

# Writing Assignment 4

Paul Beggs

April 28, 2025

## 1 Introduction

In multivariable calculus, the second derivative test provides a powerful tool for classifying critical points of a function of two variables, telling us whether a critical point corresponds to a local maximum, a local minimum, or a saddle point. While applying the test itself can be straightforward, understanding *why* it works requires delving into the underlying mathematical principles. The goal of this paper is not merely to state the second derivative test, but to provide a clear and justified derivation of its conditions. We will accomplish this by building our understanding from a simple case and progressively generalizing the results.

We will begin in Section 2 by investigating the behavior of a basic quadratic function of two variables at its critical point at the origin. By analyzing this fundamental case without relying on the second derivative test itself. This will lay the groundwork for understanding the key quantities that determine the nature of a critical point.

In Section 3, we will show how to transform any general twice-differentiable function with a critical point at  $(x_0, y_0)$  into a new function that has its critical point conveniently located at the origin. This transformation allows us to focus our analysis on the behavior near the origin, which simplifies the problem.

Section 4 will connect our work to the concept of Taylor polynomials from Calculus II. We will use the second-order Taylor polynomial to create a quadratic approximation of our transformed function near its critical point at the origin. By leveraging the results from Section 2, we will show how the behavior of this quadratic approximation dictates the behavior of the original function near the critical point.

Finally, in Section 5, we will synthesize the results from the preceding sections to formally state and justify the conditions of the second derivative test for a general function  $f(x, y)$  at a critical point  $(x_0, y_0)$ . Through this process, we aim to provide a thorough explanation that illuminates the *why* behind this essential classification tool.

## 2 Investigating the Quadratic Function

Let's start by looking at a quadratic function of the form  $q(x, y) = ax^2 + bxy + cy^2$ , where we assume that  $a, c \neq 0$  and  $b^2 - 4ac \neq 0$ . We want to look into how  $q$  behaves at its critical point. Remember that a critical point for a function of two variables can be found by finding the points where the gradient is zero. Thus, to find the critical points, we need to find the partial derivatives of  $q$  with respect to  $x$  and  $y$  and set them equal to zero:

$$\begin{aligned}\frac{\partial q}{\partial x} &= \frac{\partial}{\partial x}[ax^2 + bxy + cy^2] = 2ax + by = 0 \\ \frac{\partial q}{\partial y} &= \frac{\partial}{\partial y}[ax^2 + bxy + cy^2] = bx + 2cy = 0\end{aligned}$$

This leaves us with the system of equations:

$$\begin{cases} 2ax + by = 0 \\ bx + 2cy = 0 \end{cases}$$

We can rewrite the first equation as  $y = -\frac{2a}{b}x$  and substitute this into the second equation to get:

$$bx + 2c\left(-\frac{2a}{b}x\right) = 0 \implies \left(b - \frac{4ac}{b}\right)x = 0.$$

This gives us two cases: either  $x = 0$  or  $b - \frac{4ac}{b} = 0$ . If  $x = 0$ , then substituting back into the first equation gives us  $y = 0$ . Thus, we have one critical point at  $(0, 0)$ . If  $b - \frac{4ac}{b} = 0$ , then we have  $b^2 - 4ac = 0$ , which is not allowed by our assumption. Therefore, the only critical point is at  $(0, 0)$ .

Now that we have identified the critical point, our next step is to understand the nature of this point – specifically, whether it corresponds to a local minimum, local maximum, or a saddle point. To analyze the function's behavior—particularly its sign—around this critical point, we will rewrite the expression for  $q(x, y)$  by completing the square with respect to the  $x$  terms and treating  $y$  as a constant:

$$q(x, y) = ax^2 + bxy + cy^2$$

Factor out  $a$  from the terms involving  $x$ :

$$q(x, y) = a\left(x^2 + \frac{b}{a}xy\right) + cy^2$$

The term needed to complete the square for  $x^2 + \frac{b}{a}xy$  is  $\left(\frac{b}{2a}y\right)^2 = \frac{b^2}{4a^2y^2}$ . Thus, we add and subtract this term:

$$q(x, y) = a\left[x^2 + \frac{b}{a}xy + \frac{b^2}{4a^2y^2} - \frac{b^2}{4a^2y^2}\right] + cy^2$$

Now, the first three terms inside the parentheses can be factored as a perfect square:

$$q(x, y) = a\left[\left(x + \frac{b}{2a}y\right)^2 - \frac{b^2}{4a^2}y^2\right] + cy^2$$

Distributing  $a$  so we can add the  $y$ -terms together gives:

$$q(x, y) = a\left(x + \frac{b}{2a}y\right)^2 - \frac{b^2}{4a}y^2 + cy^2$$

Now, we can combine the  $y^2$  terms:

$$q(x, y) = a\left(x + \frac{b}{2a}y\right)^2 + \left(c - \frac{b^2}{4a}\right)y^2$$

