Chapter 4
Maxima and Minima on Open Sets

Summary *We derive, using critical points and the Hessian, a method of locating local maxima, local minima and saddle points of a real-valued function defined on an open subset of* $\mathbb{R}^n$.

We turn to the problem of finding local maxima and local minima of a real-valued function f defined on an open subset U of $\mathbb{R}^n$. The set of critical points of f on U, $\{X; \nabla f(X) = 0\}$, will include all points where f achieves a local maximum or minimum but may contain additional points such as saddle points. A critical point P is a *saddle point* if f, restricted to some curve passing through P, has a local maximum at P, while it has a local minimum at P along some other curve passing through P.

The *Hessian* of f at $P \in U$, $H_{f(P)}$, is defined as the $n \times n$ matrix $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(P)\right)_{1 \leq i, j \leq n}$. To define the Hessian[1] we are, of course, assuming that all first- and second-order partial derivatives of f exist. We use the convention that the order of differentiation is from right to left, i.e.

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right).$$

If all these second-order partial derivatives exist and are continuous then the order of differentiation is immaterial and

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(P) = \frac{\partial^2 f}{\partial x_j \partial x_i}(P)$$

for all i and j. In this case $H_{f(P)}$ is a symmetric $n \times n$ matrix. We will not prove this result but provide, in Exercise 4.6, a practical method which proves that all functions you will probably ever encounter have this property. If $\mathbf{v} = (v_1, \ldots, v_n)$ is a row

[1] The notation $\nabla^2(f)$ is also used for the Hessian of f.

S. Dineen, *Multivariate Calculus and Geometry*,
Springer Undergraduate Mathematics Series, DOI: 10.1007/978-1-4471-6419-7_4,

vector and ${}^t\mathbf{v}$ is the corresponding column vector then

$$\frac{\partial f}{\partial \mathbf{v}}(P) = \sum_{i=1}^{n} v_i \frac{\partial f}{\partial x_i}(P)$$

and

$$\begin{aligned}
\frac{\partial^2 f}{\partial \mathbf{v}^2}(P) = \frac{\partial}{\partial \mathbf{v}}\left(\frac{\partial f}{\partial \mathbf{v}}\right)(P) &= \sum_{i=1}^{n} v_i \frac{\partial}{\partial \mathbf{v}}\left(\frac{\partial f}{\partial x_i}\right)(P) \\
&= \sum_{i=1}^{n} v_i \Big(\sum_{j=1}^{n} v_j \frac{\partial^2 f}{\partial x_j \partial x_i}(P)\Big) \\
&= \sum_{i,j=1}^{n} v_i v_j \frac{\partial^2 f}{\partial x_i \partial x_j}(P) \\
&= \mathbf{v} H_{f(P)} {}^t\mathbf{v}.
\end{aligned}$$

The main theoretical result on the existence of local maxima, local minima and saddle points is the following theorem.

Theorem 4.1 *If U is an open subset of $\mathbb{R}^n$, $f: U \to \mathbb{R}$ is a twice continuously differentiable function on U and P is a critical point of f, i.e.*

$$\nabla f(P) = \left(\frac{\partial f}{\partial x_1}(P), \ldots, \frac{\partial f}{\partial x_n}(P)\right) = 0$$

then

(1) *f has a strict local maximum at P if $\frac{\partial^2 f}{\partial \mathbf{v}^2}(P) < 0$ for all $\mathbf{v} \neq 0$,*
(2) *f has a strict local minimum at P if $\frac{\partial^2 f}{\partial \mathbf{v}^2} > 0$ for all $\mathbf{v} \neq 0$,*
(3) *f has a saddle point at P if there exist $\mathbf{v}$ and $\mathbf{w}$ such that*

$$\frac{\partial^2 f}{\partial \mathbf{v}^2}(P) < 0 < \frac{\partial^2 f}{\partial \mathbf{w}^2}(P).$$

