Contents

Chapter 1

Theory of maxima and minima

The classical theory of maxima and minima provides analytical methods for finding solutions of optimization problems involving continuous and differentiable functions and making use of differential calculus. Applications of these techniques may be limited since many of the practical problems involve functions that are not continuous or differentiable. But the theory of maxima and minima is the fundamental starting point for numerical methods of optimization and a basis for advanced topics like calculus of variations or optimal control.

This chapter presents the necessary and sufficient conditions in locating the optimum solutions for unconstrained and constrained optimization problems for single-variable functions or multivariable functions.

1.1 Statement of an optimization problem. Terminology

Given a vector of n independent variables:

$$\mathbf{x} = (x_1 \; x_2 \; \dots \; x_n)^T \tag{1.1}$$

and a scalar function:

$$f : \mathbb{R}^n \to \mathbb{R} \tag{1.2}$$

an optimization problem (P) can be formulated as follows:

$$\min_{\mathbf{x}} f(\mathbf{x}) \quad \Big(\text{or} \quad \max_{\mathbf{x}} f(\mathbf{x})\Big) \tag{1.3}$$

subject to:

$$g_i(\mathbf{x}) = 0, \quad i = \overline{1, m} \tag{1.4}$$

$$h_j(\mathbf{x}) \leq 0, \quad j = \overline{1, p} \tag{1.5}$$

The goal of the problem is to find the vector of parameters $\mathbf{x}$ which minimizes (or maximizes) a given scalar function, possibly subject to some restrictions on the allowed parameter values. The function f to be optimized is termed *the objective function*; the elements of vector $\mathbf{x}$ are *the control* or *decision variables*; the restrictions (1.4) and (1.5) are the equality or inequality *constraints*.

The value $\mathbf{x}^*$ of the variable which solves the problem is a *minimizer* (or *maximizer*) of function f subject to the constraints (1.4) and (1.5), and $f(\mathbf{x}^*)$ is the minimum (or maximum) value of the function subject to the same constraints.

If the number of constraints $m + p$ is zero, the problem is called an *unconstrained optimization problem*.

The *admissible set* or *feasible region* of (P), denoted S, is defined as:

$$S = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) = 0, h_j(\mathbf{x}) \leq 0\} \tag{1.6}$$

Example 1.1 *Consider the problem:*

$$\min_{(x_1,x_2)} (1 - x_1)^2 + (1 - x_2)^2 \tag{1.7}$$

subject to

$$x_1 + x_2 - 1 \leq 0, \tag{1.8}$$
$$x_1^3 - x_2 \leq 0 \tag{1.9}$$

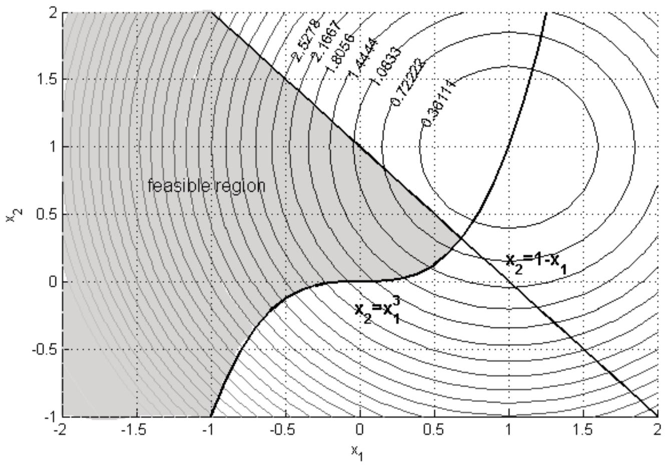

Figure 1.1: Feasible region

The function to be minimized is $f(x_1, x_2) = (1 - x_1)^2 + (1 - x_2)^2$ *and the constraint functions:* $h_1(x_1, x_2) = x_1 + x_2 - 1$ *and* $h_2(x_1, x_2) = x_2^3 - x_2$.

Figure 1.1 shows the contour plot of the objective function, i.e. the curves in two dimensions on which the value of the function $f(x_1, x_2)$ *is constant. The feasible region, obtained as the area on which the inequalities (1.8) and (1.9) hold, is shown in the same figure.*

As shown in Figure 1.2, the minimum values of a function f are the maximum of $-f$. Therefore, the optimization problems will be stated, in general, as minimization problems. The term *extremum* includes both maximum and minimum.

A point x_0 is a *global minimum* of $f(x)$ if:

$$f(x_0) < f(x), \quad \forall x \in S \tag{1.10}$$

A point x_0 is a *strong local minimum* if there exists some $\Delta > 0$, such that:

$$f(x_0) < f(x), \quad \text{when} \quad |x - x_0| < \Delta \tag{1.11}$$

A point x_0 is a *weak local minimum* if there exists some $\Delta > 0$, such that:

$$f(x_0) \leq f(x), \quad \text{when} \quad |x - x_0| < \Delta \tag{1.12}$$

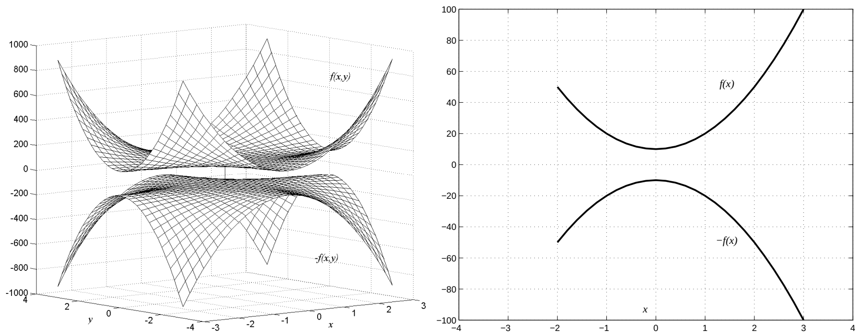

Figure 1.2: min $f(x)$ = max$(-f(x))$

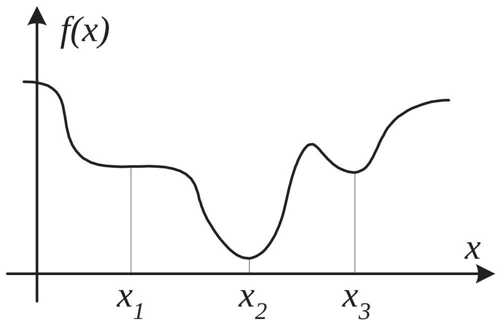

Figure 1.3: Minimum points. x_1: weak local minimum, x_2: global minimum, x_3: strong local minimum

1.2 Unconstrained optimization

1.2.1 Necessary conditions for maxima and minima

The existence of a solution to an optimization problem (P), for a continuous function f, is guaranteed by the *extreme value theorem* of Weierstrass, which states:

Theorem 1.1 *If a function $f(x)$ is continuous on a closed interval $[a, b]$, then $f(x)$ has both a maximum and a minimum on $[a, b]$. If $f(x)$ has an extremum on an open interval (a, b), then the extremum occurs at a critical point, (Renze and Weisstein, 2004).*

A single variable function $f(x)$ has *critical points* at all points x_0, where the first derivative is zero ($f'(x_0) = 0$), or $f(x)$ is not differentiable.

A function of several variables $f(\mathbf{x})$ has critical points, where the gradient is zero or partial derivatives are not defined.

In general, a *stationary point* of a function $f(\mathbf{x})$ is a point for which the gradient vanishes:

$$\nabla f(\mathbf{x}_0) = 0 \tag{1.13}$$

where

$$\nabla f(\mathbf{x}_0) = \left(\frac{\partial f}{\partial x_1} \; \frac{\partial f}{\partial x_2} \; \dots \; \frac{\partial f}{\partial x_n} \right)^T \tag{1.14}$$

A stationary point of a single variable function is a point where the first derivative $f(x_0)$ equals zero.

Example 1.2 *Consider the cases in the Figure 1.4.*

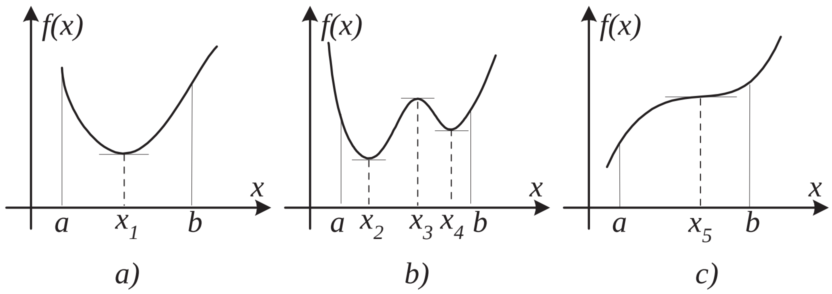

Figure 1.4: Stationary points

a) *In Figure 1.4 a) the point x_1 is the global minimizer.*

b) *In Figure 1.4 b) there are three stationary points: x_2 is a global minimizer, x_3 - a local maximizer and x_4 - a local minimizer.*

c) *In Figure 1.4 c) x_5 is a stationary point because the first derivative of f vanishes, but it is not a maximizer nor a minimizer; x_5 is an inflection point. For $f\colon [a, b] \to \mathbb{R}$ the point $x = a$ is the global minimizer on the interval $[a, b]$.*

The only case, when the minimizer (or maximizer) is not a stationary point, is when the point is one of the endpoints of the interval $[a, b]$, on which f is defined. That is, any point interior to this interval, that is a maximum must be a stationary point.

The first-order condition: *The necessary condition for a point x to be a minimizer (or a maximizer) of the function $f\colon [a, b] \to \mathbb{R}$ is: if $x \in (a, b)$, then x is a stationary point of f.*

Local extrema of a function $f\colon [a, b] \to \mathbb{R}$ may occur only at:

- boundaries
- stationary points (the first derivative of f is zero)

In case the function is non-differentiable at some points in $[a, b]$, the function has also extreme values.

For a function of n independent variables $f\colon \mathbb{R}^n \to \mathbb{R}$, the *necessary condition* for a point $\mathbf{x}_0 = (x_1\ x_2\ \ldots x_n)^T$ to be an extremum is the gradient equals zero.

For functions of n variables, the stationary points can be: minima, maxima or saddle points.

1.2.2 Sufficient conditions for maxima and minima

Since not all stationary points are necessarily minima or maxima (they can be also inflection or saddle points) we may be able to determine their character by examining the second derivative of the function at the stationary point. These sufficient conditions will be developed for single variable functions and then extended for two or n variables based on the same concepts. The global minimum or maximum has to be located by comparing all local maxima and minima.

1.2.2.1 Sufficient conditions for single variable functions

Second-order conditions for optimum. Let f be a single variable function with continuous first and second derivatives, defined on an interval S, $f\colon S \to \mathbb{R}$, and x_0 is a stationary point of f so that $f'(x_0) = 0$.