Find common denominator for the  $y^2$  terms:

$$q(x, y) = a\left(x + \frac{b}{2a}y\right)^2 + \left(\frac{4ac - b^2}{4a}\right)y^2$$

Finally, we factor out  $a$  from the entire expression:

$$q(x, y) = a\left[\left(x + \frac{b}{2a}y\right)^2 + \left(\frac{4ac - b^2}{4a^2}\right)y^2\right].$$

This process of completing the square allows us to rewrite the quadratic function in a form where the signs of the critical components ( $a$  and  $4ac - b^2$ ) can be analyzed more easily. The expression now clearly shows how the function behaves around the critical point  $(0, 0)$ . In short, it directly showed *why*  $a$  and  $4ac - b^2$  are the quantities that matter for determining the nature of the critical point.

For sake of simplicity and notation, we will let  $D = 4ac - b^2$ . Our goal now is to go through the three cases for  $D$  and  $a$  to determine the nature of the critical point at  $(0, 0)$ . These cases are outlined on the next page.

## 2.1 Case 1: $D > 0$ and $a > 0$

For this case, we know the following:

- $q(x, y) = a \left[ \left( x + \frac{b}{2a}y \right)^2 + \frac{D}{4a^2}y^2 \right]$ ;
- $q(0, 0) = 0$ ;
- for any  $(x, y) \neq (0, 0)$ ;
- the term  $\left( x + \frac{b}{2a}y \right)^2$  is always non-negative;
- since  $D > 0$  and  $a > 0$ , the term  $\frac{D}{4a^2}y^2$  is also non-negative;
- the sum of two non-negative terms  $\left( x + \frac{b}{2a}y \right)^2 + \frac{D}{4a^2}y^2$  is zero if and only if both terms are zero, which occurs only at  $(0, 0)$ ; and
- since  $a > 0$ , when we multiply the positive sum by  $a$ , we get a positive result.

This leads us to conclude that for  $(x, y) \neq (0, 0)$ ,  $q(x, y) = a \left[ \left( x + \frac{b}{2a}y \right)^2 + \frac{D}{4a^2}y^2 \right] > 0$ . Thus, since  $q(x, y) > q(0, 0)$  for all  $(x, y) \neq (0, 0)$ , we conclude that  $q$  has a local minimum at  $(0, 0)$ .

## 2.2 Case 2: $D > 0$ and $a < 0$

We share a lot of the same reasoning as the previous case. The only difference is that  $a < 0$ . This change in sign will affect our outcome when we multiply the positive value of the sum  $\left( x + \frac{b}{2a}y \right)^2 + \frac{D}{4a^2}y^2$  by  $a$ . That is, for all  $(x, y) \neq (0, 0)$ ,  $q(x, y) = a \left[ \left( x + \frac{b}{2a}y \right)^2 + \frac{D}{4a^2}y^2 \right] < 0$ . Predictably, this results leads us to the conclusion that  $q$  has a local maximum at  $(0, 0)$ .

## 2.3 Case 3: $D < 0$

It's important to note that when  $D < 0$  (and  $y \neq 0$ ), the term  $\frac{D}{4a^2}y^2$  is negative (since  $4a^2$  is always positive). This means that the expression  $\left( x + \frac{b}{2a}y \right)^2 + \frac{D}{4a^2}y^2$  can take on both positive and negative values depending on the values of  $x$  and  $y$ . Since  $(0, 0)$  is a critical point and  $q(0, 0) = 0$ , the tangent plane to the surface  $z = q(x, y)$  at  $(0, 0)$  is the plane  $z = 0$ . A saddle point occurs when the surface crosses this tangent plane, meaning  $q(x, y)$  takes on both positive and negative values arbitrarily close to  $(0, 0)$ . To show this, we can analyze two different paths through the critical point  $(0, 0)$ :

- Consider the path where  $y = 0$ . Along this path,

$$q(x, 0) = a \left( x + \frac{b}{2a}(0) \right)^2 + \left( c - \frac{b^2}{4a} \right) (0)^2 = ax^2 > 0 \text{ for } x \neq 0.$$

Thus,  $q(x, 0)$  has the same sign as  $a$ . Therefore, if  $a > 0$ , then  $q(x, 0) > 0$  for  $x \neq 0$ , or if  $a < 0$ , then  $q(x, 0) < 0$  for  $x \neq 0$ .