To derive a practical test from this result we use linear algebra and Lagrange multipliers. To simplify matters we change our notation and let $\mathbf{v} = X = (x_1, \ldots, x_n)$, $a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(P)$ and $A = (a_{ij})_{1 \le i,j \le n}$. With this notation

$$\frac{\partial^2 f}{\partial \mathbf{v}^2}(P) = \mathbf{v} H_{f(P)} {}^t\mathbf{v} = X A {}^t X.$$

Since

$$\frac{X}{\|X\|}A^t\Big(\frac{X}{\|X\|}\Big) = \frac{1}{\|X\|^2}XA^tX$$

for $X \neq 0$ we have

$$XA^tX > 0 \quad \text{for all } X \neq 0 \iff \min_{\|X\|=1} XA^tX > 0$$
$$XA^tX < 0 \quad \text{for all } X \neq 0 \iff \max_{\|X\|=1} XA^tX < 0$$

there exists $X,\ Y \in \mathbb{R}^n$ such that

$$XA^tX < 0 < YA^tY \iff \min_{\|X\|=1} XA^tX < 0 < \max_{\|X\|=1} XA^tX.$$

We thus need to examine the extreme values of XA^tX on the set

$$\|X\|^2 = \langle X, X\rangle = \sum_{i=1}^{n} x_i^2 = 1.$$

Let

$$h(X) = h(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij}x_ix_j = XA^tX$$

and

$$g(X) = g(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} x_i^2 = \langle X, X\rangle.$$

Since the set $g^{-1}(1)$ is compact and h is continuous the fundamental existence theorem implies that h has a maximum and minimum on $g^{-1}(1)$. Using the coordinate expansion of g we see that $\nabla g(X) = (2x_1, \ldots, 2x_n) = 2X$ and $\nabla g(X) \neq 0$ on $g^{-1}(1)$. Hence we may apply the method of Lagrange multipliers to find the maximum and the minimum of h on the set $g^{-1}(1)$. We have

$$\nabla h(X) = \Bigg(\sum_{j=1}^{n} 2a_{1j}x_j, \ldots, \sum_{j=1}^{n} 2a_{nj}x_j\Bigg) = 2XA.$$

By the method of Lagrange multipliers there exists at the maximum and minimum points of h on $g^{-1}(1)$ a real number λ such that

$$2XA = 2\lambda X.$$

Taking the transpose we get

$$A^tX = \lambda^tX = \lambda I^tX$$

i.e.

$$(A - \lambda I)^tX = 0. \tag{4.1}$$

Since $\|X\| = 1$, any λ which satisfies (4.1) is an *eigenvalue*[2] of A and, moreover,

$$h(X) = XA^tX = \lambda X^tX = \lambda\langle X, X\rangle = \lambda.$$

Thus the maximum and minimum values of h are eigenvalues of A and are achieved at the corresponding unit eigenvectors. If all eigenvalues are positive then h is always positive and f has a local minimum at P, if all eigenvalues are negative then h is always negative and f has local maximum at P and if some are positive and some negative then h takes positive and negative values and f has a saddle point at P. If λ is an eigenvalue of A the set

$$E_\lambda := \{X \in \mathbb{R}^n : A^tX = \lambda^tX\}$$

is a subspace of $\mathbb{R}^n$, called the λ-*eigenspace* of A, and the dimension of E_λ is called the *multiplicity* of the eigenvalue λ. An $n \times n$ symmetric matrix has n eigenvalues when eigenvalues are counted according to multiplicity, i.e. if E_λ is j–dimensional then λ is counted j times, and $\det(A) = \lambda_1 \cdots \lambda_n$. Since a particular case of the above result will play an important role in our study of Gaussian curvature in Chap. 16 we display it separately.

