Chapter 14. Partial Derivatives

14.7. Extreme Values and Saddle Points

Definition. Let $f(x, y)$ be defined on a region R containing the point (a, b). Then

1. $f(a, b)$ is a *local maximum* value of f if $f(a, b) \geq f(x, y)$ for all domain points (x, y) in an open disk centered at (a, b).

2. $f(a, b)$ is a *local minimum* value of f if $f(a, b) \leq f(x, y)$ for all domain points (x, y) in an open disk centered at (a, b).

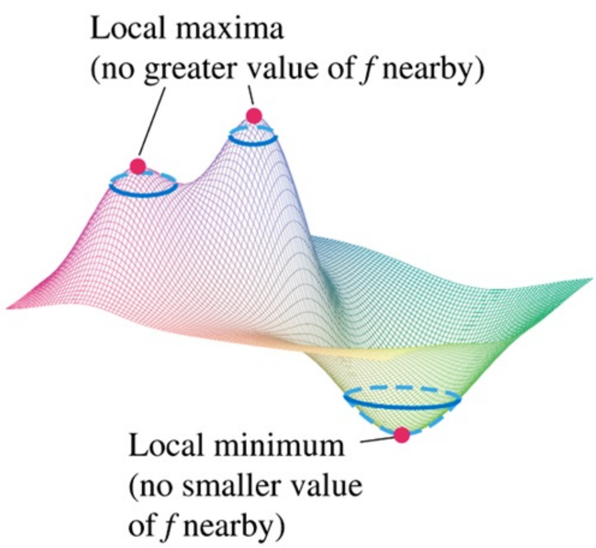


Figure 14.40, Page 821

Theorem 10. First Derivative Test for Local Extreme Values

If $f(x, y)$ has a local maximum or minimum value at an interior point (a, b) of its domain and if the first partial derivatives exist there, then $f_x(a, b) = 0$ and $f_y(a, b) = 0$.

Proof. If f has a local extremum at (a, b), then the function $g(x) = f(x, b)$ has a local extremum at $x = a$. Therefore $g'(a) = 0$. Now $g'(a) = f_x(a, b)$, so $f_x(a, b) = 0$. A similar argument with the function $h(y) = f(a, y)$ shows that $f_y(a, b) = 0$. *Q.E.D.*

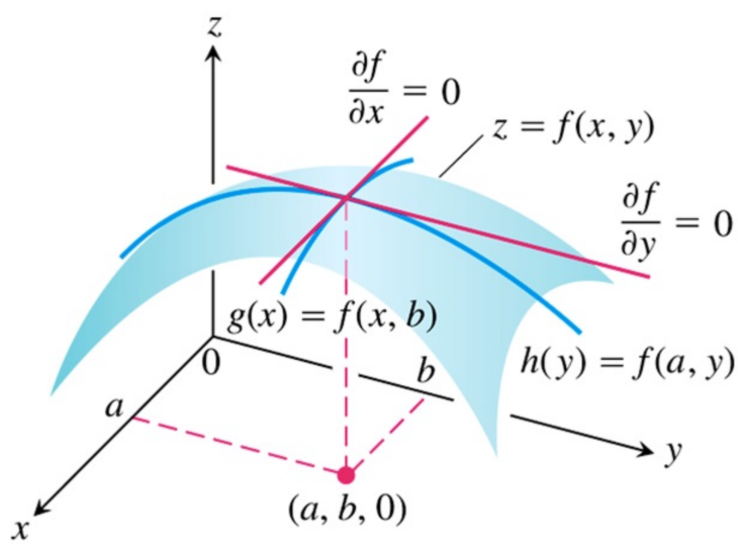
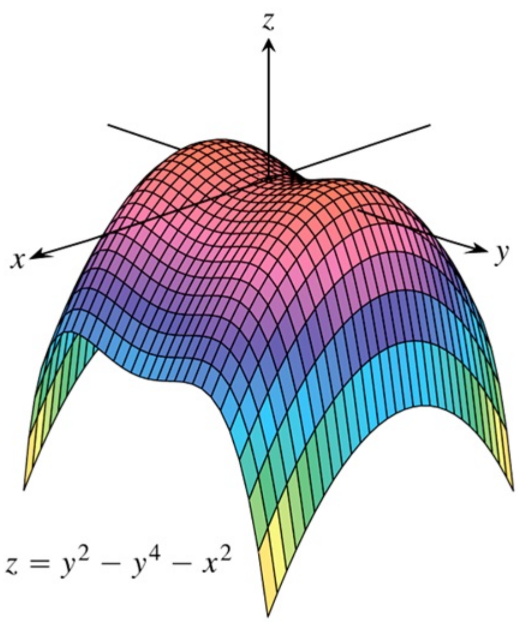
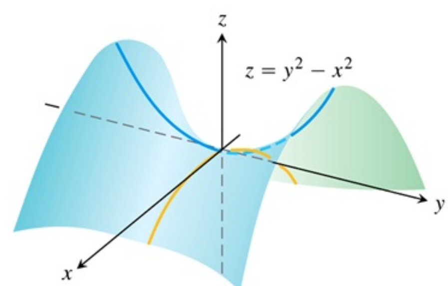