- Now, consider the path  $x = -\frac{b}{2a}y$ . Along this path, we have:

$$q \left( -\frac{b}{2a}y, y \right) = a \left[ \left( -\frac{b}{2a}y + \frac{b}{2a}y \right)^2 + \frac{D}{4a^2}y^2 \right] = a \left[ 0 + \frac{D}{4a^2}y^2 \right] = \frac{D}{4a}y^2 < 0 \text{ for } y \neq 0.$$

From this equation, we can see that for  $y \neq 0$ , the sign of  $q(-\frac{b}{2a}, y)$  is determined by the sign of  $\frac{D}{4a}$ . Since  $D < 0$  and  $4a > 0$ , the sign of  $\frac{D}{4a}$  is the opposite of the sign of  $a$ . Thus, if  $a > 0$ , then  $q(-\frac{b}{2a}, y) < 0$  for  $y \neq 0$ , or if  $a < 0$ , then  $q(-\frac{b}{2a}, y) > 0$  for  $y \neq 0$ .

So, along the path  $y = 0$  (except at the origin),  $q(x, y)$  has the sign of  $a$ . Along the path  $x + \frac{b}{2a}y = 0$  (except at the origin),  $q(x, y)$  has the sign of  $-a$ . Since  $a$  and  $-a$  have opposite signs, we can conclude the function takes both positive and negative values around the origin. This confirms that the critical point  $(0, 0)$  is a saddle point when  $D < 0$ .

## 2.4 Conclusion of Quadratic Analysis

In summary, we have shown that the value  $D = 4ac - b^2$  is the determinant of the nature of the critical point  $(0, 0)$  for the quadratic function  $q(x, y) = ax^2 + bxy + cy^2$ . The results match the second derivative test for classifying critical points. We can show this by computing the second partial derivatives of  $q$  at  $(0, 0)$ :

- $q_x = 2ax + by \Rightarrow q_{xx} = \frac{\partial}{\partial x}(2ax + by) = 2a$ ,
- $q_y = bx + 2cy \Rightarrow q_{yy} = \frac{\partial}{\partial y}(bx + 2cy) = 2c$ , and
- $q_{xy} = \frac{\partial}{\partial y}(2ax + by) = b$ .

Then, since none of these partial derivatives depend on  $x$  or  $y$ , when we evaluate them at  $(0, 0)$ , we get the same values. Now, we can compute the discriminant  $D$  for  $q(0, 0)$  using the second partial derivatives:

$$D(0, 0) = q_{xx}(0, 0)q_{yy}(0, 0) - (q_{xy}(0, 0))^2 = (2a)(2c) - b^2 = 4ac - b^2.$$

This is exactly the value of  $D$  that we used for analysis in the previous section.

Now, we can check our previous work by using the second derivative test. Remember that the second derivative test tells us that for critical points  $(x_0, y_0)$ :

- if  $D > 0$  and  $f_{xx}(x_0, y_0) > 0$ , then  $(x_0, y_0)$  is a local minimum,
- if  $D > 0$  and  $f_{xx}(x_0, y_0) < 0$ , then  $(x_0, y_0)$  is a local maximum, or
- if  $D < 0$ , then  $(x_0, y_0)$  is a saddle point.

For our function  $q(x, y)$  at the critical point  $(0, 0)$ , we have:

- $D = 4ac - b^2 > 0$  and  $q_{xx}(0, 0) = 2a > 0$ , our analysis in Case 1 showed a local minimum.
- $D = 4ac - b^2 > 0$  and  $q_{xx}(0, 0) = 2a < 0$ , our analysis in Case 2 showed a local maximum.
- $D = 4ac - b^2 < 0$ , our analysis in Case 3 showed a saddle point.

These results match our previous analysis of the quadratic function  $q(x, y)$ . Thus, we have shown that the second derivative test works in the special case of a quadratic function with a critical point at the origin.

## 3 Investigating Any Twice-Differentiable Function

We have seen how the second derivative test works for a specific type of quadratic function centered at the origin. Now, we want to extend this understanding to any general twice-differentiable function  $f(x, y)$  near one of its critical points. The challenge is that a general function is not as simple as our quadratic example. However, we can use a clever transformation and the idea of a Taylor polynomial to connect the general case back to the simple quadratic one.

Let's consider a general continuous function  $f(x, y)$  with continuous first and second partial derivatives around a critical point  $(x_0, y_0)$ . Since  $(x_0, y_0)$  is a critical point, we know that the first partial derivatives of  $f$  are zero at this point:

$$f_x(x_0, y_0) = 0, \quad f_y(x_0, y_0) = 0.$$

To simplify our analysis and relate it to the quadratic case we just studied, we want to define a new function that has its critical point at the origin  $(0, 0)$  and has a value of zero at that point. We can achieve this by defining the function  $g(x, y)$  as a shifted version of  $f$ :

$$g(x, y) = f(x + x_0, y + y_0) - f(x_0, y_0).$$

This definition effectively shifts the domain so that the point  $(x_0, y_0)$  in the context of  $f$  corresponds to the point  $(0, 0)$  in the context of  $g$ . Subtracting the constant value  $f(x_0, y_0)$  ensures that the function value at this shifted critical point is zero.

