2 Second Derivatives

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

\[
 f_{xx}(x, y), \quad f_{xy}(x, y), \quad f_{yx}(x, y), \quad f_{yy}(x, y).
\]

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

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

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

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

Specifically, \( Hf \) is a function that takes \( x \) and \( y \) as input and outputs a \( 2 \times 2 \) 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

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

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

\[
 Hf = \begin{bmatrix}
 f_{xx} & f_{xy} & f_{xz} \\
 f_{yx} & f_{yy} & f_{yz} \\
 f_{zx} & f_{zy} & f_{zz}
\end{bmatrix}
\]
EXAMPLE 2

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

**SOLUTION**

The first partial derivatives are

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

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 $\nabla f$ 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:

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

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

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

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 $u$, recall that the directional derivative of $f$ in the direction of $u$ is given by the formula

$$D_u f = u \cdot \nabla f.$$ 

As with many kinds of derivatives, the directional derivative $D_u f$ is actually a function:

$$D_u f(x, y) = u \cdot \nabla f(x, y).$$

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

The **second directional derivative** of $f$ in the direction of $u$ is the directional derivative of the directional derivative:
\[ D^2_u f = D_u[D_u f]. \]

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

\[ D^2_i f = \frac{\partial^2 f}{\partial x^2}, \quad D^2_j f = \frac{\partial^2 f}{\partial y^2}. \]

**EXAMPLE 3**

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

**SOLUTION**

Using the formula \( D_u f = u \cdot \nabla f \), we have

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

Using the same formula again, we get

\[ D^2_u f(x, y) = \left\langle \frac{3}{5}, \frac{4}{5} \right\rangle \cdot (30y + 40x, 30x) = 48x + 18y. \]

Here \( (30y + 40x, 30x) \) 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 \( D^2_u f \))**

If \( f \) is a twice differentiable function of \( x \) and \( y \) and \( u = \langle a, b \rangle \) is a unit vector, then

\[ D^2_u f = [a \ b] \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} [a \ b]. \]

**Proof.** Using the formula \( D_u f = u \cdot \nabla f \), we have

\[ D_u f = \langle a, b \rangle \cdot \langle f_x, f_y \rangle = af_x + bf_y. \]

Taking the directional derivative again gives

\[ D^2_u f = \langle a, b \rangle \cdot (af_{xx} + bf_{xy}, af_{xy} + bf_{yy}) = a^2 f_{xx} + 2ab f_{xy} + b^2 f_{yy}. \]

But

\[ [a \ b] \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} [a \ b] = [a \ b] \begin{bmatrix} af_{xx} + bf_{xy} \\ af_{xy} + bf_{yy} \end{bmatrix} = a^2 f_{xx} + 2ab f_{xy} + b^2 f_{yy} \]

as well, so the two sides of the given equation are equal. ■
EXAMPLE 4
Let \( f \) be a twice differentiable function, and suppose that
\[
H f(2, 3) = \begin{bmatrix} 4 & 7 \\ 7 & 5 \end{bmatrix}.
\]
Compute the directional derivative of \( f \) at the point \((2, 3)\) in the direction of the vector \( u = \langle 0.6, -0.8 \rangle \).

**SOLUTION**  According to the previous theorem,
\[
D_u^2 f(2, 3) = \begin{bmatrix} 0.6 & -0.8 \end{bmatrix} \begin{bmatrix} 4 & 7 \\ 7 & 5 \end{bmatrix} \begin{bmatrix} 0.6 \\ -0.8 \end{bmatrix} = \begin{bmatrix} -3.2 \\ 0.2 \end{bmatrix} = -2.08.
\]

If we think of a unit vector \( u = \langle a, b \rangle \) as a column vector, then the corresponding row vector is the transpose of \( u \):
\[
u = \begin{bmatrix} a \\ b \end{bmatrix}, \quad u^T = \begin{bmatrix} a \\ b \end{bmatrix}.
\]

Using this notation, we can write our Hessian formula for \( D_u^2 f \) as follows:
\[
D_u^2 f = u^T (H f) u
\]

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 \( f \).

