} & f_{zz} \end{bmatrix}$$

Second Directional Derivatives

Given a function $f(x, y)$ and a unit vector $\mathbf{u}$, recall that the directional derivative of f in the direction of $\mathbf{u}$ is given by the formula

$$D_{\mathbf{u}} f = \mathbf{u} \cdot \nabla f.$$

As with many kinds of derivatives, the directional derivative $D_{\mathbf{u}} f$ is actually a function:

$$D_{\mathbf{u}} f(x, y) = \mathbf{u} \cdot \nabla f(x, y).$$

This function takes x and y as input and outputs the directional derivative of f in the direction of $\mathbf{u}$ at the point (x, y).

The **second directional derivative** of f in the direction of $\mathbf{u}$ is the directional derivative of the directional derivative:

$$\boxed{D_{\mathbf{u}}^2 f = D_{\mathbf{u}}[D_{\mathbf{u}} f].}$$

Note that $D_{\mathbf{u}}^2 f$ is again a function of x and y.

In the special case where $\mathbf{u}$ is either $\mathbf{i} = \langle 1, 0\rangle$ or $\mathbf{j} = \langle 0, 1\rangle$, the second directional derivative is the same as a second partial derivative:

$$D_{\mathbf{i}}^2 f = \frac{\partial^2 f}{\partial x^2}, \qquad D_{\mathbf{j}}^2 f = \frac{\partial^2 f}{\partial y^2}.$$

EXAMPLE 3

Find the second directional derivative of the function $f(x, y) = 25x^2 y$ in the direction of the unit vector $\mathbf{u} = \langle 3/5, 4/5\rangle$.

SOLUTION Using the formula $D_{\mathbf{u}} f = \mathbf{u} \cdot \nabla f$, we have

$$D_{\mathbf{u}} f(x, y) = \left\langle \frac{3}{5}, \frac{4}{5} \right\rangle \cdot \langle 50xy, 25x^2\rangle = 30xy + 20x^2.$$

Using the same formula again, we get

$$D_{\mathbf{u}}^2 f(x, y) = \left\langle \frac{3}{5}, \frac{4}{5} \right\rangle \cdot \langle 30y + 40x, 30x\rangle = 48x + 18y$$

Here $\langle 30y + 40x, 30x\rangle$ is the gradient of $30xy + 20x^2$.

The Second Directional Derivative and the Hessian

There is a nice formula for the second directional derivative involving the Hessian.

Theorem (Hessian Formula for $D_{\mathbf{u}}^2 f$)

If f is a twice differentiable function of x and y and $\mathbf{u} = \langle a, b \rangle$ is a unit vector, then

$$D_{\mathbf{u}}^2 f = \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}.$$

Note that the product of a row vector, a matrix, and a column vector is a scalar.

Proof. Using the formula $D_{\mathbf{u}} f = \mathbf{u} \cdot \nabla f$, we have

$$D_{\mathbf{u}} f = \langle a, b \rangle \cdot \langle f_x, f_y \rangle = a f_x + b f_y.$$

Taking the directional derivative again gives

$$D_{\mathbf{u}}^2 f = \langle a, b \rangle \cdot \langle a f_{xx} + b f_{xy}, a f_{xy} + b f_{yy} \rangle = a^2 f_{xx} + 2ab f_{xy} + b^2 f_{yy}.$$

Here $\langle a f_{xx} + b f_{xy}, a f_{xy} + b f_{yy} \rangle$ is the gradient of $a f_x + b f_y$.

But

$$\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} a f_{xx} + b f_{xy} \\ a f_{xy} + b f_{yy} \end{bmatrix} = a^2 f_{xx} + 2ab f_{xy} + b^2 f_{yy}$$

as well, so the two sides of the given equation are equal. ■

EXAMPLE 4

Let f be a twice differentiable function, and suppose that

$$Hf(2,3) = \begin{bmatrix} 4 & 7 \\ 7 & 5 \end{bmatrix}.$$

Compute the directional derivative of f at the point $(2,3)$ in the direction of the vector $\mathbf{u} = \langle 0.6, -0.8 \rangle$.

SOLUTION According to the previous theorem,

$$D_{\mathbf{u}}^2 f(2,3) = \begin{bmatrix} 0.6 & -0.8 \end{bmatrix} \begin{bmatrix} 4 & 7 \\ 7 & 5 \end{bmatrix} \begin{bmatrix} 0.6 \\ -0.8 \end{bmatrix} = \begin{bmatrix} 0.6 & -0.8 \end{bmatrix} \begin{bmatrix} -3.2 \\ 0.2 \end{bmatrix} = -2.08.$$

If we think of a unit vector $\mathbf{u} = \langle a, b \rangle$ as a column vector, then the corresponding row vector is the **transpose** of $\mathbf{u}$:

$$\mathbf{u} = \begin{bmatrix} a \\ b \end{bmatrix}, \qquad \mathbf{u}^T = \begin{bmatrix} a & b \end{bmatrix}.$$

Using this notation, we can write our Hessian formula for $D_{\mathbf{u}}^2 f$ as follows:

$$\boxed{D_{\mathbf{u}}^2 f = \mathbf{u}^T (Hf)\, \mathbf{u}}$$

This formula can be thought of as an analog of the formula $D_{\mathbf{u}} f = \mathbf{u} \cdot \nabla f$ for second derivatives.

This version of the formula applies equally well to functions of three variables, or indeed to functions that take any number of variables as input.

- **References**

1. Boas, Mary L. *Mathematical methods in the physical sciences*. John Wiley & Sons, 2006.
2. Arfken George, Hans J. Weber, and F. Harris. "Mathematical Methods for Physicists. A Comprehensive Guide." (2013).