### 3.1 Analyzing the Shifted Function $g$ at the Origin

Let's verify the properties of our new function  $g(x, y)$  at the origin  $(0, 0)$ . First, let's check the function value at the origin. By plugging in  $x = 0$  and  $y = 0$  into the definition of  $g$ , we get:

$$g(0, 0) = f(0 + x_0, 0 + y_0) - f(x_0, y_0) = f(x_0, y_0) - f(x_0, y_0) = 0.$$

So,  $g(0, 0) = 0$  as intended.

Next, we need to confirm that the origin  $(0, 0)$  is indeed a critical point for  $g$ . To do this, we must find the partial derivatives of  $g$  and evaluate them at  $(0, 0)$ . Since the inputs to  $f$  in the expression for  $g$  are functions of  $x$  and  $y$ , we will use the chain rule. For clarity in applying the chain rule, let's consider the intermediate variables  $u = x + x_0$  and  $v = y + y_0$ , so  $g(x, y) = f(u(x, y), v(x, y)) - f(x_0, y_0)$ .

To find the partial derivative of  $g$  with respect to  $x$ , we can differentiate  $f(u, v)$  with respect to  $x$ . Remember from Section 4.5 (in the book) that the chain rule for two independent variables is defined as

$$\frac{\partial f}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial x}.$$

Where the constant term  $f(x_0, y_0)$  differentiates to zero, so it is not included. Thus, we have:

$$\frac{\partial u}{\partial x} = \frac{\partial}{\partial x}(x + x_0) = 1, \quad \text{and} \quad \frac{\partial v}{\partial x} = \frac{\partial}{\partial x}(y + y_0) = 0.$$

Thus,

$$g_x(x, y) = \frac{\partial f}{\partial u}(u, v) \cdot 1 + \frac{\partial f}{\partial v}(u, v) \cdot 0 = \frac{\partial f}{\partial u}(u, v).$$

The notation  $\frac{\partial f}{\partial u}(u, v)$  represents the partial derivative of  $f$  with respect to its first variable, evaluated at  $(u, v)$ . This is precisely  $f_x(u, v)$  in our standard notation for  $f$ . Substituting back  $u = x + x_0$  and  $v = y + y_0$ , we find:

$$g_x(x, y) = f_x(x + x_0, y + y_0).$$

Now, we evaluate this at the origin  $(0, 0)$ :

$$g_x(0, 0) = f_x(0 + x_0, 0 + y_0) = f_x(x_0, y_0).$$

Since  $(x_0, y_0)$  is a critical point of  $f$ , we know that  $f_x(x_0, y_0) = 0$ . Therefore,  $g_x(0, 0) = 0$ . An analogous argument applies to differentiating  $g$  with respect to  $y$  using  $\frac{\partial u}{\partial y} = \frac{\partial}{\partial y}(x + x_0) = 0$  and  $\frac{\partial v}{\partial y} = \frac{\partial}{\partial y}(y + y_0) = 1$ . This shows that  $g_y(0, 0) = f_y(x_0, y_0) = 0$ .

Because  $g_x(0, 0) = 0$  and  $g_y(0, 0) = 0$ , the origin  $(0, 0)$  we have identified a critical point for the function  $g(x, y)$ .

### 3.2 Connecting the Behavior of $f$ and $g$

Before we proceed, let's explore the relationship between analyzing  $g$  at  $(0, 0)$  and analyzing  $f$  at  $(x_0, y_0)$ . The function  $g(x, y) = f(x + x_0, y + y_0) - f(x_0, y_0)$  represents the transformation of the surface  $z = f(x, y)$ . Essentially, we have shifted the surface horizontally so that the point  $(x_0, y_0)$  is now located above  $(0, 0)$ , and we have shifted it vertically so the function value at this point is 0. It's important to note that this transformation does not change the shape of the surface around the critical point. That is, if  $f$  had a local minimum at  $(x_0, y_0)$ ,  $g$  will have a corresponding minimum at  $(0, 0)$ . The same follows for if  $f$  had a local maximum or saddle point at  $(0, 0)$ . Therefore, classifying the critical point of  $g$  at  $(0, 0)$  will directly tell us the classification of the critical point of  $f$  at  $(x_0, y_0)$ .

### 3.3 Second Partial Derivatives of $g$ at the Origin

To classify the critical point  $(0,0)$  for  $g$ , we will eventually use a test similar to the second derivative test we saw for quadratics. This requires evaluating the second-order partial derivatives of  $g$  at the critical point  $(0,0)$ .

We already found the first partial derivatives:  $g_x(x, y) = f_x(x + x_0, y + y_0)$  and  $g_y(x, y) = f_y(x + x_0, y + y_0)$ . We need to differentiate these expressions again with respect to  $x$  or  $y$ . Thus, we will use the chain rule once more, with the same intermediate variables  $u = x + x_0$  and  $v = y + y_0$ .

