SIMPLE MULTIVARIATE OPTIMIZATION

1. Definition of local maxima and local minima

1.1. Functions of 2 variables. Let f(x_1, x_2) be defined on a region D in $\Re^2$ containing the point (a, b). Then

a: f(a, b) is a **local maximum** value of f if f(a, b) $\geq$ f(x_1, x_2) for all domain points (x_1, x_2) in an open disk centered at (a, b).
b: f(a, b) is a **local minimum** value of f if f(a, b) $\leq$ f(x_1, x_2) for all domain points (x_1, x_2) in an open disk centered at (a, b).

1.2. Functions of n variables.

a: Consider a real valued function f with domain D in $\Re^n$. Then f is said to have a local minimum at a point $x^* \in$ D if there exists a real number $\delta > 0$ such that

$$f(x) \;\geq\; f(x^*) \;\forall\, x \;\in\; D \text{ satisfying } \|x \;-\; x^*\| \;<\delta \tag{1}$$

b: Consider a real valued function f with domain D in $\Re^n$. Then f is said to have a local maximum at a point $x^* \in$ D if there exists a real number $\delta > 0$ such that

$$f(x) \;\leq\; f(x^*) \;\forall\, x \;\in\; D \text{ satisfying } \|x \;-\; x^*\| \;<\delta \tag{2}$$

c: A real valued function f with domain D in $\Re^n$ is said to have a local extremum at a point $x^* \in$ D if either equation 1 or equation 2 holds.
d: The function f has a global maximum at $x^* \epsilon$ D if 2 holds for all x ϵ D and similarly for a global minimum.

2. Definition of the gradient and Hessian of a function of n variables

2.1. Gradient of f. The gradient of a function of n variables $f(x_1, x_2, \cdots, x_n)$ is defined as follows.

$$\begin{aligned}\nabla f(x) \;&=\; \left(\frac{\partial f}{\partial x_1}\;\frac{\partial f}{\partial x_2}\cdots\frac{\partial f}{\partial x_n}\right)' \\ &= \begin{pmatrix}\frac{\partial f}{\partial x_1}\\ \frac{\partial f}{\partial x_2}\\ \vdots \\ \frac{\partial f}{\partial x_n}\end{pmatrix}\end{aligned} \tag{3}$$

Date: September 1, 2005.

2.2. **Hessian matrix of f.** The Hessian matrix of a function of n variables $f(x_1, x_2, \cdots, x_n)$ is as follows.

$$\nabla^2 f(x) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right] \; i,\, j = 1,\, 2,\, \ldots,\, n$$

$$= \begin{bmatrix} \frac{\partial^2 f(x^0)}{\partial x_1 \partial x_1} & \frac{\partial^2 f(x^0)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x^0)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x^0)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x^0)}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f(x^0)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 f(x^0)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x^0)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x^0)}{\partial x_n \partial x_n} \end{bmatrix} \tag{4}$$

3. Necessary conditions for extreme points

3.1. First derivative test for local extreme values.

Theorem 1. *Let x^* be an interior point of a domain D in R^n and assume that f is twice continuously differentiable on D. It is necessary for a local minimum of f at x^* that*

$$\nabla f(x^*) = 0 \tag{5}$$

This implies that

$$\begin{aligned} \frac{\partial f(x^*)}{\partial x_1} &= 0 \\ \frac{\partial f(x^*)}{\partial x_2} &= 0 \\ &\vdots \\ \frac{\partial f(x^*)}{\partial x_n} &= 0 \end{aligned} \tag{6}$$

Theorem 2. *Let x^* be an interior point of a domain D in R^n and assume that f is twice continuously differentiable on D. It is necessary for a local maximum of f at x^* that*

$$\nabla f(x^*) = 0 \tag{7}$$

3.2. **Graphical illustration.** Figure 1 shows the relative maximum of a function. The plane that is tangent to the surface at the point where the gradient is equal to zero is horizontal in the $f(x_1,x_2)$ plane.

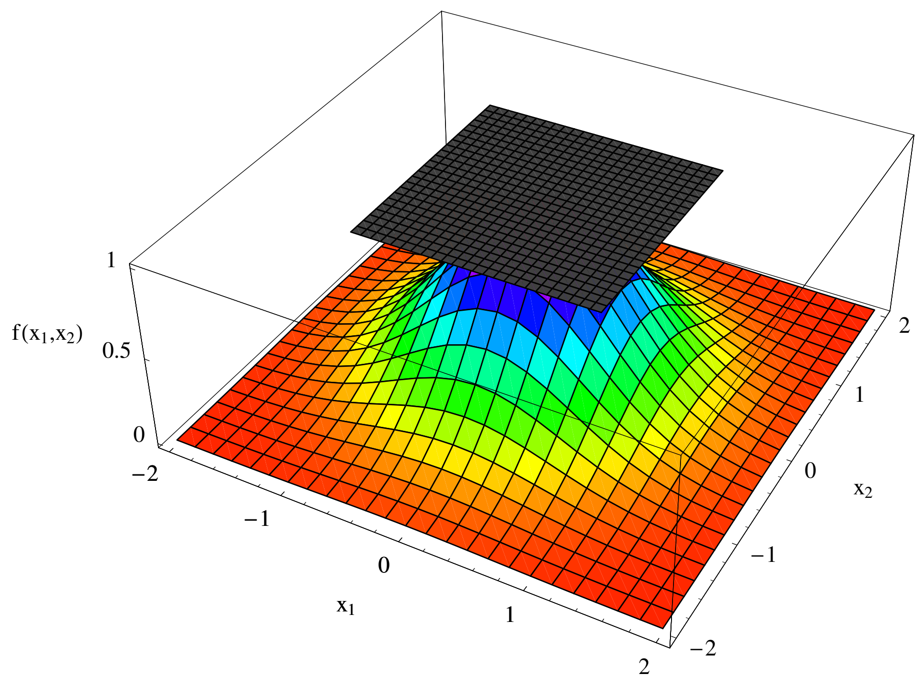
3.3. Critical points and saddle points.

3.3.1. *Critical point.* An interior point of the domain of a function $f(x_1, x_2, \ldots, x_n)$ where all first partial derivatives are zero or where one or more of the first partials does not exist is a **critical point** of f.

3.3.2. *Saddle point.* A critical point that is not a local extremum is called a saddle point. We can say that a differentiable function $f(x_1, x_2, \ldots, x_n)$ has a saddle point at a critical point (x^*) if we can partition the vector x^* into two subvectors (x^{1*}, x^{2*}) where $x^{1*} \in X^1 \subseteq R^q$ and $x^{2*} \in X^2 \subseteq R^p$ (n = p + q) with the following property

$$f(x^1,\, x^{2*}) \;\leq\; f(x^{1*},\, x^{2*}) \;\leq\; f(x^{1*},\, x^2) \tag{8}$$

FIGURE 1. Local maximum of function $f(x_1, x_2) = e^{-(x_1^2+x_2^2)}$



for all $x^1 \in X^1$ and $x^2 \in X^2$.

The idea is that a saddle point attains a maximum in one direction and a minimum in the other direction. Figure 2 shows a function with a saddle point. The function reaches a maximum along the x_1 axis and a minimum along the x_2 axis. A contour plot for the function $f(x_1, x_2) = x_1^2 - x_2^2$ in contained in figure 3.

3.3.3. *Example problems.* Find all the critical points for each of the following functions.

1: $y = 16x_1 + 12x_2 - x_1^2 - x_2^2$
2: $y = 25\,x_1 + 18\,x_2 - x_1^2 - x_2^2$
3: $y = 100\,x_1 + 76\,x_2 - 4\,x_1^2 - 2\,x_2^2$
4: $y = 20\,x_1 + 40\,x_2 - x_1^2 + 2\,x_1x_2 - 2x_2^2$
5: $y = 30\,x_1 + 20\,x_2 - 2\,x_1^2 + 2\,x_1x_2 - x_2^2$
6: $y = 30\,x_1^{0.4}\,x_2^{0.2} - 6\,x_1 - 2\,x_2$
7: $y = 20\,x_1^{0.3}\,x_2^{0.5} - 5\,x_1 - 6\,x_2$

4. SECOND DERIVATIVE TEST FOR LOCAL EXTREME VALUES

4.1. A theorem on second derivatives of a function and local extreme values.

Theorem 3. *Suppose that f(x_1, x_2) and its first and second partial derivatives are continuous throughout a disk centered at (a, b) and that* $\frac{\partial f}{\partial x_1}(a, b) = \frac{\partial f}{\partial x_2}(a, b) = 0$. *Then*

a: *f has a* **local maximum** *at (a, b) if* $\frac{\partial^2 f}{\partial x_1^2}(a, b) < 0$ *and* $\frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2} - \left[\frac{\partial^2 f}{\partial x_1\,\partial x_2}\right]^2 > 0$ *at* (a, b).