The Taylor series expansion about the stationary point x_0 is one possibility to justify the second-order conditions:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + higher\ order\ terms \tag{1.15}$$

For the points x sufficiently close to x_0 so the higher order terms become negligible compared to the second-order terms, and knowing the first derivative is zero at a stationary point, the equation (1.15) becomes:

$$f(x) = f(x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 \tag{1.16}$$

Since $(x - x_0)^2$ is always positive, we can determine whether x_0 is a local maximum or minimum by examining the value of the second derivative $f''(x_0)$:

- If $f''(x_0) > 0$, the term $\frac{1}{2} f''(x_0)(x - x_0)^2$ will add to $f(x_0)$ in the equation (1.16), so the value of f at the neighboring points x is greater than $f(x_0)$. In this case x_0 is a local minimum.
- If $f''(x_0) < 0$, the term $\frac{1}{2} f''(x_0)(x - x_0)^2$ will subtract from $f(x_0)$ and the value of f at the neighboring points x is less than $f(x_0)$. In this case x_0 is a local maximum.
- If $f''(x_0) = 0$ it is necessary to examine the higher order derivatives. In general if $f''(x_0) = ... = f^{(k-1)}(x_0) = 0$, the Taylor series expansion becomes

$$f(x) = f(x_0) + \frac{1}{k!} f^{(n)}(x_0)(x - x_0)^k \tag{1.17}$$

  - If k is even, then $(x - x_0)^k$ is positive. Thus, if $f^{(k)}(x_0) > 0$, then $f(x_0)$ is a minimum; if $f^{(k)}(x_0) < 0$ then $f(x_0)$ is a maximum (Figure 1.5).
  - If k is odd, then $(x - x_0)^k$ changes sign. It is positive for $x > x_0$ and negative for $x < x_0$. If $f^{(k)}(x_0) > 0$, the second term in the equation (1.17) is positive for $x > x_0$ and negative for $x < x_0$. If $f^{(k)}(x_0) < 0$, the second term in the equation (1.17) is negative for $x > x_0$ and positive for $x < x_0$. The stationary point is an inflection point (Figure 1.5).

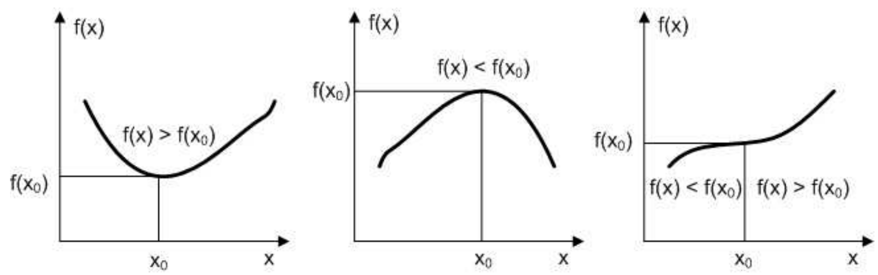

Figure 1.5: Minimum, maximum and inflection points

These results can be summarized in the following rules:

- If $f''(x_0) < 0$, then x_0 is a local maximizer.
- If $f''(x_0) > 0$, then x_0 is a local minimizer.
- If $f''(x_0) = ... = f^{(k-1)}(x_0) = 0$ and:
  - If k is even and
    - If $f^{(k)}(x_0) < 0$, then x_0 is a local maximizer.
    - If $f^{(k)}(x_0) > 0$, then x_0 is a local minimizer.
  - If k is odd, then x_0 is an inflection point.

Example 1.3 *Locate the extreme points of the following function:*

$$f(x) = \frac{x^5}{5} - \frac{x^3}{3} \tag{1.18}$$

The first derivative is:

$$f'(x) = x^4 - x^2 = x^2(x^2 - 1) = x^2(x-1)(x+1) \tag{1.19}$$

The stationary points are obtained by setting the first derivative equal to zero:

$$x_1 = 1; \;\; x_2 = -1; \;\; x_3 = x_4 = 0; \tag{1.20}$$

The second derivative:

$$f''(x) = 4x^3 - 2x \tag{1.21}$$

calculated at points $x_{1,2,3,4}$ is:

$$f''(1) = 2, \;\; f''(-1) = -2; \;\; f''(0) = 0 \tag{1.22}$$

Because $f''(1) > 0$ and $f''(-1) < 0$, the stationary point $x_1 = 1$ is a local minimum and $x_2 = -1$ is a local maximum. Since the second derivative is zero at $x_{3,4} = 0$, an analysis of higher order derivatives is necessary. The third derivative of f:

$$f'''(x) = 12x^2 - 2 \tag{1.23}$$

is nonzero for $x_{3,4} = 0$. Since the order of the first nonzero derivative is 3, i.e. it is odd, the stationary points $x_3 = x_4 = 0$ are inflection points. A plot of the function showing the local minimum, maximum and inflection points is shown in Figure 1.6.

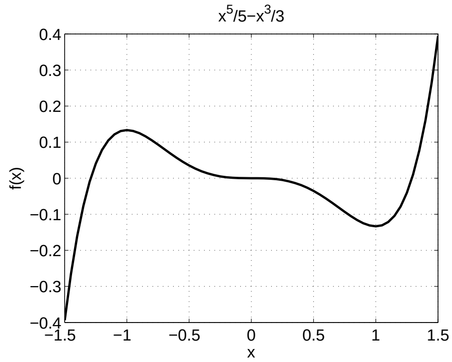

Figure 1.6: Plot of $f(x) = x^5/5 - x^3/3$

Example 1.4 *Consider the function*

$$f(x) = 4x - x^3/3 \tag{1.24}$$

Determine the maximum and minimum points in the region:

$$-3 \leq x \leq 1 \tag{1.25}$$

Compute the first derivative of $f(x)$ and set it equal to zero:

$$f'(x) = 4 - 2x^2 = 0 \tag{1.26}$$

The stationary points of the function are $x_{10} = -2$ and $x_{20} = 2$. Since the variable is constrained by (1.25) and x_{20} is out of the bounds, we shall analyze the other stationary point and the boundaries. The second derivative:

$$f''(x_{10}) = -2x_{10} = -2 \cdot (-2) = 4 > 4 \tag{1.27}$$

is positive at x_{10}, thus -2 is a local minimum, as shown in Figure 1.7. According to the theorem of Weierstrass, the function must have a maximum value in the interval $[-3, 1]$. On the boundaries, the function takes the values: $f(-3) = -3$ and $f(1) = 11/3$. Thus, the point $x = 1$ is the maximizer in $[-3, 1]$.

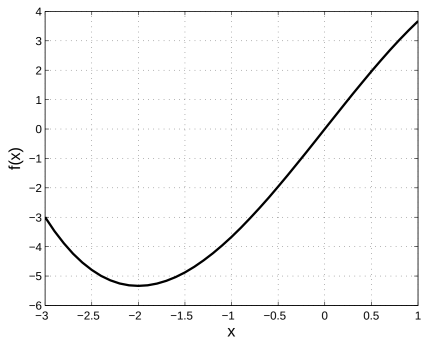

Figure 1.7: Plot of $f(x) = 4x - x^3/3$

1.2.2.2 Sufficient conditions for two independent variables

Second derivative test

Let $f(x_1, x_2)$ be a two-dimensional function and denote the first-order partial derivatives of f with respect to x_1 and x_2 by:

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = f_{x_1}, \quad \frac{\partial f(x_1, x_2)}{\partial x_2} = f_{x_2} \tag{1.28}$$

and the second-order partial derivatives:

$$\frac{\partial f(x_1, x_2)}{\partial x_i \partial x_j} = f_{x_i x_j}, \quad i, j = 1, 2 \tag{1.29}$$

If $f(x_1, x_2)$ has a local extremum at a point (x_{10},x_{20}) and has continuous partial derivatives at this point, then

$$f_{x_1}(x_{10}, x_{20}) = 0, \quad f_{x_2}(x_{10}, x_{20}) = 0 \tag{1.30}$$

The second partial derivatives test classifies the point as a local maximum or local minimum.

Define the second derivative test discriminant as:

$$D_2 = f_{x_1 x_1} f_{x_2 x_2} - f^2_{x_1 x_2} \tag{1.31}$$

Then, (Weisstein, 2004)

- If $D_2(x_{10}, x_{20}) > 0$ and $f_{x_1 x_1}(x_{10}, x_{20}) > 0$, the point is a local minimum
- If $D_2(x_{10}, x_{20}) > 0$ and $f_{x_1 x_1}(x_{10}, x_{20}) < 0$, the point is a local maximum
- If $D_2(x_{10}, x_{20}) < 0$, the point is a saddle point
- If $D_2(x_{10}, x_{20}) = 0$ the test is inconclusive and higher order tests must be used.

Note that the second derivative test discriminant, D_2, is the determinant of the Hessian matrix:

$$H_2 = \begin{bmatrix} f_{x_1 x_1} & f_{x_1 x_2} \\ f_{x_1 x_2} & f_{x_2 x_2} \end{bmatrix} \tag{1.32}$$

Example 1.5 *Locate the stationary points of the function:*

$$f(x_1, x_2) = x_1^2 - x_2^2 \tag{1.33}$$

and determine their character.

The stationary points are computed by setting the gradient equal to zero:

$$f_{x_1} = 2x_1 = 0 \quad (1.34)$$
$$f_{x_2} = -2x_2 = 0 \quad (1.35)$$

The function has only one stationary point: $(x_{10}, x_{20}) = (0, 0)$. *Compute the second derivatives:*

$$f_{x_1x_1} = 2, \quad f_{x_1x_2} = 0, \quad f_{x_2x_2} = -2 \quad (1.36)$$

and the determinant of the Hessian matrix is:

$$D_2 = \begin{vmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_1x_2} & f_{x_2x_2} \end{vmatrix} = f_{x_1x_1}f_{x_2x_2} - f^2_{x_1x_2} = 2 \cdot (-2) - 0^2 = -4 < 0 \quad (1.37)$$

According to the second derivative test, the point $(0, 0)$ *is a saddle point. A mesh and contour plot of the function is shown in Figure 1.8.*

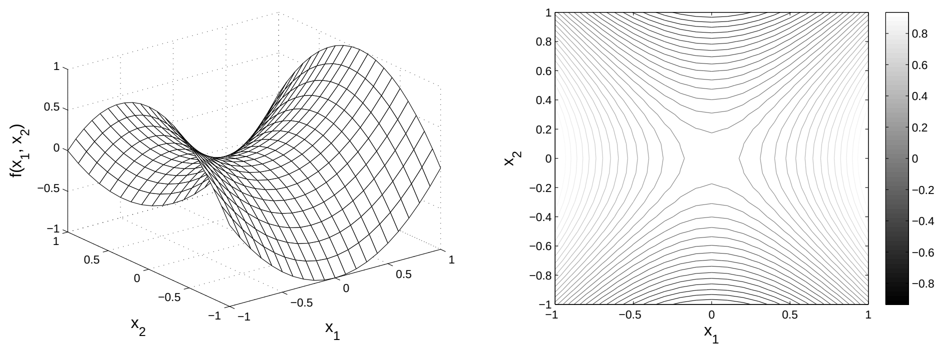

Figure 1.8: Mesh and contour plot of $f(x_1, x_2) = x_1^2 - x_2^2$

Example 1.6 *Locate the stationary points of the function:*

$$f(x_1, x_2) = 1 - x_1^2 - x_2^2 \quad (1.38)$$

and determine their character.