Proposition 4.2 *If $A = \left(a_{ij}\right)_{1\le i,j\le 2}$ is a symmetric 2×2 matrix with eigenvalues λ_1 and λ_2, $\lambda_1 \ge \lambda_2$, then*

$$\max.\{a_{11}x^2 + 2a_{12}xy + a_{22}y^2 : x^2 + y^2 = 1\} = \lambda_1$$

and

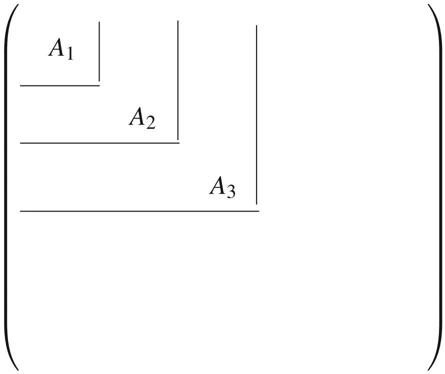
$$\min.\{a_{11}x^2 + 2a_{12}xy + a_{22}y^2 : x^2 + y^2 = 1\} = \lambda_2.$$

Moreover, the maximum and minimum are achieved at eigenvectors of A and

$$\det(A) = \lambda_1\lambda_2.$$

For our test we just need to know the sign of the largest and smallest eigenvalues. Since eigenvalues may be difficult to calculate we will use a reasonably well-known result from linear algebra. If A is a square matrix then the $k \times k$ matrix A_k obtained by deleting all except the first k rows and k columns of A is called the $k \times k$ *principal minor* of A. We have $A_n = A$.

[2] We could refocus this analysis to show that any symmetric $n \times n$ matrix with real entries admits eigenvalues.



We require the following result:

if A is a symmetric $n \times n$ matrix then all eigenvalues of A are positive if and only if $\det(A_k) > 0$ *for all k.*

Proposition 4.3 *If $f: U \subset \mathbb{R}^n \to \mathbb{R}$, U open, has continuous first- and second-order partial derivatives, P is a critical point of f, and $A = H_{f(P)}$ is the Hessian of f at P then the following hold:*

(a) *if* $\det(A_{2k}) < 0$ *for some k then P is a saddle point of f,*
(b) *if* $\det(A_n) \neq 0$ *then*

  (b1) *if* $\det(A_k) > 0$ *for all k then f has a strict local minimum at P,*
  (b2) *if* $(-1)^k \det(A_k) > 0$ *for all k then f has a strict local maximum at P,*
  (b3) *if the conditions on the determinants in (b1) and (b2) do not apply then f has a saddle point at P.*

Proof (a) Suppose $\det(A_{2k}) < 0$ for some positive integer k, $2k \leq n$. If $P = (p_1, \ldots, p_n)$ let $Q = (p_1, \ldots, p_{2k})$. Consider the function $g: V \subset \mathbb{R}^{2k} \to \mathbb{R}$ defined by

$$g(x_1, \ldots, x_{2k}) = f(x_1, \ldots, x_{2k}, p_{2k+1}, \ldots, p_n)$$

where V is the open set in $\mathbb{R}^{2k}$ consisting of all $(x_1, \ldots, x_{2k})$ such that

$$(x_1, \ldots, x_{2k}, p_{2k+1}, \ldots, p_n) \in U.$$

It is easily checked that

$$H_{g(Q)} = \left(\frac{\partial^2 g}{\partial x_i \partial x_j}(Q)\right)_{1 \leq i,j \leq 2k} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(P)\right)_{1 \leq i,j \leq 2k} = A_{2k}.$$