FIGURE 2. Saddle point of the function $f(x_1,\ x_2) = x_1^2\ -\ x_2^2$

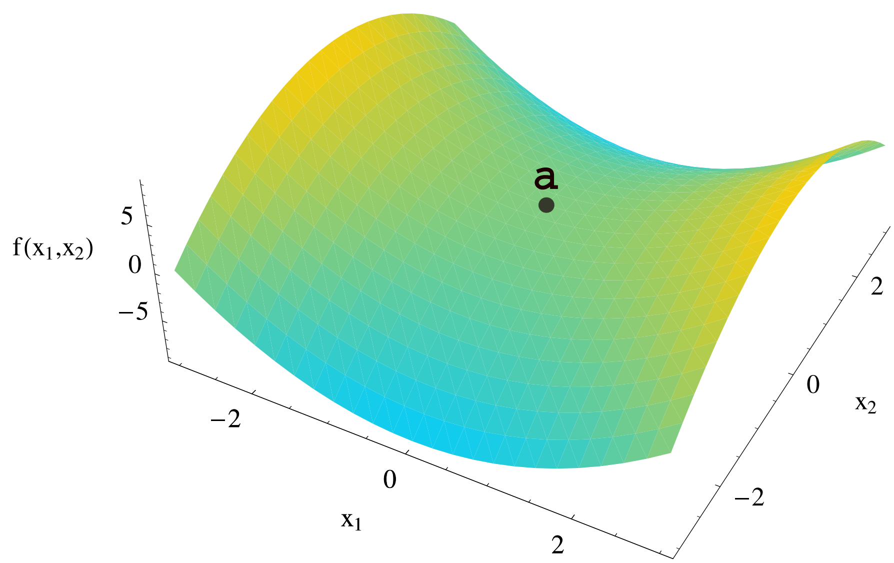
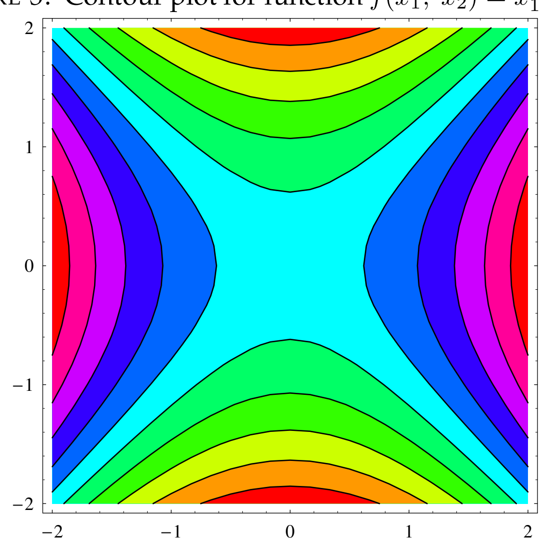


FIGURE 3. Contour plot for function $f(x_1,\ x_2) = x_1^2\ -\ x_2^2$



We can also write this as $f_{11}\ <\ 0$ *and* $f_{11}\, f_{22}\ -\ f_{12}^2\ >\ 0$ *at* $(a,\ b)$.

b: *f has a* **local minimum** *at (a, b) if*

$\frac{\partial^2 f}{\partial x_1^2}\,(a,\, b) \;>\; 0$ *and* $\frac{\partial^2 f}{\partial x_1^2}\,\frac{\partial^2 f}{\partial x_2^2} \;-\; \left[\frac{\partial^2 f}{\partial x_1\,\partial x_2}\right]^2 \;>\; 0$ *at* $(a,\; b)$.
We can also write this as $f_{11} \;>\; 0$ *and* $f_{11}\, f_{22} \;-\; f_{12}^2 \;>\; 0$ *at* $(a,\, b)$.

c: *f has a* **saddle point** *at (a, b) if* $\frac{\partial^2 f}{\partial x_1^2}\,\frac{\partial^2 f}{\partial x_2^2} \;-\; \left[\frac{\partial^2 f}{\partial x_1\,\partial x_2}\right]^2 \;<\; 0$ *at* $(a,\, b)$.
We can also write this as $f_{11}\, f_{22} \;-\; f_{12}^2 \;<\; 0$ *at* $(a,\, b)$.

d: *The* **test is inconclusive** *at (a, b) if* $\frac{\partial^2 f}{\partial x_1^2}\,\frac{\partial^2 f}{\partial x_2^2} \;-\; \left[\frac{\partial^2 f}{\partial x_1\,\partial x_2}\right]^2 \;=\; 0$ *at* $(a,\, b)$.
In this case we must find some other way to determine the behavior of f at (a, b).

The expression $\frac{\partial^2 f}{\partial x_1^2}\,\frac{\partial^2 f}{\partial x_2^2} \;-\; \left[\frac{\partial^2 f}{\partial x_1\,\partial x_2}\right]^2$ is called the **discriminant** of f. It is sometimes easier to remember it by writing it in determinant form

$$\frac{\partial^2 f}{\partial x_1^2}\,\frac{\partial^2 f}{\partial x_2^2} \;-\; \left[\frac{\partial^2 f}{\partial x_1\,\partial x_2}\right]^2 \;=\; \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} \tag{9}$$

The theorem says that if the discriminant is positive at the point (a, b), then the surface curves the same way in all directions; downwards if $f_{11} < 0$, giving rise to a local maximum, and upwards if $f_{11} > 0$, giving a local minimum. On the other hand, if the discriminant is negative at (a, b), then the surface curves up in some directions and down in others and we have a saddle point.

4.2. **Example problems.**

4.2.1. *Example 1.*

$$y \;=\; 2 \;-\; x_1^2 \;-\; x_2^2$$

Taking the first partial derivatives we obtain

$$\frac{\partial f}{\partial x_1} \;=\; -\,2\,x_1 \;= 0 \tag{10a}$$

$$\frac{\partial f}{\partial x_2} \;=\; -\,2\,x_2 \;= 0 \tag{10b}$$

$$\Rightarrow\; x_1 \;=\; 0, \qquad x_2 \;= 0 \tag{10c}$$

Thus we have a critical point at (0,0). Taking the second derivatives we obtain

$$\frac{\partial^2 f}{\partial x_1^2} \;=\; -\,2, \qquad \frac{\partial^2 f}{\partial x_1\,\partial x_2} \;=\; 0,$$

$$\frac{\partial^2 f}{\partial x_2\,\partial x_1} \;= 0, \qquad \frac{\partial^2 f}{\partial x_2^2} \;=\; -2$$

The **discriminant** is

$$\frac{\partial^2 f}{\partial x_1^2}\,\frac{\partial^2 f}{\partial x_2^2} \;-\; \left[\frac{\partial^2 f}{\partial x_1\,\partial x_2}\right]^2 \;=\; \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} \;=\; \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} \;=\; (-2)\,(-2) \;-\; (0)\,(0) \;=\; 4.$$

Thus this function has a maximum at the point (0, 0) because $\frac{\partial^2 f}{\partial x_1^2}$ and the discriminant is positive. The graph of the function is given in figure 4. The level curves are given in figure 5. The tangent plane is given in figure 6.

FIGURE 4. Graph of the function $2 - x_1^2 - x_2^2$

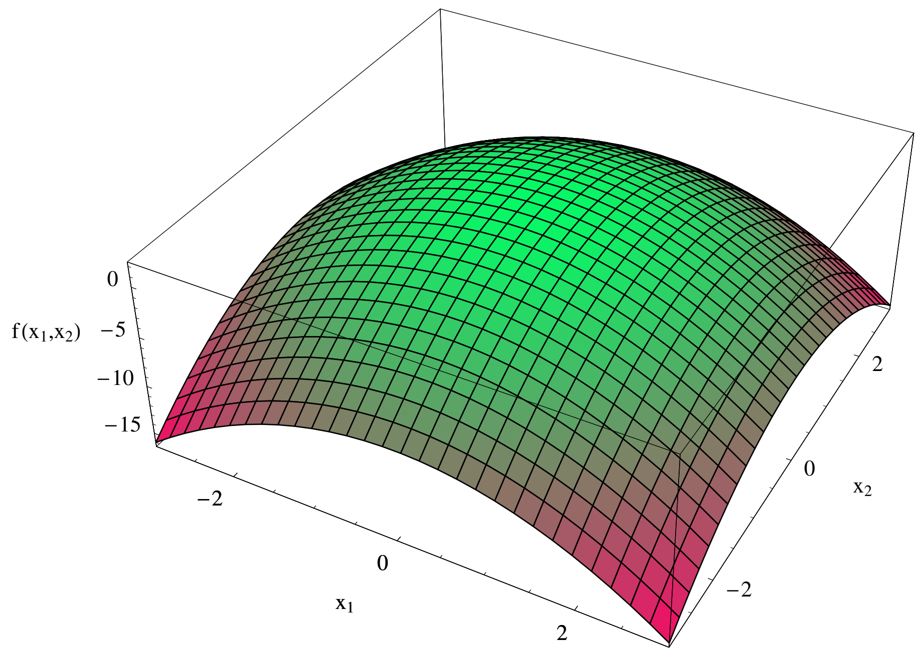
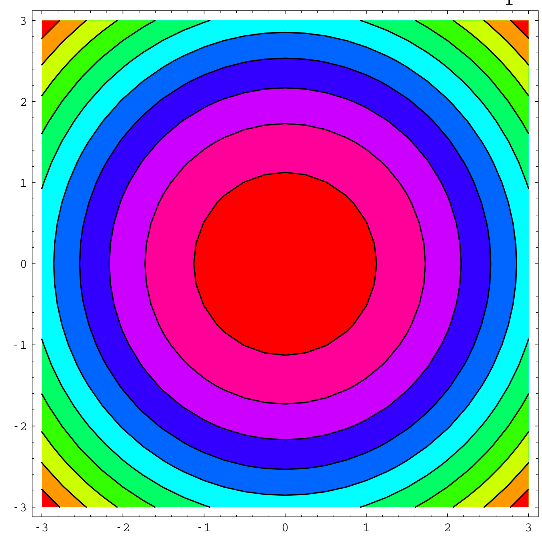