Compute the stationary points from:

$$f_{x_1} = -2x_1 = 0 \quad (1.39)$$
$$f_{x_2} = -2x_2 = 0 \quad (1.40)$$

The function has only one stationary point: $(x_{10}, x_{20}) = (0, 0)$. *The second derivatives are:*

$$f_{x_1x_1} = -2 < 0, \quad f_{x_1x_2} = 0, \quad f_{x_2x_2} = -2 \quad (1.41)$$

and the discriminant:

$$D_2 = \begin{vmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_1x_2} & f_{x_2x_2} \end{vmatrix} = f_{x_1x_1}f_{x_2x_2} - f^2_{x_1x_2} = (-2) \cdot (-2) - 0^2 = 4 > 0 \quad (1.42)$$

Thus, the function has a maximum at $(0, 0)$, *because* $f_{x_1x_1} < 0$ *and* $D_2 > 0$. *The graph of the function is shown in Figure 1.9.*

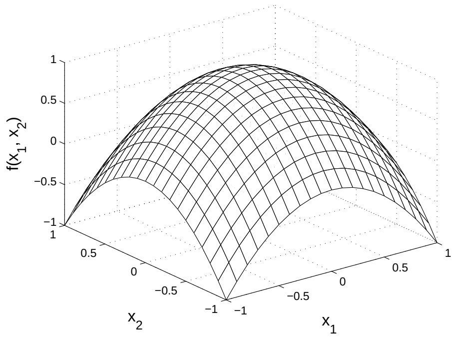

Figure 1.9: Mesh plot of $f(x_1, x_2) = 1 - x_1^2 - x_2^2$

Example 1.7 *Locate the stationary points of the function:*

$$f(x_1, x_2) = x_1^2 + 2x_1x_2 + x_2^4 - 2 \tag{1.43}$$

and determine their character.

Compute the stationary points from:

$$\begin{aligned} f_{x_1} &= 2x_1 + 2x_2 = 0 && (1.44)\\ f_{x_2} &= 2x_1 + 4x_2^3 = 0 && (1.45)\end{aligned}$$

From (1.44) and (1.45):

$$x_1 = -x_2, \;\; x_1 - 2x_1^3 = 0, \text{ or } \; x_1(1-\sqrt{2}x_1)(1+\sqrt{2}x_1) = 0$$

and the stationary points are: $(0,0)$, $(1/\sqrt{2}, -1/\sqrt{2})$ *and* $(-1/\sqrt{2}, 1/\sqrt{2})$.

The second derivatives are:

$$f_{x_1x_1} = 2, \;\; f_{x_1x_2} = 2, \;\; f_{x_2x_2} = 12x_2^2 \tag{1.46}$$

and the discriminant:

$$D_2 = \begin{vmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_1x_2} & f_{x_2x_2} \end{vmatrix} = \begin{vmatrix} 2 & 2 \\ 2 & 12x_2^2 \end{vmatrix} = 24x_2^2 - 4 \tag{1.47}$$

- *For* $x_2 = 0$*,* $D_2 = -4 < 0$ *and* $(0,0)$ *is a saddle point*
- *For* $x_2 = -1/\sqrt{2}$*,* $D_2 = 24/2 - 4 = 8 > 0$ *and* $(1/\sqrt{2}, -1/\sqrt{2})$ *is a minimum*
- *For* $x_2 = 1/\sqrt{2}$*,* $D_2 = 8 > 0$ *and* $(-1/\sqrt{2}, 1/\sqrt{2})$ *is also a minimum*

A plot of the function and the contour lines are shown in Figure 1.10.

1.2.2.3 Sufficient conditions for n independent variables

Second derivative test

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of n independent variables, and $\mathbf{x_0} = [x_{10}\; x_{20}\; \dots x_{n0}]^T$ a stationary point. The Taylor series expansion about $\mathbf{x_0}$ is:

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x_0}) + \nabla f(\mathbf{x_0})^T(\mathbf{x} - \mathbf{x_0}) + \\ &+ \frac{1}{2}(\mathbf{x} - \mathbf{x_0})^T \mathbf{H_2}(\mathbf{x} - \mathbf{x_0}) + \textit{higher order terms} \end{aligned} \tag{1.48}$$

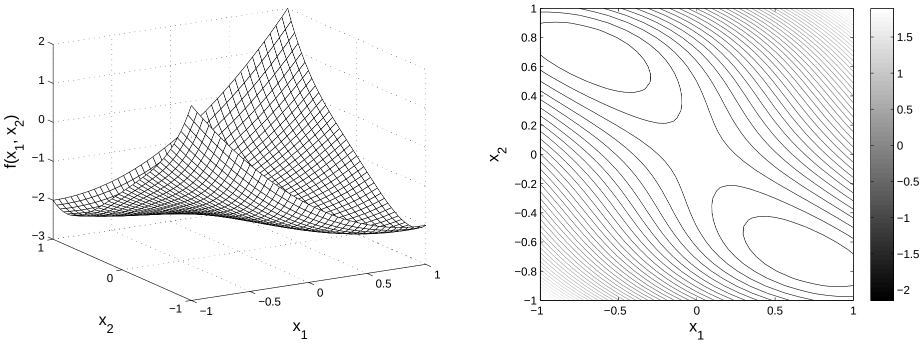

Figure 1.10: Mesh and contour plot of $f(x_1, x_2) = x_1^2 + 2x_1x_2 + x_2^4 - 2$

where $\mathbf{H_2}$ is the Hessian matrix defined by:

$$\mathbf{H_2} = \begin{bmatrix} f_{x_1x_1} & f_{x_1x_2} & \cdots & f_{x_1x_n} \\ f_{x_2x_1} & f_{x_2x_2} & \cdots & f_{x_2x_n} \\ \cdots & \cdots & \cdots & \cdots \\ f_{x_nx_1} & f_{x_nx_2} & \cdots & f_{x_nx_2} \end{bmatrix} \tag{1.49}$$

If $\mathbf{x}$ is sufficiently close to $\mathbf{x_0}$, the terms containing $(x_i - x_{i0})^k,\ k > 2$ become very small and higher order terms can be neglected. The first derivatives of f are zero at a stationary point, thus the relation (1.48) can be written as:

$$f(\mathbf{x}) = f(\mathbf{x_0}) + \frac{1}{2}(\mathbf{x} - \mathbf{x_0})^T\mathbf{H_2}(\mathbf{x} - \mathbf{x_0}) \tag{1.50}$$

The sign of the quadratic form which occurs in (1.50) as the second term in the right hand side will decide the character of the stationary point $\mathbf{x_0}$, in an analog way as for single variable functions.

According to (Hancock, 1960), we can determine weather the quadratic form is positive or negative by evaluating the signs of the determinants of the upper-left sub-matrices of $\mathbf{H_2}$:

$$\mathbf{D_i} = \begin{vmatrix} f_{x_1x_1} & f_{x_1x_2} & \cdots & f_{x_1x_i} \\ f_{x_2x_1} & f_{x_2x_2} & \cdots & f_{x_2x_i} \\ \cdots & \cdots & \cdots & \cdots \\ f_{x_ix_1} & f_{x_ix_2} & \cdots & f_{x_ix_2} \end{vmatrix}, \quad i = \overline{1, n} \tag{1.51}$$

- If $D_i(\mathbf{x_0}) > 0, \quad i = \overline{1, n}$, or $\mathbf{H_2}$ is positive definite, the quadratic form is positive and $\mathbf{x_0}$ is a local minimizer of f.
- If $D_i(\mathbf{x_0}) > 0, \ \ i = 2, 4, ...$ and $D_i(\mathbf{x_0}) < 0, \ \ i = 1, 3, ...$, or $\mathbf{H_2}$ is negative definite, the quadratic form is negative and $\mathbf{x_0}$ is a local maximizer of f.
- If $\mathbf{H_2}$ has both positive and negative eigenvalues, $\mathbf{x_0}$ is a saddle point
- otherwise, the test is inconclusive.

Note that if the hessian matrix is positive semidefinite or negative semidefinite, the test is also inconclusive.

The sufficient conditions presented above are the same as the ones stated for two independent variables, when $n = 2$. For example, a stationary point (x_{10}, x_{20}) is a local maximizer of a function f when $f_{x_1x_1}(x_{10}, x_{20}) = D_1(x_{10}, x_{20}) < 0$ and $D_2(x_{10}, x_{20}) > 0$.

Example 1.8 *Compute the stationary points of the function:*

$$f(x_1, x_2, x_3) = x_1^3 - 3x_1 + x_2^2 + x_3^2 \tag{1.52}$$

and classify them.

The first derivatives are:

$$f_{x_1} = 3x_1^2 - 3, \quad f_{x_2} = 2x_2, \quad f_{x_3} = 2x_3 \tag{1.53}$$

If they are set to zero we obtain two stationary points:

$$\begin{aligned} x_{10} = 1, \quad x_{20} = 0, \quad x_{30} = 0 \\ x_{10} = -1, \quad x_{20} = 0, \quad x_{30} = 0 \end{aligned} \tag{1.54}$$

The second-order derivatives:

$$f_{x_1x_1} = 6x_1, \quad f_{x_2x_2} = 2, \quad f_{x_3x_3} = 2, \quad f_{x_1x_2} = f_{x_1x_3} = f_{x_2x_3} = 0 \tag{1.55}$$

will form the Hessian matrix:

$$\mathbf{H_2} = \begin{bmatrix} 6x_1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \tag{1.56}$$

For $(1, 0, 0)$ *the determinants of the upper-left submatrices:*

$$\begin{aligned} D_1 &= 6 > 0, \quad D_2 = \begin{vmatrix} 6 & 0 \\ 0 & 2 \end{vmatrix} = 12 > 0, \\ D_3 &= \begin{vmatrix} 6 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{vmatrix} = 24 > 0 \end{aligned} \tag{1.57}$$

are all positive, so the stationary point is a minimizer of $f(x_1, x_2, x_3)$*.*

For $(-1, 0, 0)$ *the determinants* $D_1,\ D_2,\ D_3$ *are all negative, as resulted from:*

$$\begin{aligned} D_1 &= -6 < 0, \quad D_2 = \begin{vmatrix} -6 & 0 \\ 0 & 2 \end{vmatrix} = -12 < 0, \\ D_3 &= \begin{vmatrix} -6 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{vmatrix} = -24 < 0 \end{aligned} \tag{1.58}$$

The Hessian matrix is diagonal, the eigenvalues are easily determined as $-6,\ 2,\ 2$*. Because they do not have the same sign and are nonzero* $(-1, 0, 0)$ *is a saddle point.*

1.3 Constrained optimization

1.3.1 Problem formulation

A typical constrained optimization problem can be formulated as:

$$\min_{\mathbf{x}} f(\mathbf{x}) \tag{1.59}$$

subject to

$$g_i(\mathbf{x}) = 0, \quad i = \overline{1, m} \tag{1.60}$$

$$h_j(\mathbf{x}) \leq 0, \quad j = \overline{1, p} \tag{1.61}$$

where $\mathbf{x}$ is a vector of n independent variables, $[x_1 \ x_2 \ \ldots x_n]^T$, and f, g_i and h_j are scalar multivariate functions.

In practical problems, the values of independent variables are often limited since they usually represent physical quantities. The constraints on the variables can be expressed in the form of equations (1.60) or inequalities (1.61). A solution of a constrained optimization problem will minimize the objective function while satisfying the constraints.