Let $\beta_1, \ldots, \beta_{2k}$ denote the $2k$ eigenvalues of the symmetric $2k \times 2k$ matrix A_{2k} counted according to multiplicity. We have

$$\beta_1 \cdots \beta_{2k} = \det(A_{2k}) < 0.$$

Since $2k$ is an even integer it follows that A_{2k} has positive and negative eigenvalues. Hence there exist $\mathbf{u} = (u_1, \ldots, u_{2k})$ and $\mathbf{v} = (v_1, \ldots, v_{2k})$ such that

$$\mathbf{u}A_{2k}{}^t\mathbf{u} < 0 < \mathbf{v}A_{2k}{}^t\mathbf{v}.$$

Let $\mathbf{w}_1 = (u_1, \ldots, u_{2k}, 0, \ldots, 0)$ and $\mathbf{w}_2 = (v_1, \ldots, v_{2k}, 0, \ldots, 0)$. Then

$$\begin{aligned}\frac{\partial^2 f}{\partial \mathbf{w}_1{}^2}(P) &= \mathbf{w}_1 H_{f(P)}{}^t\mathbf{w}_1 \\ &= \mathbf{u}A_{2k}{}^t\mathbf{u} < 0 < \mathbf{v}A_{2k}{}^t\mathbf{v} \\ &= \mathbf{w}_2 H_{f(P)}{}^t\mathbf{w}_2 = \frac{\partial^2 f}{\partial \mathbf{w}_2{}^2}(P)\end{aligned}$$

and f has a saddle point at P. This proves (a).
(b) The first part, (b1), follows directly from the linear algebra result quoted above. Since f has a local maximum at P if and only if $(-f)$ has a local minimum at the same point and $H_{-f(P)} = -H_{f(P)}$ it follows, from (b1), that f has a local maximum at P if $\det(-A_k) > 0$ for all k. Hence (b2) follows since $\det(-A_k) = (-1)^k \det(A_k)$. If (b1) and (b2) do not apply then, since all eigenvalues of A are non-zero, A has positive and negative eigenvalues and hence f has a saddle point at P. This completes the proof. □

If $\det(H_{f(P)}) = 0$ we call P a *degenerate critical point* of f (all other critical points are called *non-degenerate*) and higher order derivatives may be required to test the nature of the critical point.

Proposition 4.3 enables us to classify all non-degenerate and some degenerate critical points. When P is a non-degenerate critical point of f and f has a local maximum or minimum at P then the determinants of the *odd* principal minors all have the same sign. This leads to the second test for saddle points given below and a little reflection shows that it may be applied at any critical point. In practice the determinants, $\det(A_i)$, are calculated in the order $i = 1, 2, 3, \ldots$ and testing for *saddle points* is carried out as the calculations proceed. The critical point is a saddle point and the calculations stop when for the first time either of the following is observed:

$$\det(A_{2k}) < 0$$

$$\det(A_{2k-1}) \det(A_{2k+1}) < 0.$$

If (b1) or (b2) are used to find a local maximum or minimum then all determinants of the Hessian must be calculated.

Example 4.4 Let $f(x, y, z) = x^2y^2 + z^2 + 2x - 4y + z$. We have

$$\nabla f(x, y, z) = (2xy^2 + 2, 2x^2y - 4, 2z + 1).$$

If P is a critical point of f then

$$\begin{aligned} 2xy^2 + 2 &= 0 \\ 2x^2y - 4 &= 0 \\ 2z + 1 &= 0 \end{aligned}$$

Hence $z = -1/2$ from the third equation. From the first two equations we see that x and y are non-zero. Hence $xy^2 = -1$ and $x^2y = 2$ imply $xy^2/x^2y = -1/2 = y/x$ and $x = -2y$. We have $-2y \cdot y^2 = -1$, i.e. $y^3 = 1/2$ and $y = 2^{-1/3}$. From $x = -2y$ we obtain $x = -2^{2/3}$ and conclude that $(-2^{2/3}, 2^{-1/3}, -1/2)$ is the only critical point of f. A simple calculation shows that

$$H_{f(x,y,z)} = \begin{pmatrix} 2y^2 & 4xy & 0 \\ 4xy & 2x^2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