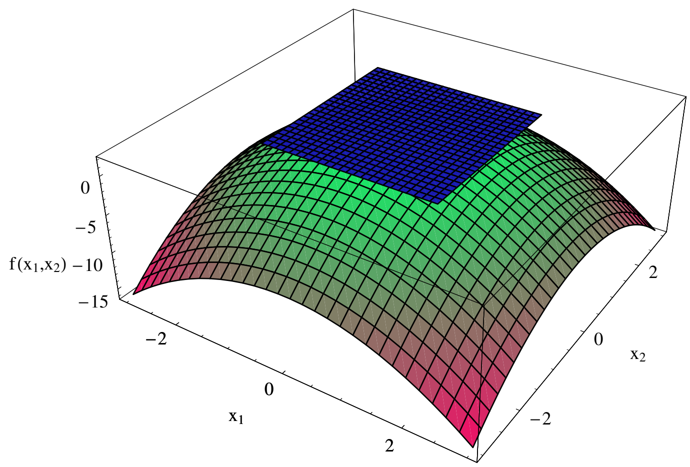
FIGURE 5. Level curves of the function $2 - x_1^2 - x_2^2$



4.2.2. *Example 2.*

$$y = 16x_1 + 12x_2 + x_1^2 + x_2^2$$

FIGURE 6. Tangent plane to the function $2 - x_1^2 - x_2^2$



Taking the first partial derivatives we obtain

$$\frac{\partial f}{\partial x_1} = 16 + 2x_1 = 0 \tag{11a}$$

$$\frac{\partial f}{\partial x_2} = 12 + 2x_2 = 0 \tag{11b}$$

These equations are easy to solve because the first equation only depends on x_1 and the second equation only depends on x_2. Solving equation 11a, we obtain

$$\begin{aligned} 16 + 2x_1 &= 0 \\ \Rightarrow 2x_1 &= -16 \\ \Rightarrow x_1 &= -8 \end{aligned}$$

Solving equation 11b, we obtain

$$\begin{aligned} 12 + 2x_2 &= 0 \\ \Rightarrow 2x_2 &= -12 \\ \Rightarrow x_1 &= -6 \end{aligned}$$

Thus we have a critical point at (-8,-6). Taking the second derivatives we obtain

$$\begin{aligned} \frac{\partial^2 f}{\partial x_1^2} &= 2, & \frac{\partial^2 f}{\partial x_1 \, \partial x_2} &= 0 \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} &= 0, & \frac{\partial^2 f}{\partial x_2^2} &= 2 \end{aligned}$$

The **discriminant** is

$$\frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2} - \left[\frac{\partial^2 f}{\partial x_1 \partial x_2}\right]^2 = \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = (2)(2) - (0)(0) = 4.$$

Thus this function has a minimum at the point (-8, -6) because $\frac{\partial^2 f}{\partial x_1^2} > 0$ and the discriminant is positive. The graph of the function is given in figure 7. The level curves are given in figure 8.

FIGURE 7. Graph of the function $f(x_1, x_2) = 16x_1 + 12x_2 + x_1^2 + x_2^2$

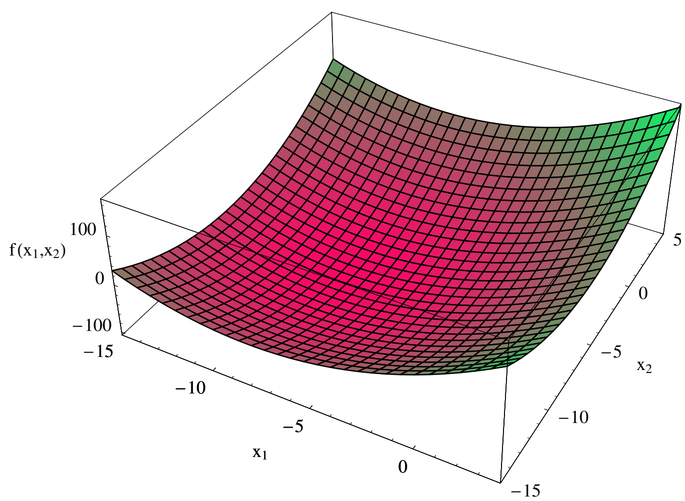
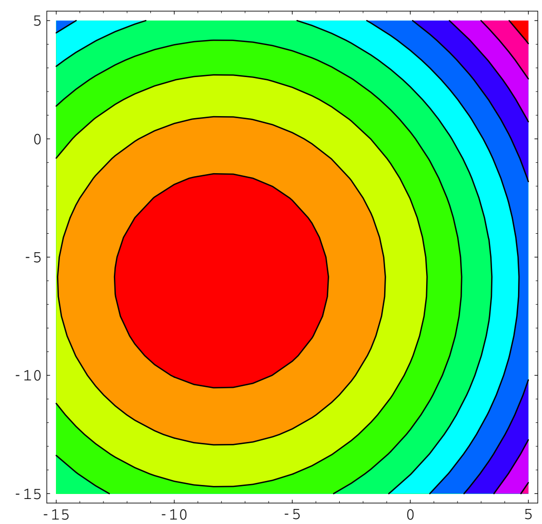
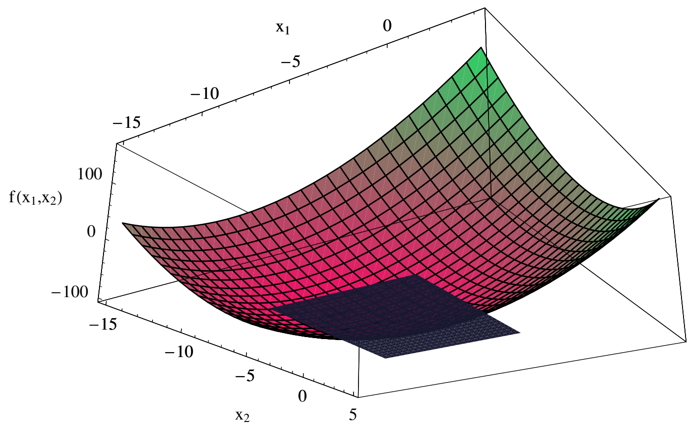


FIGURE 8. Level curves of the function $16x_1 + 12x_2 + x_1^2 + x_2^2$



The tangent plane is given in figure 9.

FIGURE 9. Plane tangent to the function $16x_1 + 12x_2 + x_1^2 + x_2^2$



4.2.3. *Example 3.*

$$y = x_1 x_2$$

Taking the first partial derivatives we obtain

$$\frac{\partial f}{\partial x_1} = x_2 = 0 \tag{12a}$$

$$\frac{\partial f}{\partial x_2} = x_1 = 0 \tag{12b}$$

Thus this function has a critical value at the point at (0,0). Taking the second derivatives we obtain

$$\frac{\partial^2 f}{\partial x_1^2} = 0, \qquad \frac{\partial^2 f}{\partial x_1\, \partial x_2} = 1$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = 1, \qquad \frac{\partial^2 f}{\partial x_2^2} = 0$$

The **discriminant** is

$$\frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2} - \left[\frac{\partial^2 f}{\partial x_1\, \partial x_2}\right]^2 = \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = (0)\,(0) - (1)\,(1) = -1.$$

The function has a saddle point at (0,0) because the discriminant is negative at this critical value. The graph of the function is given in figure 10. The level curves are given in figure 11. The tangent plane is shown from two different angles in figures 12 and 13.