In a well formulated problem, the number of equality constraints must be less or equal to the number of variables, $m \leq n$. The constraints (1.60) form a system of m (in general) nonlinear equations with n variables. When $m \geq n$, the problem is overdetermined and the solution may not exist because there are more equations than variables. If $m = n$ the values of the variables are uniquely determined and there is no problem of optimization.

At a feasible point $\mathbf{x}$, an inequality constraint is said to be *active* (or *tight*) if $h_j(\mathbf{x}) = 0$ and *inactive* (or *loose*) if the strict inequality $h_j(\mathbf{x}) < 0$ is satisfied.

Example 1.9 *Minimize*

$$f(x_1, x_2) = x_1^2 + x_2^2 \tag{1.62}$$

subject to

$$g(x_1, x_2) = x_1 + 2x_2 + 4 = 0 \tag{1.63}$$

The plot of the function and the constraint are illustrated in Figure 1.11, as a 3D surface and contour lines.

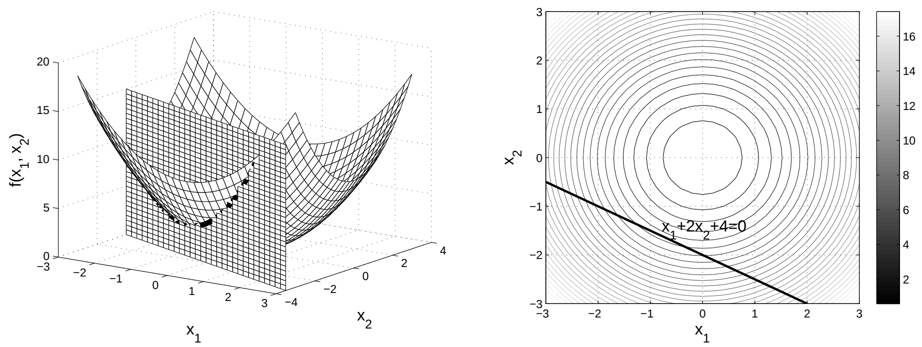

Figure 1.11: Mesh and contour plot of $f(x_1, x_2) = x_1^2 + x_2^2$ and the constraint $x_1 + 2x_2 + 4 = 0$

The purpose now is to minimize the function $x_1^2 + x_2^2$ subject to the condition that the variables x_1 and x_2 lie on the line $x_1 + 2x_2 + 4 = 0$. Graphically, the minimizer must be located on the curve obtained as the intersection of the function surface and the vertical plane that passes through the constraint line. In the $x_1 - x_2$ plane, the minimizer can be found as the point where the line $x_1 + 2x_2 + 4 = 0$ and a level curve of $x_1^2 + x_2^2$ are tangent.

1.3.2 Handling inequality constraints

Optimization problems with inequality constraints can be converted into equality-constrained problems and solved using the same approach. In this course, the method of slack variables is presented.

The functions $h_j(\mathbf{x})$ from the inequalities (1.61) can be set equal to zero if a positive quantity is added:

$$h_j(\mathbf{x}) + s_j^2 = 0, \quad j = \overline{1, p} \tag{1.64}$$

Notice that the *slack variables* s_j are squared to be positive. They are additional variables, so the number of the unknown variables increases to $n + p$.

An inequality given in the form:

$$h_j(\mathbf{x}) \geq 0, \quad j = \overline{1, p} \tag{1.65}$$

can also be turned into an equality if a positive quantity is subtracted from the left hand side:

$$h_j(\mathbf{x}) - s_j^2 = 0, \quad j = \overline{1, p} \tag{1.66}$$

Example 1.10 *Maximize*

$$f(x_1, x_2) = -x_1^2 - x_2^2 \tag{1.67}$$

subject to

$$x_1 + x_2 \leq 4 \tag{1.68}$$

The inequality (1.68) is converted into an equality by introducing a slack variable s_1^2*:*

$$x_1 + x_2 - 4 + s_1^2 = 0 \tag{1.69}$$

The unknown variables are now: x_1*,* x_2 *and* s_1*.*

1.3.3 Analytical methods for optimization problems with equality constraints. Solution by substitution

This method can be applied when it is possible to solve the constraint equations for m independent variables, where the number of constraints is less than the total number of variables $m \leq n$. The solution of the constraint equations is then substituted into the objective function. The new problem will have $n - m$ unknowns, it will be unconstrained and the techniques for unconstrained optimization can be applied.

Example 1.11 *Let* w*,* h*,* d *be the width, height and depth of a box (a rectangular parallelepiped). Find the optimal shape of the box to maximize the volume, when the sum* $w + h + d$ *is 120.*

The problem can be formulated as:

$$\max_{(w,h,d)} whd \tag{1.70}$$

subject to

$$w + h + d - 120 = 0 \tag{1.71}$$

We shall solve (1.71) for one of the variables, for example d and then substitute the result into (1.70):

$$d = 120 - w - h \tag{1.72}$$

The new problem is:

$$\max_{(w,h)} wh(120 - w - h) \tag{1.73}$$

Let:

$$f(w, h) = wh(120 - w - h) = 120wh - w^2h - wh^2 \tag{1.74}$$

Compute the stationary points from:

$$\frac{\partial f}{\partial w} = 120h - 2wh - h^2 = h(120 - 2w - h) = 0 \tag{1.75}$$

$$\frac{\partial f}{\partial h} = 120w - w^2 - 2wh = w(120 - w - 2h) = 0 \tag{1.76}$$

The solutions are: $w = h = 0$ *(not convenient) and* $w = h = 40$*.*

Determine whether $(40, 40)$ *is a minimum or maximum point. Write the determinant:*

$$\begin{aligned} D_2 &= \begin{vmatrix} \frac{\partial^2 f}{\partial w^2} & \frac{\partial^2 f}{\partial w \partial h} \\ \frac{\partial^2 f}{\partial w \partial h} & \frac{\partial^2 f}{\partial h^2} \end{vmatrix} = \begin{vmatrix} -2h & 120 - 2w - 2h \\ 120 - 2w - 2h & -2w \end{vmatrix} \\ &= \begin{vmatrix} -80 & -40 \\ -40 & -80 \end{vmatrix} \end{aligned} \tag{1.77}$$

Since

$$D_1 = \frac{\partial^2 f}{\partial w^2} = -80 < 0 \text{ and } D_2 = 4800 > 0$$

the point $(40, 40)$ *is a maximum. From (1.72) we have:* $d = 120 - 40 - 40 = 40$*, thus the box should have the sides equal:* $w = d = h = 40$*.*

Example 1.12 *Maximize*

$$f(x_1, x_2) = -x_1^2 - x_2^2 \tag{1.78}$$

subject to:

$$x_1 + x_2 = 4 \tag{1.79}$$

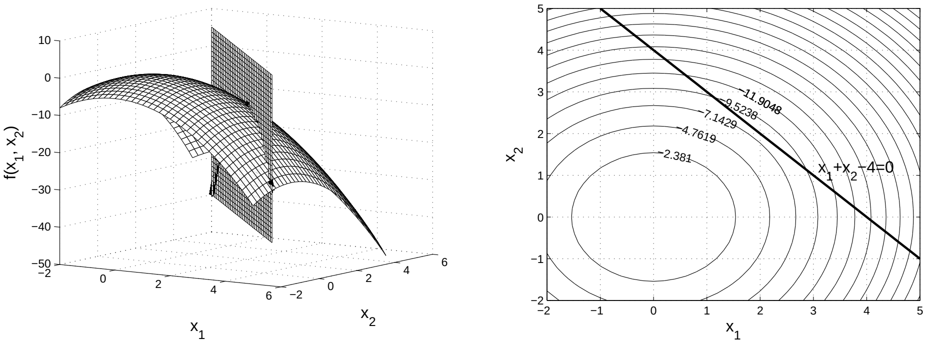

Figure 1.12: Mesh and contour plot of $f(x_1, x_2) = -x_1^2 - x_2^2$ and the constraint $x_1 + x_2 - 4 = 0$

As shown in Figure 1.12, the constrained maximum of $f(x_1, x_2)$ must be located on the curve resulted as the intersection of the function surface and the vertical plane that passes through the constraint line. In the plot showing the level curves of $f(x_1, x_2)$, the point that maximizes the function is located where the line $x_1+x_2+4 = 0$ is tangent to one level curve.

Analytically, this point can be determined by the method of substitution, as follows:

Solve (1.79) for x_2:

$$x_2 = 4 - x_1 \tag{1.80}$$

and replace it into the objective function. The new unconstrained problem is:

$$\max_{x_1} \left(-x_1^2 - (4 - x_1)^2\right) \tag{1.81}$$

The stationary point of $f(x_1) = -x_1^2 - (4 - x_1)^2$ is calculated by letting the first derivative be zero:

$$-2x_1 - 2(4 - x_1)(-1) = 0 \tag{1.82}$$

which, gives $x_{10} = 2$.

The second derivative $f''(x_{10}) = -4$ is negative, thus $x_{10} = 2$ is a maximizer of $f(x_1)$. From (1.80), $x_{20} = 4 - x_{10} = 2$, so the solution of the constrained optimization problem is $(2, 2)$.

Example 1.13 *Minimize*

$$f(x_1, x_2) = 20 - x_1 x_2 \tag{1.83}$$

subject to:

$$x_1 + x_2 = 6 \tag{1.84}$$

Substitute $x_2 = 6 - x_1$ from (1.84) into (1.83) and obtain the unconstrained problem:

$$\max_{x_1} 20 - x_1(6 - x_1) \tag{1.85}$$

The stationary point is computed from the first derivative of $f(x_1)$:

$$f'(x_1) = -6 + 2x_1 = 0, \quad x_{10} = 3 \tag{1.86}$$

Because the second derivative $f''(x_{10}) = 2$ is positive, the stationary point is a minimizer of $f(x_1)$. Because $x_2 = 6 - x_1$, the point ($x_{10} = 3$, $x_{20} = 3$) minimizes the function $f(x_1, x_2)$ subject to the constraint (1.84). As shown in Figure 1.13, the minimum obtained is located on the parabola resulted as the intersection of the function surface and the vertical plane containing the constraint line, or, in the contour plot, it is the point where the constraint line is tangent to a level curve.