Figure 14.41, Page 821

Definition. An interior point of the domain of a function $f(x, y)$ where both f_x and f_y are zero or where one or both of f_x and f_y do not exist is a *critical point* of f. A differentiable function $f(x, y)$ has a *saddle point* at a critical point (a, b) if in every open disk centered at (a, b) there are domain points (x, y) where $f(x, y) > f(a, b)$ and domain points (s, y) where $f(x, y) < f(a, b)$. The corresponding point $(a, b, f(a, b))$ on the surface $z = f(x, y)$ is called a *saddle point* of the surface.





Figures 14.42 and 14.44, Pages 822 and 823

Example. Page 826, number 6 (find the critical points).

Theorem 11. Second Derivative Test for Local Extreme Values

Suppose that $f(x, y)$ and its first and second partial derivatives are continuous throughout a disk centered at (a, b) and that $f_x(a, b) = f_y(a, b) = 0$. Then

(i) f has a *local maximum* at (a, b) if $f_{xx} < 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at (a, b).

(ii) f has a *local minimum* at (a, b) if $f_{xx} > 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at (a, b).

(iii) f has a *saddle point* at (a, b) if $f_{xx}f_{yy} - f_{xy}^2 < 0$ at (a, b).

(iv) The test is inconclusive at (a, b) if $f_{xx}f_{yy} - f_{xy}^2 = 0$ at (a, b). In this case, we must find some other way to determine the behavior of f at (a, b).

Note. The expression $f_{xx}f_{yy} - f_{xy}^2$ is called the *discriminant* or *Hessian* of f. It is sometimes easier to remember it in determinant form:

$$f_{xx}f_{yy} - f_{xy}^2 = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix}.$$

Theorem 11 makes the most since if we explore the topic of *curvature*! The "Hessian" is really (related to) the curvature of the surface. When curvature is positive at point (a, b) (as in cases (i) and (ii) of Theorem 11), then the surface lies entirely one one side of a tangent plane to the surface at point (a, b) (in some neighborhood of (a, b)). Since at the critical point we have $f_x(a, b) = f_y(a, b) = 0$, then the tangent plane is a horizontal plane so the critical point corresponds to a local maximum if the surface lies above the tangent plane, and the critical point corresponds to a local minimum if the surface lies below the tangent plane. We can determine which is the case be considering the sign of f_{xx} to determine the "concavity" of the surface. A surface is of negative curvature at point (a, b) (as in case (iii) of Theorem 11) if part of the surface lies on one side of the tangent plane to the surface at (a, b) and another part of the surface lies on the other side of the tangent plane (in all open neighborhoods of (a, b)). This is why there is no local extremum in case (iii) (consider Figure 14.40 again.) In case (iv) of Theorem 11, the surface has zero curvature and we cannot determine whether the surface has a local maximum, local minimum, or saddle point at (a, b). For more details on curvature, see my online Differential Geometry (MATH 5310) notes on 1-2. Gauss Curvature and 1-6. The Gauss Curvature in Detail (for a result relevant to Theorem 11, see the example on pages 3 and 4 of the 1-6. The Gauss Curvature in Detail notes).

Example. Page 826, number 6 (again).

Note. Absolute Maxima and Minima of Closed Bounded Regions

We organize the search for the absolute extrema of a continuous function $f(x, y)$ on a closed and bounded region R into three steps:

1. *List the interior points of R* where f may have local maxima and minima and evaluate f at these points. These are the critical points of f.

2. *List the boundary points of R* where f has local maxima and minima and evaluate f at these points. (Details to follow in the next example.)

3. *Look through the lists* for the maximum and minimum values of f. These will be the absolute maximum and minimum values of f on

R. Since absolute maxima and minima are also local maxima and minima, the absolute maximum and minimum values of f appear somewhere in the lists made in Steps 1 and 2.

Example. Page 827, number 32.

Note. Solving the extreme value problems with algebraic constraints on the variables usually requires the method of Lagrange multipliers introduced in the next section. But sometimes we can solve such problems directly.

Example. Page 828, number 58.

Note. Summary of Max-Min Tests

The extreme values of $f(x, y)$ can occur only at

(i) boundary points of the domain of f, and

(ii) critical points (interior points where $f_x = f_y = 0$ or points where f_x or f_y fails to exist).

If the first- and second-order partial derivatives of f are continuous throughout a disk centered at a point (a, b) and $f_x(a, b) = f_y(a, b) = 0$, the nature of $f(a, b)$ can be tested with the Second Derivative Test:

(i) $f_{xx} < 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at $(a, b) \Rightarrow$ local maximum.

(ii) $f_{xx} > 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at $(a, b) \Rightarrow$ local minimum.