FIGURE 10. Graph of the function $f(x_1, x_2) = x_1\, x_2$

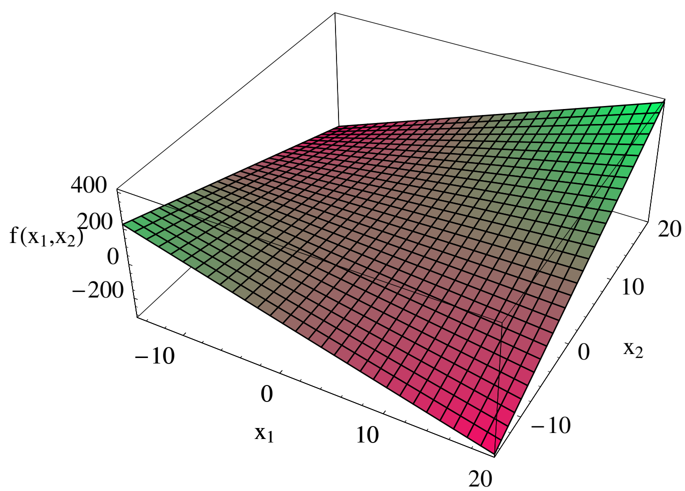
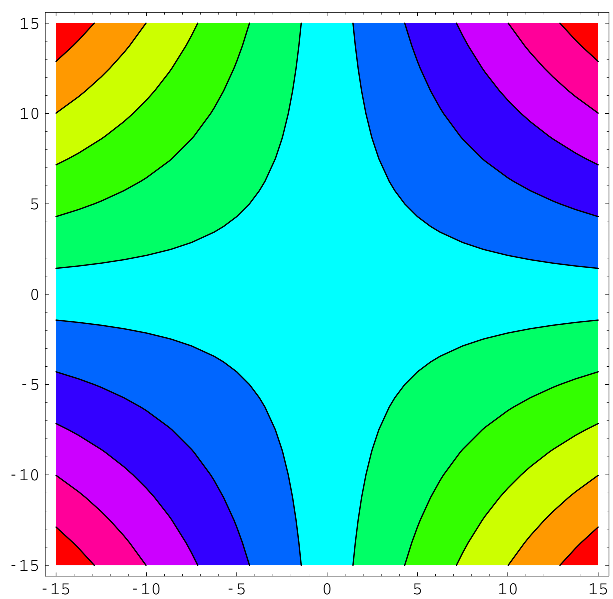


FIGURE 11. Let sets of the function $f(x_1, x_2) = x_1\, x_2$



4.2.4. *Example 4.*

$$y = 20x_1 + 40x_2 - x_1^2 + 4x_1x_2 - 2\,x_2^2$$

Taking the first partial derivatives we obtain

FIGURE 12. Saddle point of the function $f(x_1, x_2) = x_1 \, x_2$

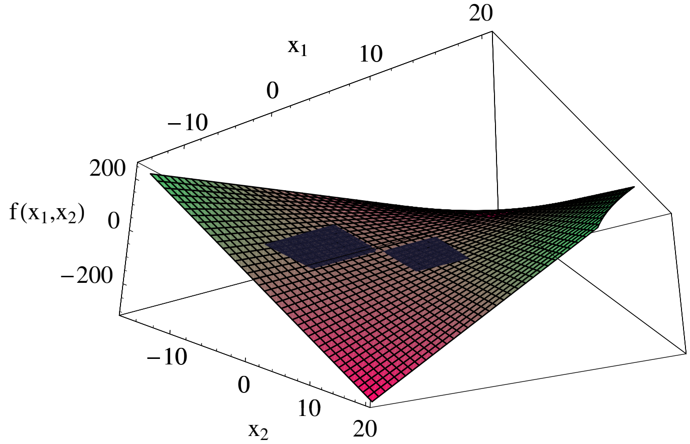
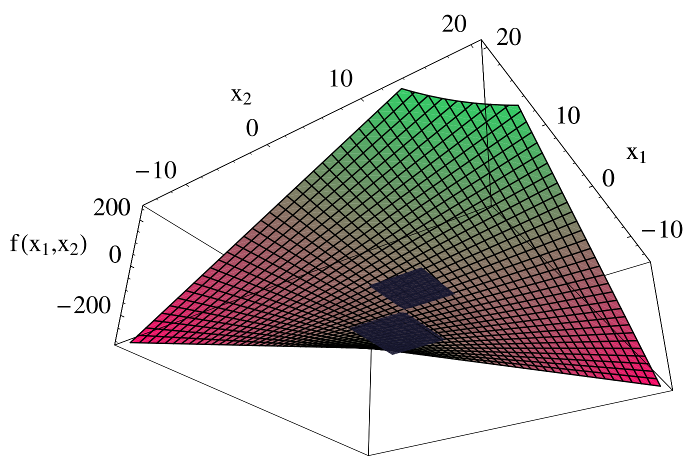


FIGURE 13. Saddle point of the function $f(x_1, x_2) = x_1 \, x_2$



$$\frac{\partial f}{\partial x_1} = 20 - 2x_1 + 4x_2 = 0 \tag{13a}$$

$$\frac{\partial f}{\partial x_2} = 40 + 4x_1 - 4x_2 = 0 \tag{13b}$$

Write this system of equations as

$$20 - 2x_1 + 4x_2 = 0 \tag{14a}$$

$$40 + 4x_1 - 4x_2 = 0 \tag{14b}$$

$$\Rightarrow 2x_1 - 4x_2 = 20 \tag{14c}$$

$$4x_1 - 4x_2 = -40 \tag{14d}$$

Multiply equation 14c by -2 and add it to equation 14d. First the multiplication

$$-2(2x_1 - 4x_2) = -2(20)$$
$$\Rightarrow -4x_1 + 8x_2 = -40$$

Then the addition

$$\begin{array}{cccc} -4x_1 & +8x_2 & = & -40 \\ 4x_1 & -4x_2 & = & -40 \\ -- & -- & -- & -- \\ & 4x_2 & = & -80 \end{array} \tag{15}$$

Now multiply the result in equation 15 by $\frac{1}{4}$ to obtain the system

$$\begin{array}{cccc} 2x_1 & -4x_2 & = & 20 \\ & x_2 & = & -20 \end{array} \tag{16}$$

Now multiply the second equation in 16 by 4 and add to the first equation as follows

$$\begin{array}{cccc} 2x_1 & -4x_2 & = & 20 \\ & 4x_2 & = & -80 \\ -- & -- & -- & -- \\ 2x_1 & & = & -60 \end{array} \tag{17}$$

Multiplying the result in equation 17 by $\frac{1}{2}$ we obtain the system

$$\begin{array}{cccc} x_1 & & = & -30 \\ & x_2 & = & -20 \end{array} \tag{18}$$

Thus we have a critical value at (-30, -20). We obtain the second derivatives by differentiating the system in equation 13

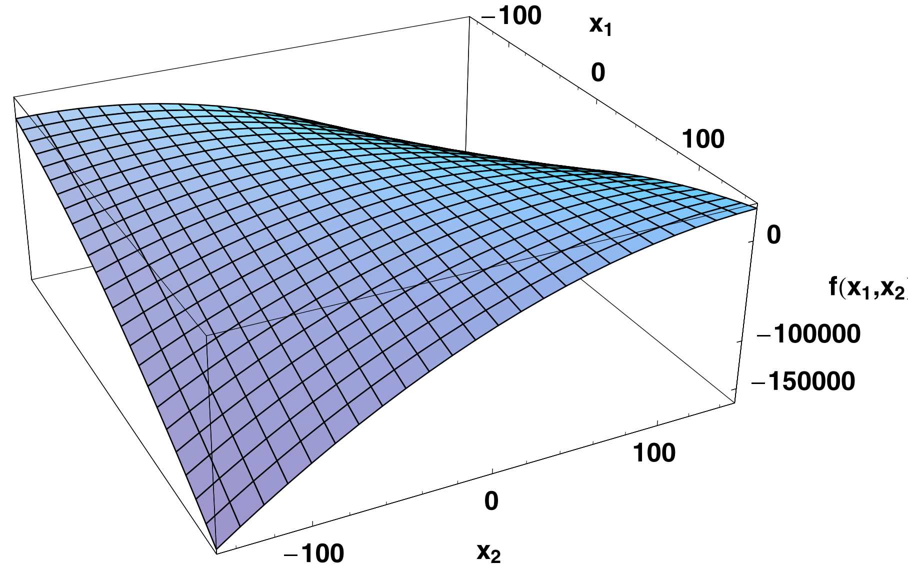
$$\frac{\partial^2 f}{\partial x_1^2} = -2, \qquad \frac{\partial^2 f}{\partial x_1 \, \partial x_2} = 4$$
$$\frac{\partial^2 f}{\partial x_2 \, \partial x_1} = 4, \qquad \frac{\partial^2 f}{\partial x_1^2} = -4$$

The **discriminant** is

$$\frac{\partial^2 f}{\partial x_1^2} \frac{\partial^2 f}{\partial x_2^2} - \left[\frac{\partial^2 f}{\partial x_1 \, \partial x_2} \right]^2 = \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} = \begin{vmatrix} -2 & 4 \\ 4 & -4 \end{vmatrix} = (-2)(-4) - (4)(4) = -8.$$

The discriminant is negative and so the function has a saddle point. The graph of the function is given in figure 14. The level curves are given in figure 15. The tangent plane is shown from two different angles in figures 16 and 17.