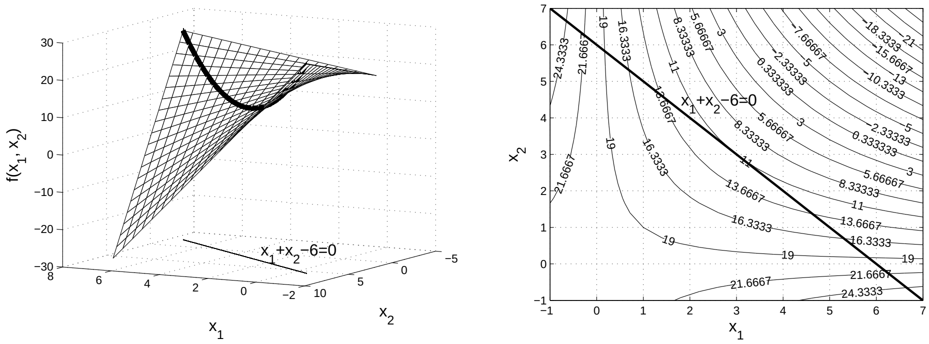

Figure 1.13: Mesh and contour plot of $f(x_1, x_2) = 20 - x_1 x_2$ and the constraint $x_1 + x_2 - 6 = 0$

1.3.4 Lagrange multipliers

In case the direct substitution cannot be applied, the method of Lagrange multipliers provides a strategy for finding the minimum or maximum value of a function subject to equality constraints. The general problem can be formulated as:

$$\min_{\mathbf{x}} f(\mathbf{x}) \tag{1.87}$$

subject to

$$g_i(\mathbf{x}) = 0, \quad i = \overline{1, m} \tag{1.88}$$

As example, consider the problem of finding the minimum of a real-valued function $f(x_1, x_2)$ subject to the constraint $g(x_1, x_2) = 0$. Let $f(x_1, x_2) = 20 - x_1 x_2$ and $g(x_1, x_2) = x_1 + x_2 - 6 = 0$, as shown in Figure 1.14. The gradient direction of $f(x_1, x_2)$ is also shown, as arrows, in the same figure.

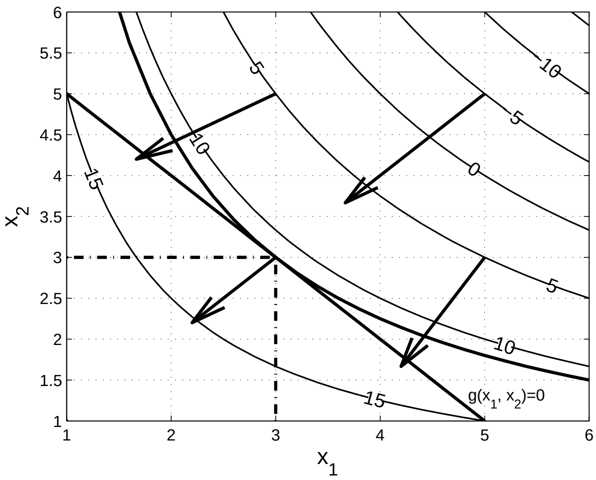

Figure 1.14: Contour plot of $f(x_1, x_2) = 20 - x_1 x_2$ and the constraint $g(x_1, x_2) = x_1 + x_2 - 6 = 0$

The aim of this example is to minimimze $f(x_1, x_2)$, or to find the point which lies on the level curve with the smallest possible value and which satisfies $g(x_1, x_2) = 0$, i.e. it must lie on the constraint line too. If we are at the point $(x_1 = 3, x_2 = 3)$, the value of the function is $f(3, 3) = 11$ and the constraint is satisfied. The constraint line is tangent to the level curve at this point. If we move on the line $g(x_1, x_2) = 0$ left or right from this point, the value of the function increases. Thus, the solution of the problem is $(3, 3)$.

In the general case, consider the point $\mathbf{x}$ where a level curve of a function $f(\mathbf{x})$ is tangent to the constraint $g(\mathbf{x})$. At this point, the gradient of f, $\nabla f(\mathbf{x})$, is parallel to the gradient of the constraint, $\nabla g(\mathbf{x})$. For two vectors to be parallel, they must be multiples of each other. Thus, there is a scalar value λ, so that:

$$\nabla f(\mathbf{x}) = -\lambda \nabla g(\mathbf{x}) \tag{1.89}$$

The equation (1.89), written as:

$$\nabla f(\mathbf{x}) + \lambda \nabla g(\mathbf{x}) = 0 \tag{1.90}$$

will provide a necessary condition for optimization of f subject to the constraint $g = 0$.

The method of *Lagrange multipliers* in case of n independent variables and m constraints, as defined by (1.87) and (1.88) can be generalized (Avriel, 2003), (Hiriart-Urruty, 1996).

Define the Lagrangian (or augmented) function:

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda^T \mathbf{g}(\mathbf{x}) \tag{1.91}$$

where $\mathbf{g}$ is a column vector function containing the m constraints $g_i(\mathbf{x})$, and λ is a column vector m of unknown values, called *Lagrange multipliers*. The function above can be written in an expanded form as:

$$\begin{aligned} L(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m) &= f(x_1, x_2, \ldots, x_n) + \lambda_1 g_1(x_1, x_2, \ldots, x_n) \\ &+ \ldots + \lambda_m g_m(x_1, x_2, \ldots, x_n) \end{aligned} \tag{1.92}$$

To locate the stationary points, the gradient of the Lagrangian function is set equal to zero:

$$\nabla L(\mathbf{x}, \lambda) = \nabla \left(f(\mathbf{x}) + \lambda^T g(\mathbf{x}) \right) = 0 \tag{1.93}$$

The necessary conditions for optimum are obtained by setting the first partial derivatives of the Lagrangian function with respect to x_i, $i = \overline{1,n}$ and λ_j, $j = \overline{1,m}$ equal to zero. There are $n + m$ nonlinear algebraic equations to be solved for $n + m$ unknowns, as follows:

$$\begin{aligned} \frac{\partial L(\mathbf{x}, \lambda)}{\partial x_1} &= \frac{\partial f(\mathbf{x})}{\partial x_1} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(\mathbf{x})}{\partial x_1} = 0 \\ \frac{\partial L(\mathbf{x}, \lambda)}{\partial x_2} &= \frac{\partial f(\mathbf{x})}{\partial x_2} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(\mathbf{x})}{\partial x_2} = 0 \\ &\ldots \\ \frac{\partial L(\mathbf{x}, \lambda)}{\partial x_n} &= \frac{\partial f(\mathbf{x})}{\partial x_n} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(\mathbf{x})}{\partial x_n} = 0 \\ \frac{\partial L(\mathbf{x}, \lambda)}{\partial \lambda_1} &= g_1(\mathbf{x}) = 0 \\ &\ldots \\ \frac{\partial L(\mathbf{x}, \lambda)}{\partial \lambda_m} &= g_m(\mathbf{x}) = 0 \end{aligned} \tag{1.94}$$

Example 1.14 *Find the stationary points of the function $f(x_1, x_2) = 20 - x_1 x_2$ subject to the constraint $x_1 + x_2 = 6$ using the method of Lagrange multipliers.*

Define the Lagrangian function:

$$L(x_1, x_2, \lambda) = 20 - x_1 x_2 + \lambda(x_1 + x_2 - 6) \tag{1.95}$$

Obtain the stationary points from:

$$\frac{\partial L(x_1, x_2, \lambda)}{\partial x_1} = -x_2 + \lambda = 0 \tag{1.96}$$

$$\frac{\partial L(x_1, x_2, \lambda)}{\partial x_2} = -x_1 + \lambda = 0 \tag{1.97}$$

$$\frac{\partial L(x_1, x_2, \lambda)}{\partial \lambda} = x_1 + x_2 - 6 = 0 \tag{1.98}$$

From (1.96) and (1.97), $x_1 = x_2 = \lambda$, which replaced in (1.98) gives the stationary point: $x_1 = 3 = x_2 = \lambda$.

This point is the solution of the constrained minimization problem from example 1.13, but the sufficient conditions for constrained optimization will be discussed in the following section.

1.3.5 Sufficient conditions for constrained optimization

The sufficient conditions for minima or maxima in constrained optimization problems are completely described and demonstrated in (Avriel, 2003) and (Hiriart-Urruty, 1996). The results in the case of equality constraints can be summarized by the following:

Corollary (Avriel, 2003): Let f, g_1, g_2, ..., g_m be twice continuously differentiable real-valued functions. If there exist vectors $\mathbf{x}_0 \in \mathbb{R}^n$, $\lambda_0 \in \mathbb{R}^m$, such that:

$$\nabla L(\mathbf{x_0}, \lambda_\mathbf{0}) = 0 \tag{1.99}$$

and if

$$(-1)^m \begin{vmatrix} \frac{\partial^2 L(\mathbf{x_0},\lambda_\mathbf{0})}{\partial x_1 \partial x_1} & \dots & \frac{\partial^2 L(\mathbf{x_0},\lambda_\mathbf{0})}{\partial x_1 \partial x_p} & \frac{\partial g_1(\mathbf{x_0})}{\partial x_1} & \dots & \frac{\partial g_m(\mathbf{x_0})}{\partial x_1} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial^2 L(\mathbf{x_0},\lambda_\mathbf{0})}{\partial x_p \partial x_1} & \dots & \frac{\partial^2 L(\mathbf{x_0},\lambda_\mathbf{0})}{\partial x_p \partial x_p} & \frac{\partial g_1(\mathbf{x_0})}{\partial x_p} & \dots & \frac{\partial g_m(\mathbf{x_0})}{\partial x_p} \\ \frac{\partial g_1(\mathbf{x_0})}{\partial x_1} & \dots & \frac{\partial g_1(\mathbf{x_0})}{\partial x_p} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial g_m(\mathbf{x_0})}{\partial x_1} & \dots & \frac{\partial g_m(\mathbf{x_0})}{\partial x_p} & 0 & \dots & 0 \end{vmatrix} > 0 \tag{1.100}$$

for $p = m+1, \dots, n$, then f has a strict local minimum at $\mathbf{x_0}$, such that

$$g_i(\mathbf{x_0}) = 0, \quad i = \overline{1, m} \tag{1.101}$$

The similar result for strict local maxima is obtained by changing $(-1)^m$ in (1.100) to $(-1)^p$, (Avriel, 2003).

For $p = n$, the matrix from (1.100) is the *bordered Hessian matrix* of the problem. The elements are in fact the second derivatives of the Lagrangian with respect to all its $n + m$ variables, x_i and λ_j. The columns in the right and the last rows can be easier recognized as second derivatives of L if we notice that:

$$\frac{\partial L(\mathbf{x}, \lambda)}{\partial \lambda_j} = g_j(\mathbf{x}) \text{ and } \frac{\partial L^2(\mathbf{x}, \lambda)}{\partial \lambda_j \partial x_i} = \frac{\partial g_j(\mathbf{x})}{\partial x_i} \tag{1.102}$$

Because $g_j(\mathbf{x})$ does not depend on λ, the zeros from lower-right corner or the matrix result from:

$$\frac{\partial L(\mathbf{x}, \lambda)}{\partial \lambda_j} = g_j(\mathbf{x}) \text{ and } \frac{\partial L^2(\mathbf{x}, \lambda)}{\partial \lambda_j \partial \lambda_i} = 0 \tag{1.103}$$

When $p < n$, the matrices can be obtained if the rows and columns between $p+1$ and $n-1$ are excluded.