Let's find  $g_{xx}(x, y)$ , the partial derivative of  $g_x$  with respect to  $x$ . Applying the chain rule to  $f_x(u(x, y), v(x, y))$ :

$$g_{xx}(x, y) = \frac{\partial}{\partial x}[f_x(u, v)] = \frac{\partial f_x}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial f_x}{\partial v} \frac{\partial v}{\partial x}.$$

Since  $\frac{\partial u}{\partial x} = 1$  and  $\frac{\partial v}{\partial x} = 0$ , and using Jacobi ( $f_{xu}$ ) for the Leibnitz ( $\frac{\partial f_x}{\partial u}$ ) notation:

$$g_{xx}(x, y) = f_{xu}(u, v) \cdot 1 + f_{xv}(u, v) \cdot 0 = f_{xu}(u, v).$$

Then, because  $u$  is just  $x$  shifted by a constant,  $f_{xu} = f_{xx}$ . Substituting back  $u = x + x_0$  and  $v = y + y_0$ :

$$g_{xx}(x, y) = f_{xx}(x + x_0, y + y_0).$$

Evaluating at the critical point  $(0,0)$ :

$$g_{xx}(0,0) = f_{xx}(0 + x_0, 0 + y_0) = f_{xx}(x_0, y_0).$$

Next, let's find the mixed partial  $g_{xy}(x, y)$ . Applying the chain rule to  $f_x(u(x, y), v(x, y))$ :

$$g_{xy}(x, y) = \frac{\partial}{\partial y}[f_x(u, v)] = \frac{\partial f_x}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial f_x}{\partial v} \frac{\partial v}{\partial y}.$$

Since  $\frac{\partial u}{\partial y} = \frac{\partial}{\partial y}(x + x_0) = 0$  and  $\frac{\partial v}{\partial y} = \frac{\partial}{\partial y}(y + y_0) = 1$ :

$$g_{xy}(x, y) = f_{xu}(u, v) \cdot 0 + f_{xv}(u, v) \cdot 1 = f_{xv}(u, v).$$

Using the same idea from before,  $f_{xv} = f_{xy}$ . Substituting back  $u = x + x_0$  and  $v = y + y_0$ :

$$g_{xy}(x, y) = f_{xy}(x + x_0, y + y_0).$$

Evaluating at  $(0,0)$ :

$$g_{xy}(0,0) = f_{xy}(0 + x_0, 0 + y_0) = f_{xy}(x_0, y_0).$$

Note that Clairaut's Theorem ensures  $g_{yx}(0,0) = f_{yx}(x_0, y_0) = f_{xy}(x_0, y_0) = g_{xy}(0,0)$ , so we only need to calculate this value once.

Finally, let's find  $g_{yy}(x, y)$ . Starting off in the same way as before by applying the chain rule to  $f_y(u(x, y), v(x, y))$ :

$$g_{yy}(x, y) = \frac{\partial}{\partial y}[f_y(u, v)] = \frac{\partial f_y}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial f_y}{\partial v} \frac{\partial v}{\partial y}.$$

Since  $\frac{\partial u}{\partial y} = 0$  and  $\frac{\partial v}{\partial y} = 1$ :

$$g_{yy}(x, y) = f_{yu}(u, v) \cdot 0 + f_{yv}(u, v) \cdot 1 = f_{yv}(u, v).$$

Making the substitutions  $u = x + x_0$ ,  $v = y + y_0$ , and  $f_{yv} = f_{yy}$ :

$$g_{yy}(x, y) = f_{yy}(x + x_0, y + y_0).$$

Evaluating at  $(0, 0)$ :

$$g_{yy}(0, 0) = f_{yy}(0 + x_0, 0 + y_0) = f_{yy}(x_0, y_0).$$

So, we have successfully shown that the values of the second-order partial derivatives of  $g$  evaluated at its critical point  $(0, 0)$  are exactly the same as the values of the second-order partial derivatives of the original function  $f$  evaluated at its critical point  $(x_0, y_0)$ :

$$\begin{aligned} g_{xx}(0, 0) &= f_{xx}(x_0, y_0), \\ g_{xy}(0, 0) &= f_{xy}(x_0, y_0), \text{ and} \\ g_{yy}(0, 0) &= f_{yy}(x_0, y_0). \end{aligned}$$

This connection between the second partial derivatives of  $f$  at  $(x_0, y_0)$  and  $g$  at  $(0, 0)$  will be crucial in Section 4 when we use the Taylor polynomial to approximate  $g$ .

### 3.4 Review and Conclusion

We have found that the function  $g(x, y) = f(x + x_0, y + y_0) - f(x_0, y_0)$  is a shifted version of the function  $f(x, y)$ . The term  $f(x_0, y_0)$  is a constant subtraction, which shifts the entire graph vertically. The change of variables  $(x, y) \rightarrow (x + x_0, y + y_0)$  is a horizontal shift of the coordinate system. Specifically, the point  $(x, y)$  in the domain of  $g$  corresponds to the point  $(x + x_0, y + y_0)$  in the domain of  $f$ .