FIGURE 14. Graph of the function $f(x_1, x_2) = 20x_1 + 40x_2 - x_1^2 + 4x_1x_2 - 2x_2^2$



4.2.5. *Example 5.*

$$y = -x_1 x_2\, e^{\frac{-(x_1^2 + x_2^2)}{2}}$$

Taking the first partial derivatives we obtain

$$\begin{aligned}\frac{\partial f}{\partial x_1} &= -\,x_2\, e^{\frac{-(x_1^2 + x_2^2)}{2}} + x_1^2 x_2 e^{\frac{-(x_1^2 + x_2^2)}{2}} \\ &= x_2(x_1^2 - 1)e^{\frac{-(x_1^2 + x_2^2)}{2}} \\ \frac{\partial f}{\partial x_2} &= -\,x_1 e^{\frac{-(x_1^2 + x_2^2)}{2}} + x_1 x_2^2\, e^{\frac{-(x_1^2 + x_2^2)}{2}} \\ &= x_1(x_2^2 - 1)e^{\frac{-(x_1^2 + x_2^2)}{2}}\end{aligned}$$

We set both of the equations equal to zero and solve for critical values of x_1 and x_2. Because

FIGURE 15. Level sets of the function $f(x_1, x_2) = 20x_1 + 40x_2 - x_1^2 + 4x_1x_2 - 2x_2^2$

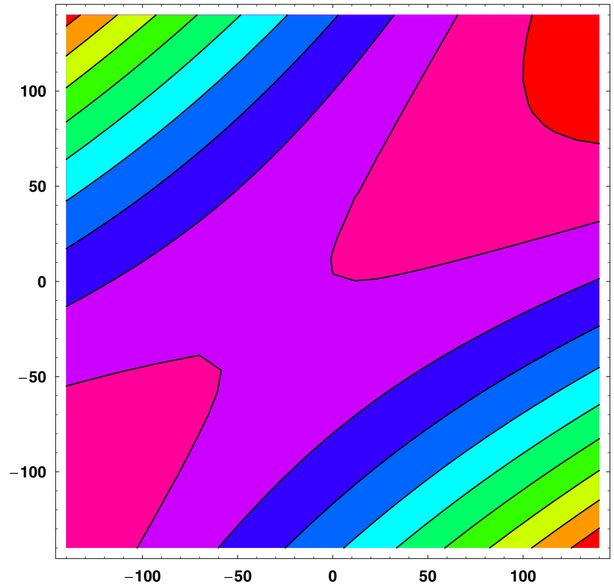
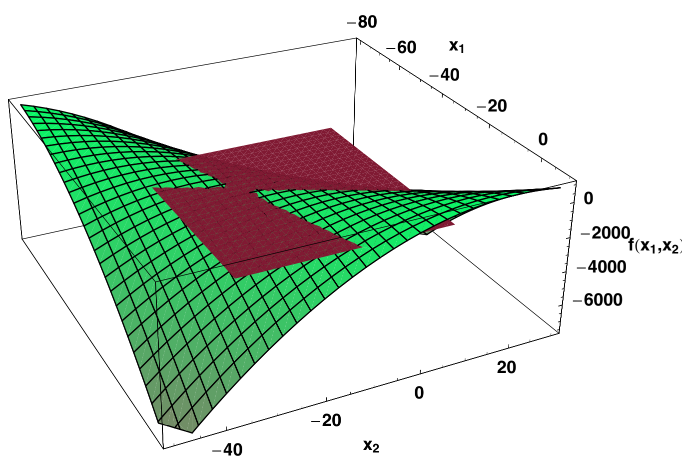


FIGURE 16. Plane through saddle point of the function $f(x_1, x_2) = 20x_1 + 40x_2 - x_1^2 + 4x_1x_2 - 2x_2^2$

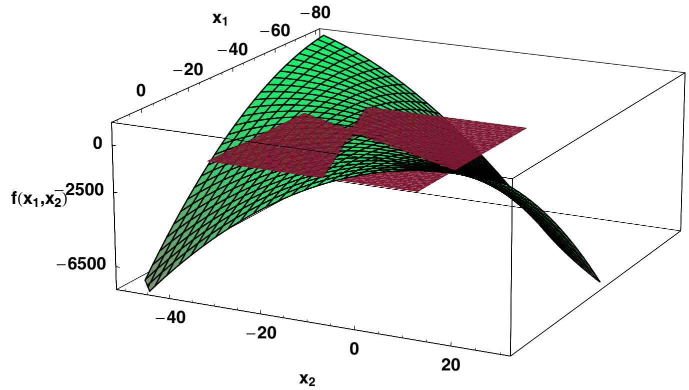


$$e^{\frac{-(x_1^2 \,+\, x_2^2)}{2}} \neq 0 \; for \; all \; (x_1, x_2)$$

the equations can be zero if and only if

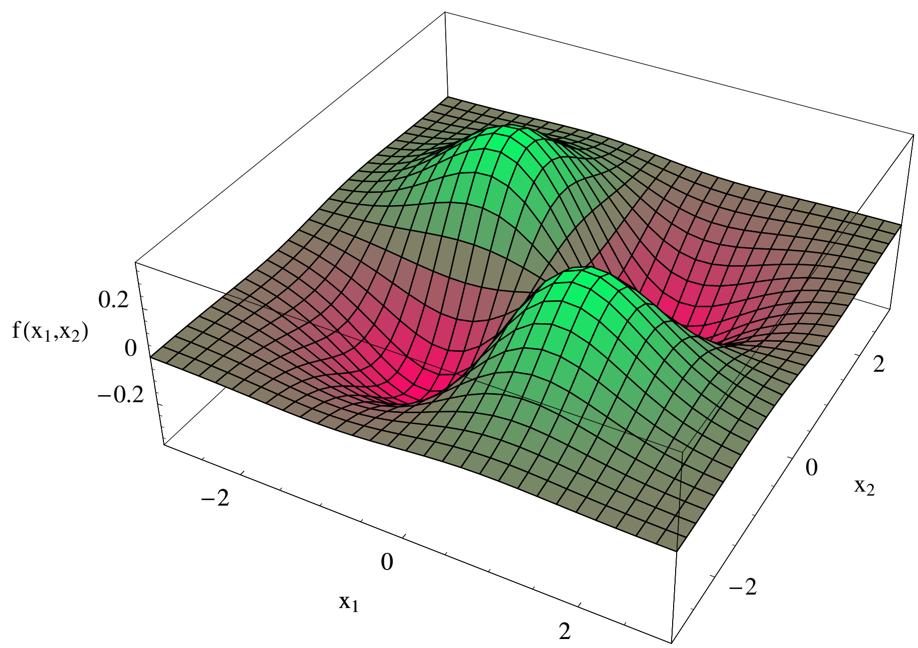
$$x_2\,(x_1^2 \,-\, 1) \,= 0$$
$$x_1\,(x_2^2 \,-\, 1) \,= 0$$

FIGURE 17. Plane through saddle point of the function $f(x_1, x_2) = 20x_1 + 40x_2 - x_1^2 + 4x_1x_2 - 2x_2^2$



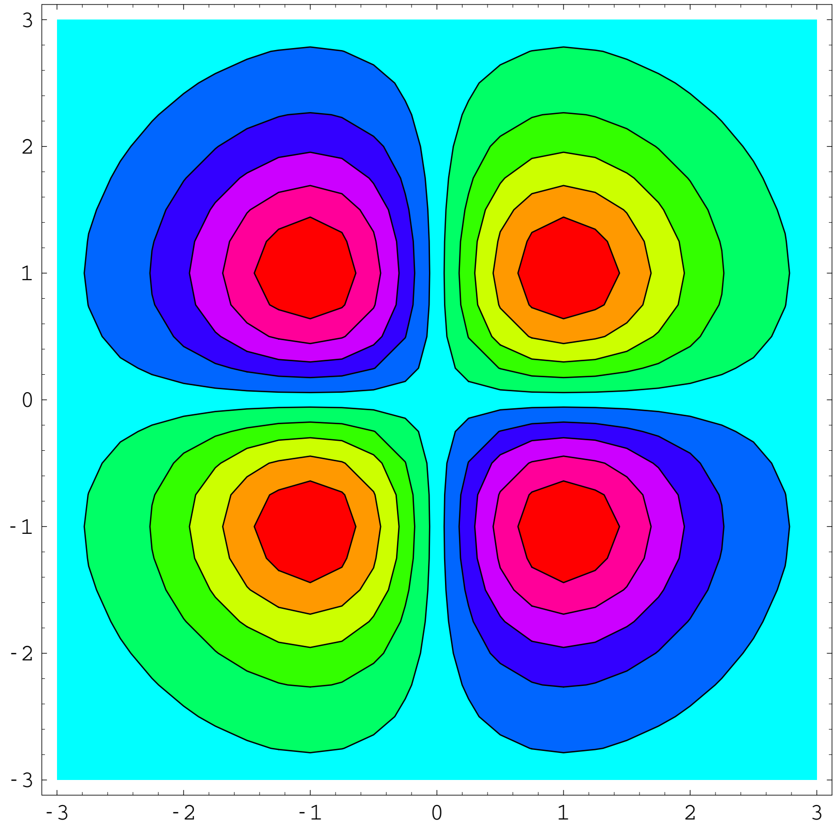
Thus we have critical values at (0, 0), (1, 1), (1, -1), (-1, 1) and (-1, -1).
A graph of the function is contained in figure 18.

FIGURE 18. Graph of the function $f(x_1,\, x_2) \;=\; -x_1\, x_2\; e^{\frac{-(x_1^2 \,+\, x_2^2)}{2}}$



A set of level curves is contained in figure 19.