Example 1.15 *For the problem from example 1.14, we shall prove that the stationary point is a minimum, according to the sufficient condition defined above. The function to be minimized is $f(x_1, x_2) = 20 - x_1x_2$ and the constraint $g(x_1, x_2) = x_1 + x_2 - 6 = 0$*

The number of variables in this case is $n = 2$, the number of constraints, $m = 1$ and $p = m + 1 = 2$. The only matrix we shall analyze is:

$$\mathbf{H_2} = \begin{bmatrix} \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial^2 x_1} & \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial x_1 \partial x_2} & \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial x_1 \partial \lambda} \\ \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial x_2 \partial x_1} & \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial^2 x_2} & \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial x_2 \partial \lambda} \\ \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial \lambda \partial x_1} & \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial \lambda \partial x_2} & \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial^2 \lambda} \end{bmatrix} \tag{1.104}$$

or

$$\mathbf{H_2} = \begin{bmatrix} \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial^2 x_1} & \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial x_1 \partial x_2} & \frac{\partial g(x_1,x_2)}{\partial x_1} \\ \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial x_2 \partial x_1} & \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial^2 x_2} & \frac{\partial g(x_1,x_2)}{\partial x_2} \\ \frac{\partial g(x_1,x_2)}{\partial x_1} & \frac{\partial g(x_1,x_2)}{\partial x_2} & 0 \end{bmatrix} \tag{1.105}$$

Using the results we have obtained in the example 1.14, the second derivatives of the Lagrangian function are:

$$\begin{aligned} \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial^2 x_1} &= \frac{\partial}{\partial x_1}(-x_2+\lambda) = 0 \\ \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial^2 x_2} &= \frac{\partial}{\partial x_2}(-x_1+\lambda) = 0 \\ \frac{\partial^2 L(x_1,x_2,\lambda)}{\partial x_1 \partial x_2} &= \frac{\partial}{\partial x_2}(-x_2+\lambda) = -1 \\ \frac{\partial g(x_1,x_2)}{\partial x_1} &= \frac{\partial}{\partial x_1}(x_1+x_2-6) = 1 \\ \frac{\partial g(x_1,x_2)}{\partial x_2} &= \frac{\partial}{\partial x_2}(x_1+x_2-6) = 1 \end{aligned} \tag{1.106}$$

and the bordered Hessian:

$$\mathbf{H_2} = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \tag{1.107}$$

The relation (1.100) is written as:

$$(-1)^1 \begin{vmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = (-1)(-2) = 2 > 0 \tag{1.108}$$

thus, the stationary point $(3,3)$ is a minimizer of the function f subject to the constraint $g = 0$.

Example 1.16 *Find the highest and lowest points on the surface $f(x_1,x_2) = x_1x_2 + 25$ over the circle $x_1^2 + x_2^2 = 18$.*

Figure 1.15 shows a graphical representation of the problem. The constraint on x_1 and x_2 places the variables on the circle with the center in the origin and with a radius equal to $\sqrt{18}$. On this circle, the values of the function $f(x_1,x_2)$ are located on the curve shown in Figure 1.15 on the mesh plot. It is clear from the picture that there are two maxima and two minima that will be determined using the method of Lagrange multipliers.

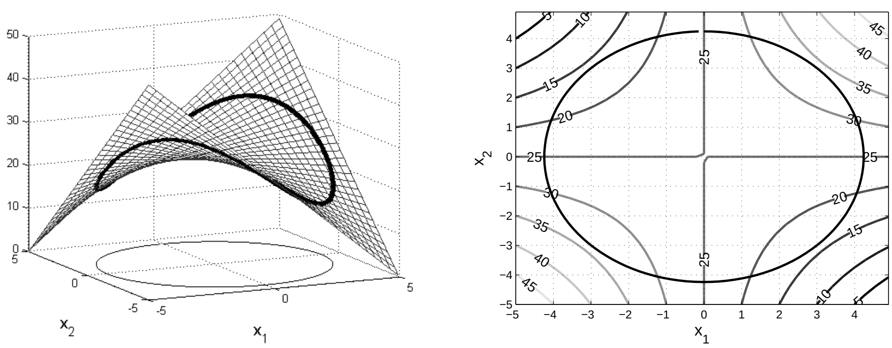

Figure 1.15: Mesh and contour plot of $f(x_1,x_2) = 25 + x_1x_2$ and the constraint $x_1^2 + x_2^2 = 18$

The problem may be reformulated as: optimize $f(x_1,x_2) = x_1x_2 + 25$ subject to the constraint $g(x_1,x_2) = x_1^2 + x_2^2 - 18 = 0$.

Write the Lagrange function:

$$L(x_1, x_2, \lambda) = x_1x_2 + 25 + \lambda(x_1^2 + x_2^2 - 18) \tag{1.109}$$

Compute the first partial derivatives of L and set them equal to zero:

$$\begin{aligned} L_{x_1} &= x_2 + 2\lambda x_1 = 0 \\ L_{x_2} &= x_1 + 2\lambda x_2 = 0 \\ L_\lambda &= x_1^2 + x_2^2 - 18 = 0 \end{aligned} \tag{1.110}$$

There are four stationary points:

$$\begin{aligned} x_{10} &= -3,\ \ x_{20} = 3,\ \ \lambda_0 = \frac{1}{2} \\ x_{10} &= 3,\ \ x_{20} = -3,\ \ \lambda_0 = \frac{1}{2} \\ x_{10} &= 3,\ \ x_{20} = 3,\ \ \lambda_0 = -\frac{1}{2} \\ x_{10} &= -3,\ \ x_{20} = -3,\ \ \lambda_0 = -\frac{1}{2} \end{aligned} \tag{1.111}$$

Build the bordered Hessian matrix and check the sufficient conditions for maxima and minima.

$$\mathbf{H_2} = \begin{bmatrix} L_{x_1x_1} & L_{x_1x_2} & g_{x_1} \\ L_{x_1x_2} & L_{x_2x_2} & g_{x_2} \\ g_{x_1} & g_{x_2} & 0 \end{bmatrix} = \begin{bmatrix} 2\lambda & 1 & 2x_1 \\ 1 & 2\lambda & 2x_2 \\ 2x_1 & 2x_2 & 0 \end{bmatrix} \tag{1.112}$$

Because the number of constraints is $m = 1$, the number of variables is $n = 2$ and $p = 2$, the sufficient condition for a stationary point to be a minimizer of f subject to $g = 0$ is:

$$(-1)^1 det(\mathbf{H_2}) > 0 \ \text{ or } \ det(\mathbf{H_2}) < 0 \tag{1.113}$$

and for a maximizer:

$$(-1)^2 det(\mathbf{H_2}) > 0 \ \text{ or } \ det(\mathbf{H_2}) > 0 \tag{1.114}$$

For all points $(x_{10}, x_{20}, \lambda_0)$ given in (1.111) compute $det(\mathbf{H_2})$ and obtain:

$$(-3, 3, \frac{1}{2}): \quad det(\mathbf{H_2}) = \begin{vmatrix} 1 & 1 & -6 \\ 1 & 1 & 6 \\ -6 & 6 & 0 \end{vmatrix} = -144 < 0 \tag{1.115}$$

$$(3, -3, \frac{1}{2}): \quad det(\mathbf{H_2}) = \begin{vmatrix} 1 & 1 & 6 \\ 1 & 1 & -6 \\ 6 & -6 & 0 \end{vmatrix} = -144 < 0 \tag{1.116}$$

$$(3, 3, -\frac{1}{2}): \quad det(\mathbf{H_2}) = \begin{vmatrix} -1 & 1 & 6 \\ 1 & -1 & 6 \\ 6 & 6 & 0 \end{vmatrix} = 144 > 0 \tag{1.117}$$

$$(-3, -3, -\frac{1}{2}): \quad det(\mathbf{H_2}) = \begin{vmatrix} -1 & 1 & -6 \\ 1 & -1 & -6 \\ -6 & -6 & 0 \end{vmatrix} = 144 > 0 \tag{1.118}$$

Thus, the function f subject to $g = 0$ has two minima at $(3, -3, \frac{1}{2})$ and $(-3, 3, \frac{1}{2})$, and two maxima at $(3, 3, -\frac{1}{2})$ and $(-3, -3, -\frac{1}{2})$.

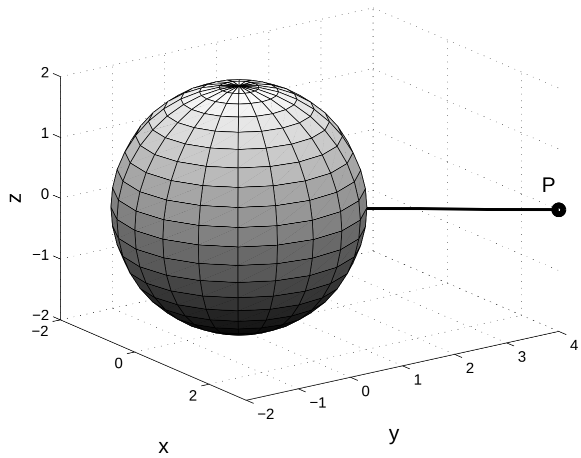

Figure 1.16: The sphere $x^2 + y^2 + z^2 = 4$ and the point P

Example 1.17 *Find the point on the sphere $x^2 + y^2 + z^2 = 4$ closest to the point $P(3, 4, 0)$ (Figure 1.16).*

We shall find the point (x, y, z) that minimizes the distance between P and the sphere. In a 3-dimensional space, the distance between any point (x, y, z) and $P(3, 4, 0)$ is given by:

$$D(x, y, z) = \sqrt{(x-3)^2 + (y-4)^2 + z^2} \tag{1.119}$$

Since the point (x, y, z) must be located on the sphere, the variables are constrained by the equation $x^2 + y^2 + z^2 = 4$. The calculus will be easier if instead of (1.119) we minimize the function under the square root. The problem to be solved is then:

$$\min_{x,y,z} (x-3)^2 + (y-4)^2 + z^2 \tag{1.120}$$

subject to:

$$g(x, y, z) = x^2 + y^2 + z^2 - 4 = 0 \tag{1.121}$$

Write the Lagrangian function first:

$$L(x, y, z, \lambda) = (x-3)^2 + (y-4)^2 + z^2 + \lambda(x^2 + y^2 + z^2 - 4) \tag{1.122}$$

and set the partial derivatives equal to zero to compute the stationary points:

$$\begin{aligned} L_x &= 2x - 6 + 2\lambda x = 0 \\ L_y &= 2y - 8 + 2\lambda y = 0 \\ L_z &= 2z + 2\lambda z = 0 \\ L_\lambda &= x^2 + y^2 + z^2 - 4 = 0 \end{aligned} \tag{1.123}$$

The system (1.123) has two solutions:

$$(S_1): \ x_{10} = \frac{6}{5}, \ y_{10} = \frac{8}{5}, \ z_{10} = 0, \ \lambda_{10} = \frac{3}{2} \tag{1.124}$$

and

$$(S_2): \ x_{10} = -\frac{6}{5}, \ y_{10} = -\frac{8}{5}, \ z_{10} = 0, \ \lambda_{10} = -\frac{7}{2} \tag{1.125}$$

It is clear from Figure 1.16 that we must find a minimum and a maximum distance between the point P and the sphere, thus the sufficient conditions for maximum or minimum have to be checked.