If  $f$  has a local minimum at  $(x_0, y_0)$ , this means that for points  $(x, y)$  near  $(x_0, y_0)$ ,  $f(x, y) \geq f(x_0, y_0)$ . Let's consider points near  $(0, 0)$  for  $g$ . These points are of the form  $(x, y)$  where  $x$  and  $y$  are small. The corresponding points for  $f$  are  $(x + x_0, y + y_0)$ , which are near  $(x_0, y_0)$ .

If  $f$  has a local minimum at  $(x_0, y_0)$ , then for small  $x, y$ :

$$f(x + x_0, y + y_0) \geq f(x_0, y_0).$$

Subtracting  $f(x_0, y_0)$  from both sides:

$$f(x + x_0, y + y_0) - f(x_0, y_0) \geq 0.$$

This means  $g(x, y) \geq g(0, 0) = 0$  for small  $x, y$ . Thus,  $g$  has a local minimum at  $(0, 0)$ .

Similarly, if  $f$  has a local maximum at  $(x_0, y_0)$ , then for small  $x, y$ :

$$f(x + x_0, y + y_0) \leq f(x_0, y_0).$$

Subtracting  $f(x_0, y_0)$ :

$$f(x + x_0, y + y_0) - f(x_0, y_0) \leq 0.$$

This means  $g(x, y) \leq g(0, 0) = 0$  for small  $x, y$ . Thus,  $g$  has a local maximum at  $(0, 0)$ .

If  $f$  has a saddle point at  $(x_0, y_0)$ , it means  $f$  takes values both greater than and less than  $f(x_0, y_0)$  in any neighborhood of  $(x_0, y_0)$ . This translates directly to  $g(x, y) = f(x + x_0, y + y_0) - f(x_0, y_0)$  taking values both greater than and less than  $g(0, 0) = 0$  in any neighborhood of  $(0, 0)$ . Thus,  $g$  has a saddle point at  $(0, 0)$ .

In summary, the shape of the graph of  $f$  near  $(x_0, y_0)$  is identical to the shape of the graph of  $g$  near  $(0, 0)$ , just shifted. Therefore, the type of critical point (minimum, maximum, or saddle) is preserved by this transformation. Classifying the critical point of  $g$  at  $(0, 0)$  is equivalent to classifying the critical point of  $f$  at  $(x_0, y_0)$ .

## 4 Approximations with Taylor Polynomials

In the previous section, we transformed the problem of classifying the critical point of  $f$  at  $(x_0, y_0)$  into the equivalent problem of classifying the critical point of  $g$  at  $(0, 0)$ , where  $g(0, 0) = 0$  and the second partial derivatives of  $g$  at  $(0, 0)$  match those of  $f$  at  $(x_0, y_0)$ . Now, we need a way to analyze the behavior of a general function like  $g$  near its critical point at the origin.

Recall from Calculus II the idea of using Taylor polynomials to approximate a function near a specific point. A Taylor polynomial is a polynomial that matches the function's value and the values of its derivatives up to a certain order at that point. For points very close to the point of approximation, the Taylor polynomial provides a good approximation of the function's behavior.

For our function  $g(x, y)$ , which has a critical point at  $(0, 0)$  and  $g(0, 0) = 0$ , we can construct its second-order Taylor polynomial centered at the origin. We choose the second-order polynomial because the second derivative test relies on information from the second partial derivatives, which capture the concavity or curvature of the function at the critical point.

The general formula for the second-order Taylor polynomial of  $g(x, y)$  centered at  $(0, 0)$  is:

$$Q(x, y) = g(0, 0) + g_x(0, 0)x + g_y(0, 0)y + \frac{1}{2}g_{xx}(0, 0)x^2 + g_{xy}(0, 0)xy + \frac{1}{2}g_{yy}(0, 0)y^2.$$

From our work in the previous section, we know that  $g(0, 0) = 0$ ,  $g_x(0, 0) = 0$ , and  $g_y(0, 0) = 0$ . Substituting these values into the formula, the constant and linear terms vanish, and the Taylor polynomial simplifies to:

$$Q(x, y) = \frac{1}{2}g_{xx}(0, 0)x^2 + g_{xy}(0, 0)xy + \frac{1}{2}g_{yy}(0, 0)y^2.$$

This simplified second-order Taylor polynomial  $Q(x, y)$  is a quadratic function, very similar in form to the quadratic function  $q(x, y)$  we analyzed in Section 2. Let's confirm that  $Q(x, y)$  matches  $g(x, y)$  in value and its first and second derivatives at the origin.