FIGURE 19. Level sets of the function $f(x_1, x_2) = -x_1\, x_2\, e^{\frac{-(x_1^2 + x_2^2)}{2}}$



We compute the second partial derivatives. The second partial derivative of f with respect to x_1 is

$$f_{x_1 x_1} = 3e^{\frac{-x_1^2-x_2^2}{2}}\, x_1 x_2 - e^{\frac{-x_1^2-x_2^2}{2}}\, x_1^3\, x_2$$

The second partial derivative of f with respect to x_2 is given by

$$f_{x_2 x_2} = 3e^{\frac{-x_1^2-x_2^2}{2}}\, x_1 x_2 - e^{\frac{-x_1^2-x_2^2}{2}}\, x_1 x_2^3$$

The cross partial derivative of f with respect to x_1 and x_2 is

$$f_{x_2 x_1} = -e^{\frac{-x_1^2 - x_2^2}{2}} + e^{\frac{-x_1^2 - x_2^2}{2}}\, x_1^2 + e^{\frac{-x_1^2 - x_2^2}{2}}\, x_2^2 - e^{\frac{-x_1^2 - x_2^2}{2}}\, x_1^2 x_2^2$$

The **discriminant** $\frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2} - \left[\frac{\partial^2 f}{\partial x_1\, \partial x_2}\right]^2$ is given by $\begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix}$.
We can write this as follows by making the substitutions.

$$\begin{vmatrix} 3e^{\frac{-x_1^2-x_2^2}{2}}\, x_1 x_2 - e^{\frac{-x_1^2-x_2^2}{2}}\, x_1^3\, x_2 & -e^{\frac{-x_1^2-x_2^2}{2}} + e^{\frac{-x_1^2-x_2^2}{2}}\, x_1^2 + e^{\frac{-x_1^2-x_2^2}{2}}\, x_2^2 - e^{\frac{-x_1^2-x_2^2}{2}}\, x_1^2\, x_2^2 \\ -e^{\frac{-x_1^2-x_2^2}{2}} + e^{\frac{-x_1^2-x_2^2}{2}}\, x_1^2 + e^{\frac{-x_1^2-x_2^2}{2}}\, x_2^2 - e^{\frac{-x_1^2-x_2^2}{2}}\, x_1^2\, x_2^2 & 3e^{\frac{-x_1^2-x_2^2}{2}}\, x_1\, x_2 - e^{\frac{-x_1^2-x_2^2}{2}}\, x_1\, x_2^3 \end{vmatrix}$$

Computing the discriminant we obtain

$$Discriminant = \left(3e^{\frac{-x_1^2-x_2^2}{2}}\, x_1 x_2 - e^{\frac{-x_1^2-x_2^2}{2}}\, x_1^3\, x_2\right)\left(3e^{\frac{-x_1^2-x_2^2}{2}}\, x_1 x_2 - e^{\frac{-x_1^2-x_2^2}{2}}\, x_1\, x_2^3\right)$$
$$- \left[-e^{\frac{-x_1^2-x_2^2}{2}} + e^{\frac{-x_1^2-x_2^2}{2}}\, x_1^2 + e^{\frac{-x_1^2-x_2^2}{2}}\, x_2^2 - e^{\frac{-x_1^2-x_2^2}{2}}\, x_1^2\, x_2^2\right]\left[-e^{\frac{-x_1^2-x_2^2}{2}} + e^{\frac{-x_1^2-x_2^2}{2}}\, x_1^2 + e^{\frac{-x_1^2-x_2^2}{2}}\, x_2^2 - e^{\frac{-x_1^2-x_2^2}{2}}\, x_1^2 x_2^2\right]$$

Multiplying out the first term we obtain

$$\begin{aligned} term\ 1 &= e^{-x_1^2-x_2^2}\, x_1^2\,(-3+x_1^2)\, x_2^2\,(-3+x_2^2) \\ &= e^{-x_1^2-x_2^2}\,\left(9\, x_1^2\, x_2^2 - 3\, x_1^4\, x_2^2 - 3\, x_1^2\, x_2^4 + x_1^4\, x_2^4\right) \end{aligned}$$

Multiplying out the second term we obtain

$$\begin{aligned} term\ 2 &= e^{-x_1^2-x_2^2}\,(-1+x_1^2)^2\,(-1+x_2^2)^2 \\ &= e^{-x_1^2-x_2^2}\,\left(1 - 2x_1^2 + x_1^4 - 2x_2^2 + 4x_1^2\, x_2^2 - 2x_1^4\, x_2^2 + x_2^4 - 2x_1^2\, x_2^4 + x_1^4\, x_2^4\right) \end{aligned}$$

Subtracting term 2 from term 1 we obtain

$$\begin{aligned} Discriminant = &\ e^{-x_1^2-x_2^2}\,\left(9\, x_1^2\, x_2^2 - 3\, x_1^4\, x_2^2 - 3\, x_1^2\, x_2^4 + x_1^4\, x_2^4\right) \\ &+ \left(-1 + 2x_1^2 - x_1^4 + 2x_2^2 - 4x_1^2\, x_2^2 + 2x_1^4\, x_2^2 - x_2^4 + 2x_1^2\, x_2^4 - x_1^4\, x_2^4\right) \\ &= e^{-x_1^2-x_2^2}\,\left(-1 + 5x_1^2\, x_2^2 - x_1^4\, x_2^2 - x_1^2\, x_2^4 + 2x_1^2 + 2x_2^2 - x_1^4 - x_2^4\right) \end{aligned}$$

Because $e^{-x_1^2\, -\, x_2^2} > 0$ for all $(x_1,\, x_2)$, we need only consider the term in parentheses to check the sign of the discriminant. Consider each of the cases in turn.

a: Consider the critical point (0, 0). In the case the relevant term is given by

$$\left(-1 + 5x_1^2\, x_2^2 - x_1^4\, x_2^2 - x_1^2\, x_2^4 + 2x_1^2 + 2x_2^2 - x_1^4 - x_2^4\right) = -1$$

This is negative so this is a saddle point. We can see this in figure 20.

b: Consider the critical point (1, -1). In the case the relevant term is given by

$$\left(-1 + 5x_1^2\, x_2^2 - x_1^4\, x_2^2 - x_1^2\, x_2^4 + 2x_1^2 + 2x_2^2 - x_1^4 - x_2^4\right) = 4$$

This is positive. The second derivative of f with respect to x_1 evaluated at (1, -1) is given by

$$\begin{aligned} f_{x_1 x_1} &= 3e^{\frac{-x_1^2-x_2^2}{2}}\, x_1 x_2 - e^{\frac{-x_1^2-x_2^2}{2}}\, x_1^3 x_2 \\ &= e^{\frac{-x_1^2-x_2^2}{2}}\,\left(3x_1 x_2 - x_1^3 x_2\right) \\ &= e^{-1}\,(-3+1) = \frac{-2}{e} < 0. \end{aligned}$$

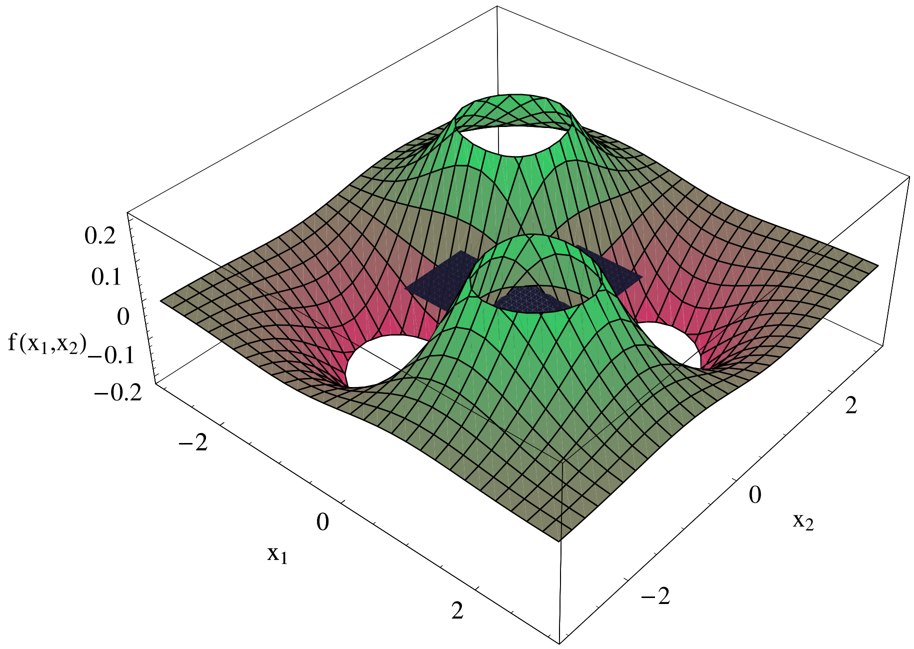
Given that the f_{11} is negative, this point is a local maximum. We can see this in figure 21.

c: Consider the critical point (-1, 1). In the case the relevant term is given by

$$\left(-1 + 5x_1^2\, x_2^2 - x_1^4\, x_2^2 - x_1^2\, x_2^4 + 2x_1^2 + 2x_2^2 - x_1^4 - x_2^4\right) = 4$$

This is positive. The second derivative of x with respect to x_1 evaluated at (1, -1) is given by