The number of variables is $n = 3$, the number of constraints $m = 1$ and p from (1.100) has two values: $p = \overline{2, 3}$. We must analyze the sign of the determinants for the following matrices:

For $p = 2$:

$$\mathbf{H_{22}} = \begin{bmatrix} L_{xx} & L_{xy} & L_{x\lambda} \\ L_{xy} & L_{yy} & L_{y\lambda} \\ L_{x\lambda} & L_{y\lambda} & L_{\lambda\lambda} \end{bmatrix} = \begin{bmatrix} L_{xx} & L_{xy} & g_x \\ L_{xy} & L_{yy} & g_y \\ g_x & g_y & 0 \end{bmatrix} \tag{1.126}$$

For $p = 3$:

$$\mathbf{H_{23}} = \begin{bmatrix} L_{xx} & L_{xy} & L_{xz} & L_{x\lambda} \\ L_{xy} & L_{yy} & L_{yz} & L_{y\lambda} \\ L_{xz} & L_{yz} & L_{zz} & L_{z\lambda} \\ L_{x\lambda} & L_{y\lambda} & L_{z\lambda} & L_{\lambda\lambda} \end{bmatrix} = \begin{bmatrix} L_{xx} & L_{xy} & L_{xz} & g_x \\ L_{xy} & L_{yy} & L_{yz} & g_y \\ L_{xz} & L_{yz} & L_{zz} & g_z \\ g_x & g_y & g_z & 0 \end{bmatrix} \tag{1.127}$$

The sufficient conditions for minimum in this case are written as:

$$(-1)^1 det(\mathbf{H_{22}}) > 0, \text{ and } (-1)^1 det(\mathbf{H_{23}}) > 0 \tag{1.128}$$

and the sufficient conditions for maximum:

$$(-1)^2 det(\mathbf{H_{22}}) > 0, \text{ and } (-1)^3 det(\mathbf{H_{23}}) > 0 \tag{1.129}$$

The second derivatives of the Lagrangian function with respect to all its variables are:

$$\begin{aligned} L_{xx} &= 2 + 2\lambda, \quad L_{yy} = 2 + 2\lambda, \quad L_{zz} = 2 + 2\lambda \\ L_{xy} &= 0, \quad L_{xz} = 0, \quad L_{x\lambda} = g_x = 2x \\ L_{yz} &= 0, \quad L_{y\lambda} = g_y = 2y, \quad L_{z\lambda} = g_z = 2z \end{aligned} \tag{1.130}$$

From (1.126) and (1.127) we obtain:

$$\mathbf{H_{22}} = \begin{bmatrix} 2+2\lambda & 0 & 2x \\ 0 & 2+2\lambda & 2y \\ 2x & 2y & 0 \end{bmatrix} \tag{1.131}$$

$$\mathbf{H_{23}} = \begin{bmatrix} 2+2\lambda & 0 & 0 & 2x \\ 0 & 2+2\lambda & 0 & 2y \\ 0 & 0 & 2+2\lambda & 2z \\ 2x & 2y & 2z & 0 \end{bmatrix} \tag{1.132}$$

For the first stationary point, (S_1), the determinants of $\mathbf{H_{22}}$ and $\mathbf{H_{23}}$ are:

$$det\mathbf{H_{22}} = \begin{vmatrix} 5 & 0 & \frac{12}{5} \\ 0 & 5 & \frac{16}{5} \\ \frac{12}{5} & \frac{16}{5} & 0 \end{vmatrix} = -80 \tag{1.133}$$

$$det\mathbf{H_{23}} = \begin{vmatrix} 5 & 0 & 0 & \frac{12}{5} \\ 0 & 5 & 0 & \frac{16}{5} \\ 0 & 0 & 5 & 0 \\ \frac{12}{5} & \frac{16}{5} & 0 & 0 \end{vmatrix} = -400 \tag{1.134}$$

The point $(\frac{6}{5}, \frac{8}{5}, 0)$ is then a minimizer of the problem because:

$$(-1)^1 det(\mathbf{H_{22}}) = 80 > 0, \text{ and } (-1)^1 det(\mathbf{H_{23}}) = 400 > 0 \tag{1.135}$$

For the second stationary point, (S_2), the determinants are:

$$det\mathbf{H_{22}} = \begin{vmatrix} -5 & 0 & -\frac{12}{5} \\ 0 & -5 & -\frac{16}{5} \\ -\frac{12}{5} & -\frac{16}{5} & 0 \end{vmatrix} = 80, \tag{1.136}$$

$$det\mathbf{H_{23}} = \begin{vmatrix} -5 & 0 & 0 & -\frac{12}{5} \\ 0 & -5 & 0 & -\frac{16}{5} \\ 0 & 0 & -5 & 0 \\ -\frac{12}{5} & -\frac{16}{5} & 0 & 0 \end{vmatrix} = -400 \tag{1.137}$$

The point $(-\frac{6}{5}, -\frac{8}{5}, 0)$ is a maximizer of the problem because:

$$(-1)^2 det(\mathbf{H_{22}}) = 80 > 0, \text{ and } (-1)^3 det(\mathbf{H_{23}}) = 400 > 0 \tag{1.138}$$

The point on the sphere closest to P is: $(\frac{6}{5}, \frac{8}{5}, 0)$.

Example 1.18 *Find the extrema of the function:*

$$f(x, y, z) = x + 2y + z \tag{1.139}$$

subject to:

$$g_1(x, y, z) = x^2 + y^2 + z^2 - 1 = 0 \tag{1.140}$$
$$g_2(x, y, z) = x + y + z - 1 = 0 \tag{1.141}$$

In this case we have two constraints and 3 variables. Two new unknowns will be introduced and the Lagrange function is written as:

$$L(x, y, z, \lambda_1, \lambda_2) = x + 2y + z + \lambda_1(x^2 + y^2 + z^2 - 1) + \lambda_2(x + y + z - 1) \tag{1.142}$$

The stationary points are computed from:

$$\begin{aligned} L_x &= 1 + 2\lambda_1 x + \lambda_2 = 0 \\ L_y &= 2 + 2\lambda_1 y + \lambda_2 = 0 \\ L_z &= 1 + 2\lambda_1 z + \lambda_2 = 0 \\ L_{\lambda_1} &= x^2 + y^2 + z^2 - 1 = 0 \\ L_{\lambda_2} &= x + y + z - 1 = 0 \end{aligned} \tag{1.143}$$

The solutions of (1.143) are:

$$(S_1): \quad x_0 = 0, \ \ y_0 = 1, \ \ z_0 = 0, \ \ \lambda_{10} = -\tfrac{1}{2}, \ \ \lambda_{20} = -1 \tag{1.144}$$
$$(S_2): \quad x_0 = \tfrac{2}{3}, \ \ y_0 = -\tfrac{1}{3}, \ \ z_0 = \tfrac{2}{3}, \ \ \lambda_{10} = \tfrac{1}{2}, \ \ \lambda_{20} = -\tfrac{5}{3} \tag{1.145}$$

The number p from (1.100) is 3 in this case, therefore we have to analyze the sign of the determinant:

$$det\mathbf{H_2} = \begin{vmatrix} L_{xx} & L_{xy} & L_{xz} & g_{1x} & g_{2x} \\ L_{xy} & L_{yy} & L_{yz} & g_{1y} & g_{2y} \\ L_{xz} & L_{yz} & L_{zz} & g_{1z} & g_{2z} \\ g_{1x} & g_{1y} & g_{1z} & 0 & 0 \\ g_{2x} & g_{2y} & g_{2z} & 0 & 0 \end{vmatrix} = \begin{vmatrix} 2\lambda_1 & 0 & 0 & 2x & 1 \\ 0 & 2\lambda_1 & 0 & 2y & 1 \\ 0 & 0 & 2\lambda_1 & 2z & 1 \\ 2x & 2y & 2z & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{vmatrix} \tag{1.146}$$

The necessary condition for minimum (when $m = 2$ and $p = 3$) is:

$$(-1)^2 det\mathbf{H_2} > 0 \text{ or } det\mathbf{H_2} > 0 \tag{1.147}$$

and for maximum:

$$(-1)^3 det\mathbf{H_2} > 0 \text{ or } det\mathbf{H_2} < 0 \tag{1.148}$$

For (S_1)*:*

$$det\mathbf{H_2} = \begin{vmatrix} -1 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 2 & 1 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 2 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{vmatrix} = -8 < 0 \tag{1.149}$$

thus, the point ($x_0 = 0$*,* $y_0 = 1$*,* $z_0 = 0$*) is a maximizer of* f *subject to* $g_1 = 0$ *and* $g_2 = 0$*.*

For (S_2)*:*

$$det\mathbf{H_2} = \begin{vmatrix} 1 & 0 & 0 & \frac{4}{3} & 1 \\ 0 & 1 & 0 & -\frac{2}{3} & 1 \\ 0 & 0 & 1 & \frac{4}{3} & 1 \\ \frac{4}{3}0 & -\frac{2}{3} & \frac{4}{3} & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{vmatrix} = 8 > 0 \tag{1.150}$$

thus, the point ($x_0 = \frac{2}{3}$*,* $y_0 = -\frac{1}{3}$*,* $z_0 = \frac{2}{3}$*) minimizes* f *subject to* $g_1 = 0$ *and* $g_2 = 0$*.*

1.3.6 Karush-Kuhn-Tucker conditions

The *Karush-Kuhn-Tucker* (KKT) conditions are an extension of the Lagrangian theory to include nonlinear optimization problems with inequality constraints.

If $\mathbf{x}$ is a vector of n variables, $\mathbf{x} = [x_1 \ \ x_2 \ \ \ldots \ \ x_n]^T$ and f a nonlinear real-valued function, $f : \mathbb{R}^n \to \mathbb{R}$, consider the constrained minimization problem:

$$(P): \qquad \min_{\mathbf{x}} f(\mathbf{x}) \tag{1.151}$$

subject to

$$g_i(\mathbf{x}) = 0, \quad i = \overline{1,m} \tag{1.152}$$
$$h_j(\mathbf{x}) \leq 0, \quad j = \overline{1,p} \tag{1.153}$$

where f, g_i, h_j are twice differentiable real-valued functions.

The Lagrangian function is written as:

$$\begin{aligned} L(\mathbf{x}, \lambda, \mu) &= f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^{p} \mu_j h_j(\mathbf{x}) \\ &= f(\mathbf{x}) + \lambda^T \mathbf{g}(\mathbf{x}) + \mu^T \mathbf{h}(\mathbf{x}) \end{aligned} \tag{1.154}$$

where:

- $\lambda = [\lambda_1 \ \ \lambda_2 \ \ \ldots \ \ \lambda_m]^T$ and $\mu = [\mu_1 \ \ \mu_2 \ \ \ldots \ \ \mu_p]^T$ are vector multipliers,
- $\mathbf{g} = [g_1(\mathbf{x}) \ \ g_2(\mathbf{x}) \ \ldots \ g_m(\mathbf{x})]^T$ and $\mathbf{h} = [h_1(\mathbf{x}) \ \ h_2(\mathbf{x}) \ \ldots \ h_p(\mathbf{x})]^T$ are vector functions.