and hence

$$H_{f(-2^{2/3},2^{-1/3},1/2)} = \begin{pmatrix} 2^{1/3} & -4 \cdot 2^{1/3} & 0 \\ -4 \cdot 2^{1/3} & 2 \cdot 2^{4/3} & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Since $\det(2^{1/3}) > 0$ and

$$\begin{aligned} \det\begin{pmatrix} 2^{1/3} & -4 \cdot 2^{1/3} \\ -4 \cdot 2^{1/3} & 2 \cdot 2^{4/3} \end{pmatrix} &= 2 \cdot 2^{5/3} - 16 \cdot 2^{2/3} \\ &= 4 \cdot 2^{2/3} - 16 \cdot 2^{2/3} < 0 \end{aligned}$$

the critical point $(-2^{2/3}, 2^{-1/3}, -1/2)$ is a saddle point of f.

Example 4.5 We wish to find and classify the non-degenerate critical points of $f(x, y, z) = x^2y + y^2z + z^2 - 2x$. We have

$$\nabla f(x, y, z) = (2xy - 2, x^2 + 2yz, y^2 + 2z)$$

and the critical points satisfy the equations

$$2xy - 2 = 0, \quad x^2 + 2yz = 0 \quad \text{and} \quad y^2 + 2z = 0.$$

Substituting $z = -y^2/2$ into the second equation implies $y^3 = x^2$. Hence, the first equation shows $y^{5/2} = 1$ and we have $y = 1$ and $z = -1/2$. From $xy = -1$ we get

$x = 1$ and $(1, 1, -1/2)$ is the only critical point of f. We have

$$H_{f(x,y,z)} = \begin{pmatrix} 2y & 2x & 0 \\ 2x & 2z & 2y \\ 0 & 2y & 2 \end{pmatrix}$$

and

$$H_{f(1,1,-1/2)} = \begin{pmatrix} 2 & 2 & 0 \\ 2 & -1 & 2 \\ 0 & 2 & 2 \end{pmatrix}.$$

Since $\det(2) > 0$ and

$$\det\begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix} = -2 - 4 < 0$$

the point $(1, 1, -1/2)$ is a saddle point of f.

A number of simple initial checks may be carried out to detect saddle points. These are based on rearranging the variables, a procedure which does not alter the nature of a critical point. Thus we may first interchange row i and row j and afterwards column i and column j in the Hessian and then apply our standard test. In this way any 2×2 sub-matrix of the Hessian, which is symmetric about the diagonal, can be moved to become the 2×2 principal minor. Since a 2×2 symmetric matrix with positive and negative diagonal entries has negative determinant this means that a critical point is a saddle point if the diagonal contains positive and negative terms. This applies, for instance, to Example 4.5. Another useful indicator is the presence of a relatively large non-diagonal term; if a_{ij}, the ijth term in the Hessian, satisfies $|a_{ij}| > |a_{kk}|$ for all k then, as we see in our next example, this also implies that the critical point is a saddle point.

It is *not necessary* to physically interchange these rows and columns as the appropriate sub-matrix can be isolated by inspection but to convince you that this process is valid we include this step in our next example.

Example 4.6 Suppose the Hessian at a critical point P is

$$A = \begin{pmatrix} 2 & 2 & 9 & 1 \\ 2 & 3 & 2 & 0 \\ 9 & 2 & 4 & 2 \\ 1 & 0 & 2 & -1 \end{pmatrix}.$$

Since the diagonal contains positive and negative terms P is a saddle point. We see this more clearly if we interchange row 2 and row 4 and column 2 and column 4. We obtain the matrix B below and $\det(B_2) < 0$.