FIGURE 20. Saddle point of the function $f(x_1, x_2) = -x_1 x_2 \, e^{\frac{-(x_1^2 + x_2^2)}{2}}$



$$
\begin{aligned}
f_{x_1 x_1} =& 3e^{\frac{-x_1^2 - x_2^2}{2}} \, x_1 x_2 - e^{\frac{-x_1^2 - x_2^2}{2}} \, x_1^3 x_2 \\
=& e^{\frac{-x_1^2 - x_2^2}{2}} \left(3x_1 x_2 - x_1^3 x_2\right) \\
=& e^{-1} \left(-3 + 1\right) = \frac{-2}{e} < 0.
\end{aligned}
$$

Given that f_{11} is negative, this point is a local maximum. We can see this in figure 22.

d: Consider the critical point (1, 1). In the case the relevant term is given by

$$\left(-1 + 5x_1^2 \, x_2^2 - x_1^4 \, x_2^2 - x_1^2 \, x_2^4 + 2x_1^2 + 2x_2^2 - x_1^4 - x_2^4\right) = 4$$

This is positive. The second derivative of f with respect to x_1 evaluated at (1, 1) is given by

$$
\begin{aligned}
f_{x_1 x_1} =& 3e^{\frac{-x_1^2 - x_2^2}{2}} \, x_1 x_2 - e^{\frac{-x_1^2 - x_2^2}{2}} \, x_1^3 x_2 \\
=& e^{\frac{-x_1^2 - x_2^2}{2}} \left(3x_1 x_2 - x_1^3 x_2\right) \\
=& e^{-1} \left(3 - 1\right) = \frac{2}{e} > 0.
\end{aligned}
$$

Given that the f_{11} derivative is positive, this point is a local minimum. We can see this in figure 23.

FIGURE 21. Local maximum of the function $f(x_1, x_2) = -x_1 \, x_2 \, e^{\frac{-(x_1^2 + x_2^2)}{2}}$

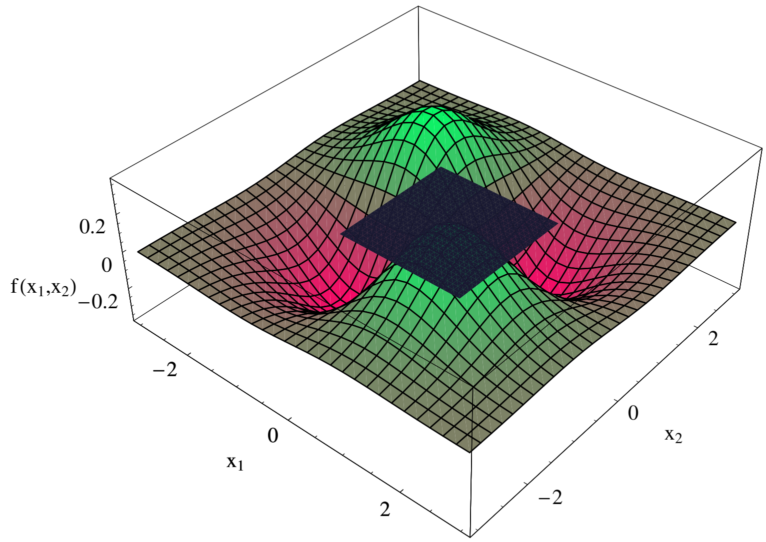
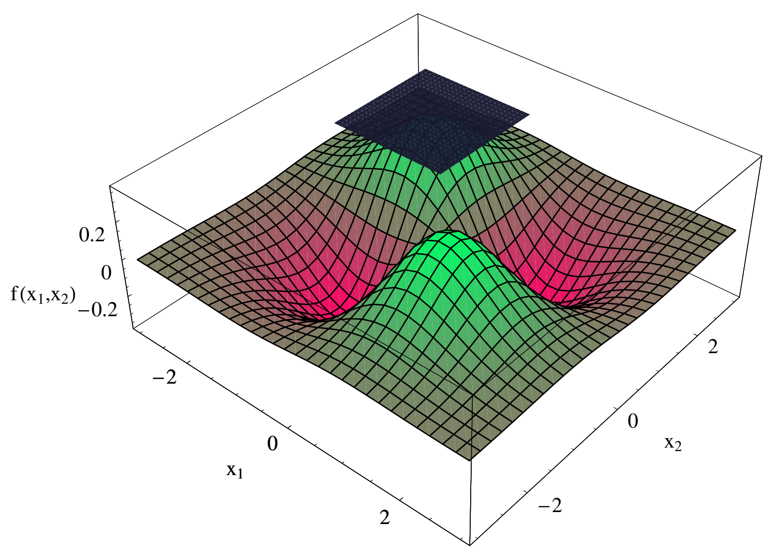


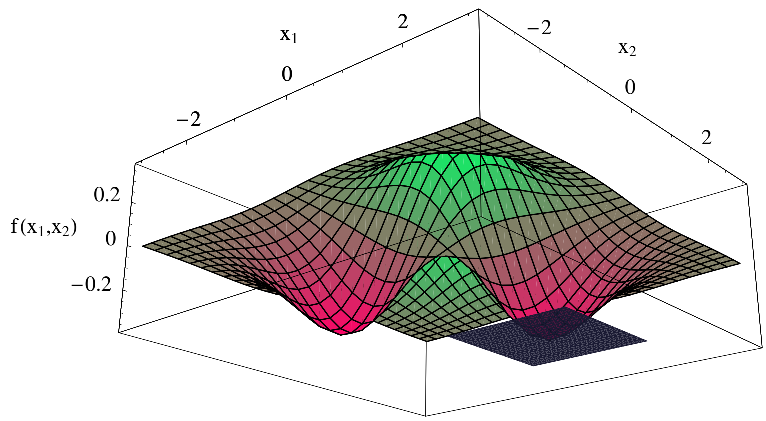
FIGURE 22. Local maximum of the function $f(x_1, x_2) = -x_1 x_2 \, e^{\frac{-(x_1^2 + x_2^2)}{2}}$



e: Consider the critical point (-1, -1). In the case the relevant term is given by

$$\left(-1 + 5x_1^2 \, x_2^2 - x_1^4 \, x_2^2 - x_1^2 \, x_2^4 + 2x_1^2 + 2x_2^2 - x_1^4 - x_2^4\right) = 4$$

FIGURE 23. Local minimum of the function $f(x_1,\, x_2) \;=\; -x_1\, x_2\; e^{\frac{-(x_1^2\;+\;x_2^2)}{2}}$

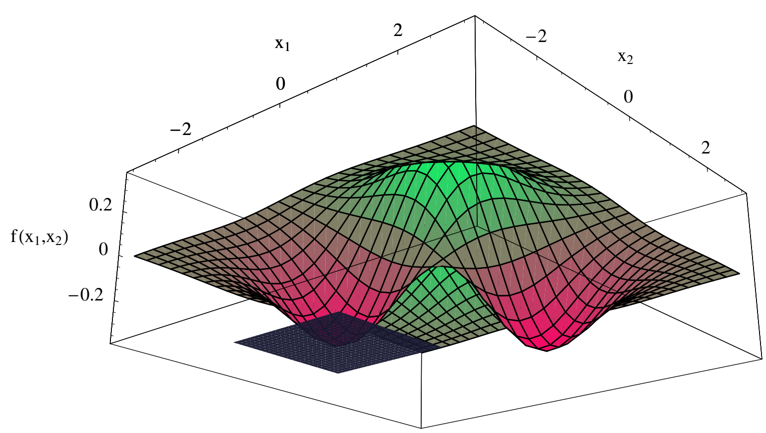


This is positive. The second derivative of f with respect to x_1 evaluated at (-1, -1) is given by

$$\begin{aligned} f_{x_1 x_1} &= 3e^{\frac{-x_1^2 - x_2^2}{2}}\, x_1\, x_2 - e^{\frac{-x_1^2 - x_2^2}{2}}\, x_1^3\, x_2 \\ &= e^{\frac{-x_1^2 - x_2^2}{2}} \left(3x_1\, x_2 - x_1^3\, x_2\right) \\ &= e^{-1}\,(3-1) = \frac{2}{e} > 0. \end{aligned}$$

Given that the f_{11} derivative is positive, this point is a local minimum. We can see this in figure 24.

FIGURE 24. Local minimum of the function $f(x_1,\, x_2) \;=\; -x_1 x_2\; e^{\frac{-(x_1^2\;+\;x_2^2)}{2}}$



4.3. **In-class problems.**

a: $y = 2x_1^2 + x_2^2 - x_1 x_2 - 7\,x_2$
b: $y = x_2^2 - x_1 x_2 + 2x_1 + x_2 + 1$
c: $y = 16\,x_1 + 10\,x_2 - 2\,x_1^2 + x_1 x_2 - x_2^2$
d: $y = 4\,x_1 x_2 - 2\,x_1^2 - x_2^4$
e: $y = 10\,x_1^{0.4}\,x_2^{0.2} - x_1 - 2\,x_2$

5. Taylor series

5.1. **Definition of a Taylor series.** Let f be a function with derivatives of all orders throughout some interval containing *a* as an interior point. The **Taylor series generated by f at x = a** is

$$\sum_{k=0}^{\infty} \frac{f^k(a)}{k!}(x - a)^k = f(a) + f'a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \frac{f'''(a)}{3!}(x - a)^3 + \cdots + \frac{f^n(a)}{n!}(x - a)^n + \cdots \quad (19)$$

5.2. **Definition of a Taylor polynomial.** The linearization of a differentiable function f at a point *a* is the polynomial

$$P_1(x) = f(a) + f'(a)(x - a) \quad (20)$$

For a given a, f(a) and $f'(a)$ will be constants and we get a linear equation in x. This can be extended to higher order terms as follows.