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 \( H f(x_0, y_0) \) is positive definite, then \( (x_0, y_0) \) is a local minimum for \( f \).
2. If \( H f(x_0, y_0) \) is negative definite, then \( (x_0, y_0) \) is a local maximum for \( f \).
3. If \( H f(x_0, y_0) \) is indefinite, then \( (x_0, y_0) \) is a saddle point for \( f \).
The reason that this test works is that the eigenvalues of the Hessian \( H = H f (x_0, y_0) \) are related to the directional second derivatives of \( f \) at \( x_0, y_0 \). In particular, if \( u \) is an eigenvector for \( H \) with eigenvalue \( \lambda \), then
\[
D_u f(x_0, y_0) = u^T H u = u^T \lambda u = \lambda u^T u = \lambda.
\]
That is, the directional derivative of the Hessian in the direction of an eigenvector \( u \) 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.

**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
\[
H f(x, y) = \begin{bmatrix} 6x + 4 & -4 \\ -4 & 4 \end{bmatrix}
\]
and in particular
\[
H f(1, 1) = \begin{bmatrix} 10 & -4 \\ -4 & 4 \end{bmatrix}
\]
The eigenvalues of this matrix are 2 and 12, so \((1, 1)\) is a local minimum.

**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
\[
H f(x, y) = \begin{bmatrix} -6 \cos(x) & 4 \cos(y) \\ 4 \cos(y) & -4x \sin(y) \end{bmatrix}
\]
and in particular
\[
H f(0, 0) = \begin{bmatrix} -6 & 4 \\ 4 & 0 \end{bmatrix}
\]
The eigenvalues of this matrix are \(-8\) and 2, so \((0, 0)\) is a saddle point.

**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. \( f(x, y, z) = \frac{16z}{\sqrt{xy}}; \ P = (4, 1, 8) \)
5. Let \( f(x, y) \) be a twice differentiable function, and suppose that
\[
   H f(x, y) = \begin{bmatrix}
   -2xy \sin(x^2) & \cos(x^2) \\
   \cos(x^2) & 0 
   \end{bmatrix}.
\]
Compute \( f_{xy}(\sqrt{\pi}, 5) \).

6. Let \( f(x, y) = x^3 + x^2y \), and let \( u = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \).
   (a) Find a formula for \( D_u f(x, y) \).
   (b) Use your formula from part (a) to find a formula for \( D_{u^2} f(x, y) \).

7. Let \( f(x, y) \) be a twice differentiable function, and suppose that
   \[
   H f(2, 3) = \begin{bmatrix} 7 & 4 \\ 4 & 5 \end{bmatrix}.
   \]
Compute \( D_{\mathbf{u}}^2 f(2, 3) \), where \( \mathbf{u} \) is the unit vector \( \mathbf{u} = \frac{1}{\sqrt{5}} (1, 2) \).

8. Let \( f(x, y) \) be a twice differentiable function, and suppose that
   \[
   H f(x, y) = \begin{bmatrix}
   0 & \sin(e^y) \\
   \sin(e^y) & xe^y \cos(e^y) 
   \end{bmatrix}.
   \]
Find a formula for \( D_{\mathbf{u}}^2 f(x, y) \), where \( \mathbf{u} \) is the unit vector \( \begin{pmatrix} 4/5 \\ 3/5 \end{pmatrix} \).

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. \( f(x, y) = \sin x \cos y; \quad P = (\pi/2, 0) \)
14. \( f(x, y) = \sin x \cos y; \quad P = (\pi/2, \pi) \)
15. \( f(x, y) = \sin x \cos y; \quad P = (\pi, \pi/2) \)
16. \( f(x, y) = 7x^2 + 4xy + 4y^2 - 48x; \quad P = (4, -2) \)
17. \( f(x, y) = 3x^2 + 4 \cos(x + y); \quad P = (0, 0) \)
18. \( f(x, y, z) = 3x^2 + (1 + z^2) \cos y; \quad P = (0, 0, 0) \)
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.