2 Second Derivatives

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

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

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

so the second partial derivatives are

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

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

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

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

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

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.

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.

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:

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

$$A = A^T,$$

where A^T denotes the transpose of A.

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

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

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

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

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:

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

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:

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

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

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.

The Second Derivative Test

In single-variable calculus, there is a simple test to determine whether a given critical point is a local maximum or a local minimum:

Second Derivative Test (Single Variable)

Let $f(x)$ be a twice differentiable function, and let x_0 be a critical point for f.

1. If $f''(x_0) > 0$, then x_0 is a local minimum for f.
2. If $f''(x_0) < 0$, then x_0 is a local maximum for f.

When $f''(x_0) = 0$, the second derivative test is inconclusive.

This test can be generalized to multivariable functions as follows.

Second Derivative Test

Let $f(x,y)$ be a twice differentiable function, and let (x_0, y_0) be a critical point for f.

1. If $Hf(x_0, y_0)$ is positive definite, then (x_0, y_0) is a local minimum for f.
2. If $Hf(x_0, y_0)$ is negative definite, then (x_0, y_0) is a local maximum for f.
3. If $Hf(x_0, y_0)$ is indefinite, then (x_0, y_0) is a saddle point for f.

Though we are only stating this test for the two-variable case, it works for any number of variables.

When $Hf(x_0, y_0)$ is neither positive definite, negative definite nor indefinite, the second derivative test is inconclusive.

The reason that this test works is that the eigenvalues of the Hessian $H = Hf(x_0, y_0)$ are related to the directional second derivatives of f at x_0, y_0. In particular, if $\mathbf{u}$ is an eigenvector for H with eigenvalue λ, then

$$D_{\mathbf{u}} f(x_0, y_0) = \mathbf{u}^T H \mathbf{u} = \mathbf{u}^T \lambda \mathbf{u} = \lambda \mathbf{u}^T \mathbf{u} = \lambda.$$

Here $\mathbf{u}^T\mathbf{u} = 1$ since $\mathbf{u}$ is a unit vector.

That is, the directional derivative of the Hessian in the direction of an eigenvector $\mathbf{u}$ is equal to the corresponding eigenvalue. Thus we expect the eigenvalues of the Hessian to be positive at a local minimum and negative at a local maximum. Moreover, if the Hessian has both positive and negative eigenvalues, the corresponding point must be a saddle point.

It is less obvious that a critical point *must* be a local minimum just because all of the eigenvalues of the Hessian are positive. This argument requires some additional linear algebra that we will not pursue here.

EXAMPLE 5

The function $f(x, y) = x^3 + 2(x - y)^2 - 3x$ has a critical point at $(1, 1)$. Classify this critical point as a local maximum, a local minimum, or a saddle point.

SOLUTION The Hessian of f is

$$Hf(x, y) = \begin{bmatrix} 6x + 4 & -4 \\ -4 & 4 \end{bmatrix}$$

and in particular

$$Hf(1, 1) = \begin{bmatrix} 10 & -4 \\ -4 & 4 \end{bmatrix}$$

The eigenvalues of this matrix are 2 and 12, so $(1, 1)$ is a local minimum.

The eigenvalues add to 14 (the trace) and multiply to 24 (the determinant), so they must be 2 and 12.

EXAMPLE 6

The function $f(x, y) = 6 \cos x + 4x \sin y$ has a critical point at $(0, 0)$. Classify this critical point as a local maximum, a local minimum, or a saddle point.

SOLUTION The Hessian of f is

$$Hf(x, y) = \begin{bmatrix} -6 \cos x & 4 \cos y \\ 4 \cos y & -4x \sin y \end{bmatrix}$$

and in particular

$$Hf(0, 0) = \begin{bmatrix} -6 & 4 \\ 4 & 0 \end{bmatrix}$$

The eigenvalues of this matrix are -8 and 2, so $(0, 0)$ is a saddle point.

The eigenvalues add to -6 (the trace) and multiply to -16 (the determinant), so they must be -8 and 2.

EXERCISES

1–2 ■ Compute the Hessian matrix for the given function f.

1. $f(x, y) = x^2 \sin y$

2. $f(x, y, z) = x^2 y^3 z^4$

3–4 ■ Compute the Hessian matrix for the given function f at the given point P.

3. $f(x, y) = x^3 + 4xy^2;\ P = (2, 3)$

4. $f(x, y, z) = \dfrac{16z}{\sqrt{xy}};\ P = (4, 1, 8)$

5. Let $f(x, y)$ be a twice differentiable function, and suppose that

$$Hf(x, y) = \begin{bmatrix} -2xy\sin(x^2) & \cos(x^2) \\ \cos(x^2) & 0 \end{bmatrix}.$$

Compute $f_{xy}(\sqrt{\pi}, 5)$.

6. Let $f(x, y) = x^3 + x^2y$, and let $\mathbf{u} = \left\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\rangle$.

(a) Find a formula for $D_{\mathbf{u}} f(x, y)$.

(b) Use your formula from part (a) to find a formula for $D_{\mathbf{u}}^2 f(x, y)$.

7. Let $f(x, y)$ be a twice differentiable function, and suppose that

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

Compute $D_{\mathbf{u}}^2 f(2, 3)$, where $\mathbf{u}$ is the unit vector $\mathbf{u} = \dfrac{1}{\sqrt{5}}\langle 1, 2 \rangle$.

8. Let $f(x, y)$ be a twice differentiable function, and suppose that

$$Hf(x, y) = \begin{bmatrix} 0 & \sin(e^y) \\ \sin(e^y) & xe^y\cos(e^y) \end{bmatrix}.$$

Find a formula for $D_{\mathbf{u}}^2 f(x, y)$, where $\mathbf{u}$ is the unit vector $\left\langle \frac{4}{5}, \frac{3}{5} \right\rangle$

9–12 ■ Find all critical points of the given function. (See Section 11.7 of the textbook.)

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

10. $f(x, y) = x^3 - 12xy + 8y^3$

11. $f(x, y) = e^x \cos y$

12. $f(x, y) = e^y(y^2 - x^2)$

13–18 ■ A function and one of its critical points are given. Use the second derivative test to determine whether the critical point is a local maximum, a local minimum, or a saddle point.

13. $f(x, y) = \sin x \cos y;\ P = (\pi/2, 0)$

14. $f(x, y) = \sin x \cos y;\ P = (\pi/2, \pi)$

15. $f(x, y) = \sin x \cos y;\ P = (\pi, \pi/2)$

16. $f(x, y) = 7x^2 + 4xy + 4y^2 - 48x;\ P = (4, -2)$

17. $f(x, y) = 3x^2 + 4\cos(x + y);\ P = (0, 0)$

18. $f(x, y, z) = 3x^2 + (1 + z^2)\cos y;\ P = (0, 0, 0)$

19. Let $f(x, y) = x^3 - 3x^2 - 2y^2$. Find the critical points of f, and classify each critical point as a local maximum, a local minimum, or a saddle point.