Evaluating  $Q(x, y)$  at  $(0, 0)$ :

$$Q(0, 0) = \frac{1}{2}g_{xx}(0, 0)(0)^2 + g_{xy}(0, 0)(0)(0) + \frac{1}{2}g_{yy}(0, 0)(0)^2 = 0.$$

This shows that  $Q(0, 0) = g(0, 0) = 0$ .

Now, let's find the first partial derivatives of  $Q(x, y)$ :

$$Q_x(x, y) = \frac{\partial}{\partial x} \left[ \frac{1}{2}g_{xx}(0, 0)x^2 + g_{xy}(0, 0)xy + \frac{1}{2}g_{yy}(0, 0)y^2 \right] = g_{xx}(0, 0)x + g_{xy}(0, 0)y$$

Evaluating  $Q_x$  at  $(0, 0)$ :

$$Q_x(0, 0) = g_{xx}(0, 0)(0) + g_{xy}(0, 0)(0) = 0.$$

Since  $g_x(0, 0) = 0$ , we have  $Q_x(0, 0) = g_x(0, 0)$ . Similarly, for  $Q_y(x, y)$ :

$$Q_y(x, y) = \frac{\partial}{\partial y} \left[ \frac{1}{2}g_{xx}(0, 0)x^2 + g_{xy}(0, 0)xy + \frac{1}{2}g_{yy}(0, 0)y^2 \right].$$

Treating the partials of  $g$  as constants:

$$Q_y(x, y) = 0 + g_{xy}(0, 0)(x) + \frac{1}{2}g_{yy}(0, 0)(2y) = g_{xy}(0, 0)x + g_{yy}(0, 0)y.$$

Evaluating  $Q_y$  at  $(0, 0)$ :

$$Q_y(0, 0) = g_{xy}(0, 0)(0) + g_{yy}(0, 0)(0) = 0.$$

Since  $g_y(0, 0) = 0$ , we have  $Q_y(0, 0) = g_y(0, 0)$ . Thus, the origin  $(0, 0)$  is a critical point for  $Q(x, y)$ , just as it is for  $g(x, y)$ , and their first derivatives match at this point. Now, when we find the second partial derivatives of  $Q$ , because

$$Q(x, y) = \frac{1}{2}g_{xx}(0, 0)x^2 + g_{xy}(0, 0)xy + \frac{1}{2}g_{yy}(0, 0)y^2,$$



its second derivatives are constant:

$$Q_{xx}(x, y) = g_{xx}(0, 0), \quad Q_{xy}(x, y) = g_{xy}(0, 0), \quad Q_{yy}(x, y) = g_{yy}(0, 0).$$

Evaluating each at  $(0, 0)$  agrees with the corresponding second partial of  $g$ .

So, we have shown that  $g$  and its second-order Taylor polynomial  $Q$  have the same function value, the same first partial derivatives, and the same second partial derivatives when evaluated at the origin  $(0, 0)$ . This confirms that  $Q(x, y)$  is an excellent approximation of  $g(x, y)$  in the immediate neighborhood of the origin.

#### 4.1 Classifying the Critical Point of $g$ using $Q$

Since  $Q(x, y)$  is a good approximation of  $g(x, y)$  for points  $(x, y)$  near  $(0, 0)$ , the behavior of  $g$  at its critical point  $(0, 0)$  is expected to be the same as the behavior of  $Q$  at its critical point  $(0, 0)$ . If  $Q$  has a local minimum, maximum, or saddle point at  $(0, 0)$ , we expect  $g$  to have one too.

Now, notice our quadratic approximation  $Q(x, y)$  has the form  $ax^2 + bxy + cy^2$ , where the coefficients are given by:

$$a = \frac{1}{2}g_{xx}(0, 0), \quad b = g_{xy}(0, 0), \quad c = \frac{1}{2}g_{yy}(0, 0).$$

This is extremely important because in Section 2, we rigorously classified the critical point of a quadratic function  $q(x, y) = ax^2 + bxy + cy^2$  based on the sign of  $a$  and the sign of the discriminant  $D = 4ac - b^2$ . Let's apply those results to our Taylor polynomial  $Q(x, y)$ .

The discriminant for  $Q(x, y)$  is:

$$\begin{aligned} 4ac - b^2 &= 4 \left( \frac{1}{2}g_{xx}(0, 0) \right) \left( \frac{1}{2}g_{yy}(0, 0) \right) - (g_{xy}(0, 0))^2 \\ &= g_{xx}(0, 0)g_{yy}(0, 0) - (g_{xy}(0, 0))^2. \end{aligned}$$

Let's call this discriminant  $D_g = g_{xx}(0, 0)g_{yy}(0, 0) - (g_{xy}(0, 0))^2$ .