The necessary conditions for a point $\mathbf{x}_0$ to be a local minimizer of f are:

$$\nabla f(\mathbf{x}_0) + \textstyle\sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}_0) + \sum_{j=1}^{p} \mu_j \nabla h_j(\mathbf{x}_0) \tag{1.155}$$
$$g_i(\mathbf{x}_0) = 0, \quad i = \overline{1,m} \tag{1.156}$$
$$h_j(\mathbf{x}_0) \leq 0, \quad j = \overline{1,p} \tag{1.157}$$
$$\mu_j h_j(\mathbf{x}_0) = 0, \quad j = \overline{1,p} \tag{1.158}$$
$$\mu_j \geq 0, \quad j = \overline{1,p} \tag{1.159}$$
$$\lambda_i, \quad \text{unrestricted in sign}, \quad i = \overline{1,m} \tag{1.160}$$

The relations (1.155) - (1.160) are called the *Karush-Kuhn-Tucker* conditions, (Boyd and Vandenberghe, 2004). For any optimization problem with differentiable objective and constraint functions, any optimal points must satisfy the KKT conditions.

- The first condition (1.155) is a system of n nonlinear equations with $n + m + p$ unknowns, obtained by setting equal to zero all the partial derivatives of the Lagrangian with respect to $x_1,\ x_2,\ \ldots\ x_n$, i.e. $\nabla L(\mathbf{x}_0, \lambda, \mu) = 0$.
- The following two conditions (1.156, 1.157), are the inequality and equality constraints which must be satisfied by the minimizer of the constrained problem, and are called *the feasibility conditions*.
- The relation (1.158) is called *the complementary slackness equation*
- The relation (1.159) is the *non-negativity* condition for the μ multipliers.

The KKT conditions are *necessary conditions* for optimum, they are means to verify solutions. Not all the points that satisfy (1.155)-(1.160) are optimal points. On the other hand, a point is not optimal if the KKT conditions are not satisfied.

If the objective and inequality constraint functions (f and h_j) are convex and g_i are affine, the KKT conditions are also *sufficient* for a minimum point.

A function h_j is affine if it has the form:

$$h_j = A_1 x_1 + A_2 x_2 + \ldots A_n x_n + b \tag{1.161}$$

In a few cases, it is possible to solve the KKT conditions (and therefore, the optimization problem) analytically. but the sufficient conditions are difficult to verify.

Example 1.19 *Minimize*

$$f(x_1, x_2) = e^{-3x_1} + e^{-2x_2} \tag{1.162}$$

subject to:

$$x_1 + x_2 \leq 2 \tag{1.163}$$
$$x_1 \geq 0 \tag{1.164}$$
$$x_2 \geq 0 \tag{1.165}$$

The constraints are re-written in the standard form:

$$x_1 + x_2 - 2 \leq 0 \tag{1.166}$$
$$-x_1 \leq 0 \tag{1.167}$$
$$-x_2 \leq 0 \tag{1.168}$$

The Lagrangian of the problem is written as:

$$\begin{aligned} L(x_1, x_2, \mu_1, \mu_2, \mu_3) &= e^{-3x_1} + e^{-2x_2} + \mu_1(x_1 + x_2 - 2) + \\ &+ \mu_2(-x_1) + \mu_3(-x_2) \end{aligned} \tag{1.169}$$

and the KKT conditions are:

$$\frac{\partial L}{\partial x_1} = -3e^{-3x_1} + \mu_1 - \mu_2 = 0 \tag{1.170}$$
$$\frac{\partial L}{\partial x_2} = -2e^{-2x_2} + \mu_1 - \mu_3 = 0 \tag{1.171}$$
$$\mu_1(x_1 + x_2 - 2) = 0 \tag{1.172}$$

$$\mu_2(-x_1) = 0 \tag{1.173}$$
$$\mu_3(-x_2) = 0 \tag{1.174}$$
$$\mu_1 \geq 0 \tag{1.175}$$
$$\mu_2 \geq 0 \tag{1.176}$$
$$\mu_3 \geq 0 \tag{1.177}$$

First we may notice that $x_1 \geq 0$ and $x_2 \geq 0$, thus they can be either zero, or strictly positive. Therefore, we have four cases:

1.) $x_1 = 0$**,** $x_2 = 0$ *The relations (1.170), (1.171), (1.172) become:*

$$-3 + \mu_1 - \mu_2 = 0 \tag{1.178}$$
$$-2 + \mu_1 - \mu_3 = 0 \tag{1.179}$$
$$\mu_1(-2) = 0 \tag{1.180}$$

and $\mu_1 = 0$, $\mu_2 = -3$, $\mu_3 = -2$. The conditions (1.176) (1.177) are violated so $x_1 = x_2 = 0$ is not a solution of the problem.

2.) $x_1 = 0$**,** $x_2 > 0$ *Because x_2 is strictly positive, from (1.174) we obtain $\mu_3 = 0$, and the relations (1.170), (1.171), (1.172) are recalculated:*

$$-3 + \mu_1 - \mu_2 = 0 \tag{1.181}$$
$$-2e^{-2x_2} + \mu_1 = 0 \tag{1.182}$$
$$\mu_1(x_2 - 2) = 0 \tag{1.183}$$

From (1.182) we obtain $\mu_1 = 2e^{-2x_2} \neq 0$ so the relation (1.183) is satisfied only for $x_2 = 2$. Then $\mu_1 = 2e^{-4}$. From (1.181) we obtain: $\mu_2 = -3 + 2e^{-4} < 0$ and the constraint (1.176) is not satisfied.

This case will not give a solution of the problem.

3.) $x_1 > 0$**,** $x_2 = 0$ *Because x_1 is strictly positive, from (1.173) we obtain $\mu_2 = 0$, and the relations (1.170), (1.171), (1.172) are recalculated:*

$$-3e^{-3x_1} + \mu_1 = 0 \tag{1.184}$$
$$-2 + \mu_1 - \mu_3 = 0 \tag{1.185}$$
$$\mu_1(x_1 - 2) = 0 \tag{1.186}$$

From (1.184) we obtain $\mu_1 = 3e^{-3x_2} \neq 0$ so the relation (1.186) is satisfied only for $x_1 = 2$. Then $\mu_1 = 3e^{-6}$. From (1.185) we obtain: $\mu_3 = -2 + 3e^{-6} < 0$ and the constraint (1.177) is not satisfied.

This situation is not a solution of the problem either.

4.) $x_1 > 0$**,** $x_2 > 0$ *Since x_1 and x_2 cannot be zero, from (1.173) and (1.174) we obtain: $\mu_2 = 0$ and $\mu_3 = 0$.*

The relations (1.170), (1.171), (1.172) become:

$$-3e^{-3x_1} + \mu_1 = 0 \tag{1.187}$$
$$-2e^{-2x_2} + \mu_1 = 0 \tag{1.188}$$
$$\mu_1(x_1 + x_2 - 2) = 0 \tag{1.189}$$

The value of μ_1 cannot be zero because this would make zero the exponentials from (1.187) and (1.188) which is not valid. Then $x_1 + x_2 - 2 = 0$, or $x_2 = 2 - x_1$.

From (1.187) and (1.188) we obtain:

$$\mu_1 = 3e^{-3x_1} = 2e^{-2x_2} \tag{1.190}$$

and then:

$$3e^{-3x_1} = 2e^{-2(2-x_1)}, \text{ or } e^{-5x_1+4} = \frac{2}{3} \tag{1.191}$$

The solution is:

$$x_1 = \frac{1}{5}(4 - \ln\frac{2}{3}) = 0.88, \quad x_2 = 2 - x_1 = 1.11 \tag{1.192}$$

and $\mu_1 = 0.21 > 0$.

The necessary conditions for the point $(0.88,\ 1.11)$ *to be a minimizer of the constrained optimization problem are satisfied.*

The contour plot of f and the linear constraint are shown in Figure 1.17. The constraint $x_1 + x_2 = 2$ *is tangent to a level curve at the point* $P(0.88,\ 1.11)$ *thus it is the global minimizer of f subject to the constraints.*

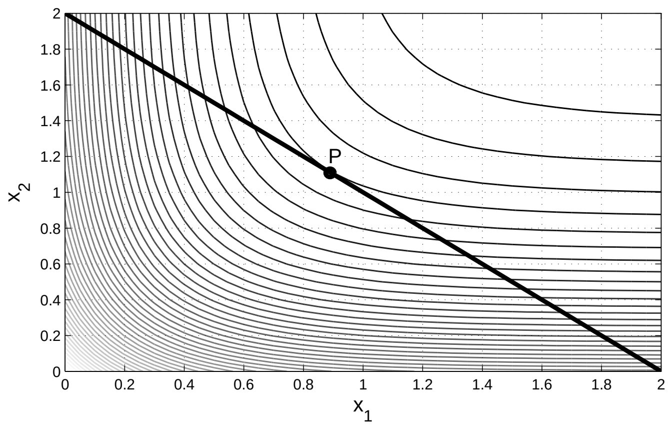

Figure 1.17: Contour plot of f and the constraint $x_1 + x_2 - 2 = 0$

1.4 Exercises

1. Locate the stationary points of the following functions and determine their character:

   a) $f(x) = x^7$

   b) $f(x) = 8x - x^4/4$

   c) $f(x) = 50 - 6x + x^3/18$

2. Consider the functions $f : \mathbb{R}^2 \to \mathbb{R}$:

   a) $f(x_1, x_2) = x_1 x_2$

   b) $f(x_1, x_2) = x_1/2 + x_2^2 - 3x_1 + 2x_2 - 5$

   c) $f(x_1, x_2) = -x_1^2 + x_2^3 + 6x_1 - 12x_2 + 5$

   d) $f(x_1, x_2) = 4x_1 x_2 - x_1^4 - x_2^4$

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

   Compute the stationary points and determine their character using the second derivative test.

3. Find the global minimum of the function:

$$f(x_1, x_2) = (x_1 - 2)^2 + (x_2 - 1)^2$$

in the region

$$0 \le x_1 \le 1$$
$$0 \le x_2 \le 2$$

4. Find the extrema of the function $f : \mathbb{R}^3 \to \mathbb{R}$:

$$f(x_1, x_2, x_3) = 2x_1^2 + 3x_2^2 + 4x_3^2 - 4x_1 - 12x_2 - 16x_3$$

5. Use the method of substitution to solve the constrained-optimization problem:

$$\min_{(x_1,x_2)} \left(x_1^2 + x_2^2 - 49\right)$$

subject to

$$x_1 + 3x_2 - 10 = 0$$

6. Use the method of Lagrange multipliers to find the maximum and minimum values of f subject to the given constraints

a) $f(x_1, x_2) = 3x_1 - 2x_2, \;\; x_1^2 + 2x_2^2 = 44$

b) $f(x_1, x_2, x_3) = x_1^2 - 2x_2 + 2x_3^3, \;\; x_1^2 + x_2^2 + x_3^2 = 1$

c) $f(x_1, x_2) = x_1^2 - x_2^2, \;\; x_1^2 + x_2^2 = 1$

7. Minimize the surface of a cylinder with a given volume.

Bibliography

Avriel, M. (2003). *Nonlinear Programming: Analysis and Methods*. Courier Dover Publications.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press.

Hancock, H. (1960). *Theory of Maxima and Minima*. Ginn and Company.

Hiriart-Urruty, J.-B. (1996). *L'Optimisation*. Que sais-je? Presses Universitaires de France.

Renze, J. and Weisstein, E. (2004). Extreme value theorem. From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/ExtremeValueTheorem.html.

Weisstein, E. (2004). Second derivative test. From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/SecondDerivativeTest.html.