The presence of 9 in the third column suggests we interchange the second and third rows and the second and third columns. This gives the matrix C below with $\det(C_2) < 0$ and confirms that P is a saddle point.

$$B = \begin{pmatrix} 2 & 1 & 9 & 2 \\ 1 & -1 & 2 & 0 \\ 9 & 2 & 4 & 2 \\ 2 & 0 & 2 & 3 \end{pmatrix}, C = \begin{pmatrix} 2 & 9 & 2 & 1 \\ 9 & 4 & 2 & 1 \\ 2 & 2 & 3 & 0 \\ 1 & 2 & 0 & -1 \end{pmatrix}.$$

The standard approach gives $\det(A_1) = 2$, $\det(A_2) = 2$, $\det(A_3) = -171$ and $\det(A_1)\det(A_3) < 0$ implies that P is a saddle point.

The above method does not identify absolute or global maxima and minima. We encountered a similar problem in Example 3.3. We now describe a useful method which can be applied to certain functions on convex open sets. A subset $U \subset \mathbb{R}^n$ is *convex* if the straight line joining any two points in U is contained in U. The interior of a circle, sphere, box, polygon, the first quadrant or octant, and the upper half-plane are typical examples of convex open sets. The exterior of a circle or polygon is not convex.

Suppose $f : U(\text{open, convex}) \longrightarrow \mathbb{R}$ has continuous first and second order partial derivatives at all points. If $P, Q \in U$ and $\mathbf{v} = Q - P$ then, by convexity, $g(x) = f(P + x\mathbf{v})$ is defined on an *open interval* in $\mathbb{R}$ and is the restriction of f to that part of the straight line through P and Q which lies in U. Moreover, $g'(x) = \dfrac{\partial f}{\partial \mathbf{v}}(P + x\mathbf{v})$ and $\dfrac{\partial^2 f}{\partial \mathbf{v}^2}(P + x\mathbf{v}) = g''(x)$. Suppose $\dfrac{\partial^2 f}{\partial \mathbf{v}^2}(X) \neq 0$ for all X in U and all $\mathbf{v} \neq 0$. By the Intermediate Value Theorem $g''(x)$ is either always positive or always negative. This implies that g' is either strictly increasing or strictly decreasing and in either case g has at most one critical point on every line in U. Since we can carry out this analysis for any pair of points P and Q in U this shows that f has at most one critical point in U. Suppose f has a local maximum at P. If $f(P) < f(Q)$ then the function g considered above must have a local minimum at some point on the line joining P and Q. This contradicts the fact that g has at most one critical point and shows that f has an absolute maximum over U at P. Similarly if f has a local minimum at P then it has an absolute minimum over U at P. To verify that $\dfrac{\partial^2 f}{\partial \mathbf{v}^2}(X) \neq 0$ when U is a convex open subset in $\mathbb{R}^n$ it suffices to show that $\det(H_{f(x,y)})$ is never zero.

Example 4.7 Let

$$f(x, y) = x + y - e^x - e^y - e^{x+y}$$

for all (x, y) in $\mathbb{R}^2$. We have

$$\nabla f(x, y) = (1 - e^x - e^{x+y}, 1 - e^y - e^{x+y}).$$

If (x, y) is a critical point of f then $e^x = 1 - e^{x+y} = e^y$ and $x = y$. Hence $e^{2x} + e^x - 1 = 0$. If $w = e^x$ then $w^2 + w - 1 = 0$ and $w = (-1 \pm \sqrt{5})/2$. Since $w = e^x > 0$ we have only one solution $w = (-1 + \sqrt{5})/2$ and f has just one critical point at $P = ((-1 + \sqrt{5})/2, (-1 + \sqrt{5})/2)$. The existence of a single critical point suggests that we consider the Hessian at all points. We have

$$H_{f(x,y)} = \begin{pmatrix} -e^x - e^{x+y} & -e^{x+y} \\ -e^{x+y} & -e^y - e^{x+y} \end{pmatrix}.$$

Hence $\det(H_{f(x,y)}) = (e^x + e^{x+y})(e^y + e^{x+y}) - e^{2x+2y} = e^{2x+y} + e^{y+2x} + e^{x+y}$ is positive at all points $(x, y) \in \mathbb{R}^2$. Since $\dfrac{\partial^2 f}{\partial x^2}(P) < 0$, f has a local and hence a global maximum at P.