(iii) $f_{xx}f_{yy} - f_{xy}^2 < 0$ at $(a, b) \Rightarrow$ saddle point.

(iv) $f_{xx}f_{yy} - f_{xy}^2 = 0$ at $(a, b) \Rightarrow$ test is inconclusive.

Example. Page 828, number 65.

When we try to fit a line $y = mx + b$ to a set of numerical data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, the *method of least squares* to determine the *regression line* is a technique which minimizes the sum of the squares of the vertical distances from the points to the line. This means finding values of m and b that minimize the value of the function

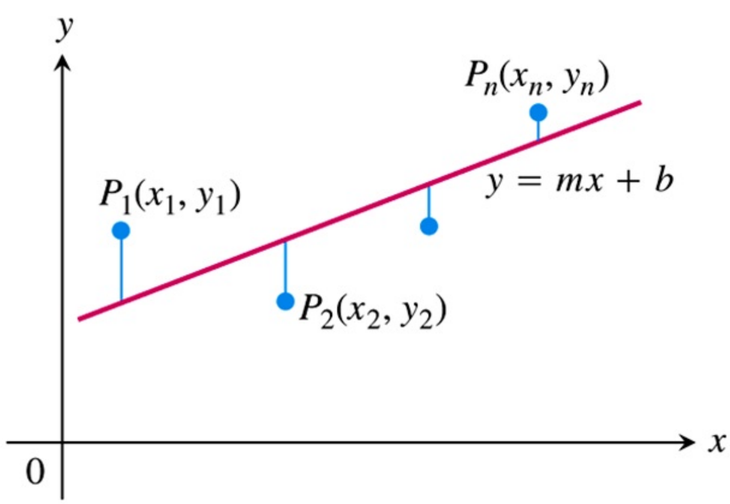
$$w = (mx_1 + b - y_1)^2 + (mx_2 + b - y_2)^2 + \cdots + (mx_n + b - y_n)^2.$$

In this problem we show that the desired values of m and b are:

$$m = \frac{\left(\sum x_k\right)\left(\sum y_k\right) - n\sum x_k y_k}{\left(\sum x_k\right)^2 - n\sum x_k^2},$$

$$b = \frac{1}{n}\left(\sum y_k - m\sum x_k\right)$$

where the sums are all taken for $K = 1$ to $k = n$. This is the usual technique encountered in an introductory statistics class.



Figures 14.48, Pages 828

Solution. We have $w = (mx_1 + b - y_1)^2 + (mx_2 + b - y_2)^2 + \cdots + (mx_n + b - y_n)^2$, or using summation notation,

$$w(m, b) = \sum (mx_k + b - y_k)^2.$$

To find critical points, we consider:

$$\frac{\partial w}{\partial m} = \sum 2(mx_k + b - y_k)[x_k] \text{ and } \frac{\partial w}{\partial b} = \sum 2(mx_k + b - y_k)[1].$$

Expanding, we have

$$\begin{aligned}
\frac{\partial w}{\partial m} &= \sum(2mx_k^2 + 2bx_k - 2x_ky_k) \\
&= \left(\sum 2mx_k^2\right) + \left(\sum 2bx_k\right) - \left(\sum 2x_ky_k\right) \\
&= 2m\sum x_k^2 + 2b\sum x_k - 2\sum x_ky_k.
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial w}{\partial b} &= \sum(2mx_k + b - y_k) = \left(\sum 2mx_k\right) + \left(\sum 2b\right) - \left(\sum 2y_k\right) \\
&= 2m\sum x_k + 2bn - 2\sum y_k.
\end{aligned}$$

Setting each partial derivative equal to 0, gives the following two equations in the two unknowns m and b:

$$\begin{aligned}
m\sum x_k^2 + b\sum x_k &= \sum x_ky_k \\
m\sum x_k + bn &= \sum y_k.
\end{aligned}$$

“Solving” this system of equations for m and b gives:

$$m = \frac{\left(\sum x_k\right)\left(\sum y_k\right) - n\sum x_ky_k}{\left(\sum x_k\right)^2 - n\sum x_k^2},$$

$$b = \frac{1}{n}\left(\sum y_k - m\sum x_k\right).$$

(Notice that we have not "solved" the system in the usual sense, since we only have a formula for b in terms of m.) Next, we need to check a second partial derivative of $w(m, b)$ at the critical point. Notice that

$$\begin{aligned}
\frac{\partial^2 w}{\partial m^2} &= \frac{\partial}{\partial m}\left[\frac{\partial w}{\partial m}\right] \\
&= \frac{\partial}{\partial m}\left[2m\sum x_k^2 + 2b\sum x_k - 2\sum x_k y_k\right] \\
&= 2\sum x_k^2 \\
&> 0,
\end{aligned}$$

and therefore, by the Second Derivative Test for Local Extreme Values, the critical point yields a local minimum. Since there is only one critical point, this must yield an absolute (or "global") minimum.

Revised: 10/31/2020