Based on our analysis of quadratics in Section 2:

- If  $D_g > 0$  and  $a = \frac{1}{2}g_{xx}(0, 0) > 0$  (which means  $g_{xx}(0, 0) > 0$ ), the quadratic  $Q$  has a local minimum at  $(0, 0)$ .
- If  $D_g > 0$  and  $a = \frac{1}{2}g_{xx}(0, 0) < 0$  (which means  $g_{xx}(0, 0) < 0$ ), the quadratic  $Q$  has a local maximum at  $(0, 0)$ .
- If  $D_g < 0$ , the quadratic  $Q$  has a saddle point at  $(0, 0)$ .

Since the behavior of  $g$  at  $(0, 0)$  is predicted by the behavior of  $Q$  at  $(0, 0)$ , these same conditions classify the critical point of  $g$ :

- If  $D_g > 0$  and  $g_{xx}(0, 0) > 0$ , then  $g$  has a local minimum at  $(0, 0)$ .
- If  $D_g > 0$  and  $g_{xx}(0, 0) < 0$ , then  $g$  has a local maximum at  $(0, 0)$ .
- If  $D_g < 0$ , then  $g$  has a saddle point at  $(0, 0)$ .

## 5 Conclusion: The Second Derivative Test

We began this investigation aiming to understand the second derivative test for classifying critical points of a function  $f(x, y)$  at a point  $(x_0, y_0)$ . We accomplished this by following a series of logical steps:

1. We first analyzed a simple quadratic function  $q(x, y)$  centered at the origin and rigorously showed that its critical point classification depends on the signs of the coefficient of  $x^2$  and the discriminant  $4ac - b^2$ .
2. We then showed how to transform any general function  $f(x, y)$  with a critical point at  $(x_0, y_0)$  into a new function  $g(x, y)$  that has a critical point at the origin  $(0, 0)$  with  $g(0, 0) = 0$ . Crucially, we proved that the second partial derivatives of  $g$  at  $(0, 0)$  are identical to those of  $f$  at  $(x_0, y_0)$ . We also established that the type of critical point for  $f$  at  $(x_0, y_0)$  is the same as the type of critical point for  $g$  at  $(0, 0)$ .
3. Next, we constructed the second-order Taylor polynomial  $Q(x, y)$  for  $g(x, y)$  centered at the origin. We showed that  $Q(x, y)$  matches  $g(x, y)$  in value, first derivatives, and second derivatives at  $(0, 0)$ .
4. Finally, we argued that because  $Q(x, y)$  closely approximates  $g(x, y)$  near  $(0, 0)$ , the behavior of  $g$  at its critical point should be the same as the behavior of  $Q$  at its critical point. By applying the classification rules from our initial quadratic analysis (Step 1) to  $Q(x, y)$ , we derived conditions based on the second partial derivatives of  $g$  at  $(0, 0)$  that classify the critical point of  $g$ .

Now, we can bring this all back to the original function  $f(x, y)$  at its critical point  $(x_0, y_0)$ . The classification of  $f$  at  $(x_0, y_0)$  is the same as the classification of  $g$  at  $(0, 0)$ . The classification of  $g$  at  $(0, 0)$  is based on the discriminant

$$D_g = g_{xx}(0, 0)g_{yy}(0, 0) - (g_{xy}(0, 0))^2,$$

and the value  $g_{xx}(0, 0)$ . From Step 2, we know that

$$g_{xx}(0, 0) = f_{xx}(x_0, y_0), \quad g_{xy}(0, 0) = f_{xy}(x_0, y_0), \quad g_{yy}(0, 0) = f_{yy}(x_0, y_0).$$

Substituting these into the discriminant for  $g$ , we get the discriminant for  $f$  at  $(x_0, y_0)$ :

$$D(x_0, y_0) = f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - (f_{xy}(x_0, y_0))^2.$$

The classification rules for the critical point  $(x_0, y_0)$  of  $f$  are therefore determined by the signs of  $D(x_0, y_0)$  and  $f_{xx}(x_0, y_0)$ . This gives us the second derivative test:

- If  $D(x_0, y_0) > 0$  and  $f_{xx}(x_0, y_0) > 0$ , then  $f$  has a local minimum at  $(x_0, y_0)$ . This occurs because the quadratic approximation of  $f$  around  $(x_0, y_0)$  (which is  $Q(x - x_0, y - y_0)$  shifted back and up by  $f(x_0, y_0)$ ) is a paraboloid opening upwards.
- If  $D(x_0, y_0) > 0$  and  $f_{xx}(x_0, y_0) < 0$ , then  $f$  has a local maximum at  $(x_0, y_0)$ . This occurs because the quadratic approximation is a paraboloid opening downwards.
- If  $D(x_0, y_0) < 0$ , then  $f$  has a saddle point at  $(x_0, y_0)$ . This occurs because the quadratic approximation is a hyperbolic paraboloid (saddle shape).

By breaking down the problem into analyzing a simple quadratic, shifting the critical point of a general function, and using the quadratic Taylor approximation, we have explained the mathematical justification for the second derivative test, showing why the signs of  $D$  and  $f_{xx}$  at a critical point are the key to classifying it.