Let f be a function with derivatives of order k for k = 1, 2, ... , N in some interval containing *a* as an interior point. Then for any integer n from 0 through N, the **Taylor polynomial of order n** generated by f at x = a is the polynomial

$$P_n(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \cdots + \frac{f^k(a)}{k!}(x - a)^k + \cdots + \frac{f^n(a)}{n!}(x - a)^n \quad (21)$$

That is, the Taylor polynomial of order n is the Taylor series truncated with the term containing the nth derivative of f. This allows us to approximate a function with polynomials of higher orders.

5.3. **Taylor's theorem.**

Theorem 4. *If a function f and its first n derivatives $f', f'', \ldots, f^{(n)}$ are continuous on [a, b] or on [b, a], and $f^{(n)}$ is differentiable on (a, b) or (b, a), then there exists a number c between a and b such that*

$$f(b) = f(a) + f'(a)(b - a) + \frac{f''(a)}{2!}(b - a)^2 + \ldots + \frac{f^n(a)}{n!}(b - a)^n + \frac{f^{n+1}(c)}{(n+1)!}(b - a)^{n+1} \quad (22)$$

What this says is that we can approximate the function f at the point b by an nth order polynomial defined at the point *a* with an error term defined by $\frac{f^{n+1}(c)}{(n+1)!}(b - a)^{n+1}$. In using Taylor's theorem we usually think of *a* as being fixed and treat b as an independent variable. If we replace b with x we obtain the following corollary.

Corollary 1. If a function f has derivatives of all orders in an open interval I containing *a*, then for each positive integer *n* and for each *x* in I,

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \cdots + \frac{f^n(a)}{n!}(x - a)^n + R_n(x),$$

where (23)

$$R_n(x) = \frac{f^{n+1}(c)}{(n+1)!}(x - a)^{n+1} \text{ for some c between a and x.}$$

If $R_n(x) \to 0$ as $n \to \infty$ for all x in I, then we say that the Taylor series generated by f at x = a converges to f on I.

6. Proof of the second derivative test for local extreme values

Let $f(x_1, x_2)$ have continuous partial derivatives in the open region R containing a point P(a, b) where $\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = 0$. Let h and k be increments small enough to put the point S(a+h, b+k) and the line segment joining it to P inside R. We parameterize the line segment PS as

$$x_1 = a + th, \; x_2 = b + tk, \; 0 \le t \le 1. \tag{24}$$

Define F(t) as f(a+th, b+tk). In this way F tracks the movement of the function f along the line segment from P to S. Now differentiate F with respect to t using the chain rule to obtain

$$F'(t) = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dt} = f_1 h + f_2 k. \tag{25}$$

Because $\frac{\partial f}{\partial x_1}$ and $\frac{\partial f}{\partial x_2}$ are differentiable, F′ is a differentiable function of t and we have

$$\begin{aligned} F''(t) &= \tfrac{\partial F'}{\partial x_1}\tfrac{dx_1}{dt} + \tfrac{\partial F'}{\partial x_2}\tfrac{dx_2}{dt} \\ &= \tfrac{\partial}{\partial x_1}(f_1 h + f_2 k)\, h + \tfrac{\partial}{\partial x_2}(f_1 h + f_2 k)\, k \\ &= h^2 f_{11} + 2hk\, f_{12} + k^2 f_{22}, \end{aligned} \tag{26}$$

where we use the fact that $f_{12} = f_{21}$. By the definition of F, which is defined over the interval $0 \le t \le 1$, and the fact that f is continuous over the open region R, it is clear that F and F' are continuous on [0, 1] and that F′ is differentiable on (0, 1). We can then apply Taylor's formula with $n = 2$ and $a = 0$. This will give

$$F(1) = F(0) + F'(0)(1 - 0) + F''(c)\frac{(1 - 0)^2}{2} \tag{27}$$

Simplifying we obtain

$$\begin{aligned} F(1) &= F(0) + F'(0)(1 - 0) + F''(c)\tfrac{(1 - 0)^2}{2} \\ &= F(0) + F'(0) + \tfrac{F''(c)}{2} \end{aligned} \tag{28}$$

for some c between 0 and 1. Now write equation 28 in terms of f. Since t = 1 in equation 28, we obtain f at the point (a+h, b+k) or

$$f(a+h, b+k) = f(a, b) + hf_1(a, b) + kf_2(a, b) + \tfrac{1}{2}(h^2 f_{11} + 2hk\, f_{12} + k^2 f_{22})|_{(a+ch,\, b+ck)} \tag{29}$$

Because $\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = 0$ at (a, b), equation 29 reduces to

$$f(a+h, b+k) - f(a, b) = \tfrac{1}{2}(h^2 f_{11} + 2hk\, f_{12} + k^2 f_{22})|_{(a+ch,\, b+ck)} \tag{30}$$

The presence of an extremum of f at the point (a, b) is determined by the sign of the left hand side of equation 30, or f(a+b, b+k) - f(a, b). If $[f(a+h, b+k) - f(a, b)] < 0$ then the point is a maximum. If

$[f(a+h,\, b+k) \; - \; f(a,\, b)] > 0$, the point is a minimum. So if the left hand side (lhs) of equation 20 < 0, we have a maximum, and if the left hand side of equation 20 > 0, we have a minimum. Specifically,

$$\begin{aligned} lhs\; &< 0, \quad f \; is \; at \; a \; maximum \\ lhs\; &> 0, \quad f \; is \; at \; a \; minimum \end{aligned}$$

By equation 30 this is the same as the sign of

$$Q(c) \; = (h^2 f_{11} \; + \; 2\, h\, k\, f_{12} \; + \; k^2\, f_{22})|_{(a\,+\,ch,\, b+ck)} \tag{31}$$

If Q(0) $\neq 0$, the sign of Q(c) will be the same as the sign of Q(0) for sufficiently small values of h and k. We can predict the sign of

$$Q(0) \; = h^2\, f_{11}\,(a,\, b) \; + \; 2\, h\, k\, f_{12}\,(a, b) \; + \; k^2\, f_{22}\,(a, b) \tag{32}$$

from the signs of $f_{11} \; and \; f_{11}\, f_{22} \; - \; f_{12}^2$ at the point (a, b). Multiply both sides of equation 32 by f_{11} and rearrange to obtain

$$\begin{aligned} f_{11}\, Q(0) \; &= f_{11}\,(h^2\, f_{11} \; + \; 2\, h\, k\, f_{12} \; + \; k^2\, f_{22}) \\ &= f_{11}^{\,2}\, h^2 \; + \; 2\, h\, k \; f_{11}\, f_{12} \; + \; k^2\, f_{11}\, f_{22} \\ &= f_{11}^{\,2}\, h^2 \; + \; 2\, h\, k \; f_{11}\, f_{12} \; + \; k^2\, f_{12}^2 \; - \; k^2\, f_{12}^2 \; + \; k^2\, f_{11}\, f_{22} \\ &= (\, h\, f_{11} \; + \; k\, f_{12}\,)^2 \; - \; k^2\, f_{12}^2 \; + \; k^2\, f_{11} \; f_{22} \\ &= (\, h\, f_{11} \; + \; k\, f_{12}\,)^2 \; + \; (\, f_{11}\, f_{22} \; - \; f_{12}^2\,)\, k^2 \end{aligned} \tag{33}$$

Now from equation 33 we can see that

a: If $f_{11} < 0 \; and \; f_{11}\, f_{22} \; - \; f_{12}^2 \; > \; 0$ at (a, b), then Q(0) < 0 for all sufficiently small values of h and k. This is clear because the right hand side of equation 33 will be positive if the discriminant is positive. With the right hand side positive, the left hand side must be positive which implies that with $f_{11} \; < \; 0$, Q(0) must also be negative. Thus f has a local maximum at (a, b) because f evaluated at points close to (a, b) is less than f at (a, b).

b: If $f_{11} \; > \; 0 \; and \; f_{11}\, f_{22} \; - \; f_{12}^2 \; > \; 0$ at (a, b), then Q(0) > 0 for all sufficiently small values of h and k. This is clear because the right hand side of equation 33 will be positive if the discriminant is positive. With the right hand side positive, the left hand side must be positive which implies that with $f_{11} \; > \; 0$, Q(0) must also be positive. Thus f has a local minimum at (a, b) because f evaluated at points close to (a, b) is greater than f at (a, b).

c: If $f_{11}\, f_{22} \; - \; f_{12}^2 \; < \; 0$ at (a, b), there are combinations of arbitrarily small nonzero values of h and k for which Q(0) > 0 and other values for which Q(0) < 0. Arbitrarily close to the point P_0(a, b, f(a, b)) on the surface y = f(x_1, x_2) there are points above P_0 and points below P_0, so f has a saddle point at (a, b).

d: If $f_{11}\, f_{22} \; - \; f_{12}^2 \; = \; 0$, another test is needed. The possibility that Q(0) can be zero prevents us from drawing conclusions about the sign of Q(c).