Exercises

4.1 Classify the non-degenerate critical points of

(a) $x^2 + xy + 2x + 2y + 1$	(f) $xyze^{-x^2-y^2-z^2}$
(b) $x^3 + y^3 - 3xy$	(g) $xy + x^2z - x^2 - y - z^2$
(c) $x^3z - 192x + y^2 - yz$	(h) $x^2y + y^2z + z^2 - 8\sqrt{2}x$
(d) $(2 - x)(4 - y)(x + y - 3)$	(i) $2x^2y^2 - z^2 + x - 2y + z$
(e) $4xyz - x^4 - y^4 - z^4$	(j) $x^2 - xy + yz^3 - 6z$.

Show, using convexity, that the function in (b) has an absolute minimum over the set $U = \{(x, y) : x > 1/2, y > 1/2\}$. Show using the exhaustion method outlined in Example 3.3 that the function in (e) has an absolute maximum and no absolute minimum over $\mathbb{R}^3$. Show that the function in (f) has an absolute maximum and an absolute minimum over $\mathbb{R}^3$.

4.2 If $f(x, y, z) = (ax^2 + by^2 + cz^2)e^{-x^2-y^2-z^2}$ and $a > b > c > 0$ show that the function has two local maxima, one local minimum and four saddle points. Find the maximum and minimum of f over $\mathbb{R}^3$.

4.3 Show that the function $xyz(x+y+z-1)$ has one non-degenerate critical point and an infinite set of degenerate critical points. Show that the non-degenerate critical point is a local minimum.

4.4 Show that every critical point of $\dfrac{x^3 + y^3 + z^3}{xyz}$ is degenerate.

4.5 Find the distance from the point $(-1, 1, 1)$ to the level set $z = xy$.

4.6 Let $U = (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n) \subset \mathbb{R}^n$ and let

$$M(U) = \Big\{ f : U \to \mathbb{R} : \text{ all first- and second-order partial derivatives of } f \text{ exist and } \frac{\partial^2 f}{\partial x_i \partial x_j}(P) = \frac{\partial^2 f}{\partial x_j \partial x_i}(P) \text{ for all } P \in U \text{ and all } i, j, 1 \le i, j \le n \Big\}.$$

Show that $M(U)$ has the following properties:

(i) if $f,\ g \in M(U)$ and $c \in \mathbb{R}$ then $f \pm g$, $f \cdot g$ and $c \cdot f \in M(U)$ and if $g \neq 0$ then $f/g \in M(U)$
(ii) if $f \in M(U)$ and $\phi\colon \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable then $\phi \circ f \in M(U)$.

Using (i) and (ii) show that

$$h(x, y, z) = \sin^2\left(\frac{e^{xyz}}{y^2 + z^2 + 1}\right)$$

lies in $M(U)$. Verify this result by calculating directly the appropriate second-order partial derivatives of h.

4.7 Let $Y_1, \ldots, Y_m$ be m points in $\mathbb{R}^n$. Show that $\sum_{i=1}^{m} \|X - Y_i\|^2$ achieves its absolute minimum at $X = \frac{1}{m}\sum_{i=1}^{m} Y_i$. Interpret your result geometrically.

4.8 If $z = \phi(x, y)$ satisfies the equation

$$x^2 + 2y^2 + 3z^2 - 2xy - 2yz = 2$$

find the points (x, y) at which ϕ has a local maximum or a local minimum.

4.9 If $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ has continuous first and second order partial derivatives show that $f''(x_1, \ldots, x_n) = (f')'(x_1, \ldots, x_n) = H_{f(x_1,\ldots,x_n)}$.

4.10 Show directly that the set $F^{-1}(2, 1)$ in Example 3.1 is compact.