Contents

1 Implicit Functions **3**
1.1 Existence Theorems for Implicit Functions 4
1.2 Existence Theorems for Systems of Implicit Functions 6
1.3 Change of Coordinates 8
1.4 Change of Variables 9

2 Extrema of Functions **13**
2.1 Local Extremum of a Function 14
2.2 Conditional Extrema 17
2.3 Exercises: Extrema of Functions 19

3 Answers to Exercises **21**
3.1 Exercises: Extrema of a Function (Solutions) 23

1.1 Existence Theorems for Implicit Functions

Let $A, B \subseteq \mathbb{R}$ let $F : A \times B \to \mathbb{R}$. Consider the equation

1.1.1
$$F(x, y) = 0.$$

1.1.2 **Definition.** *If, for all $x \in A$, equation (1.1.1) has a unique solution with respect to y, then a function $f : A \to B$ satisfying the equality $F(x, f(x)) = 0$, for all $x \in A$, is called an implicit function defined by equation (1.1.1) or, briefly,* `implicit function`.

The set of all points $(x, y) \in \mathbb{R}^2$ satisfying equation (1.1.1) represents the graph of the function f.

1.1.3 **Example.** Let $F : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$,

$$F(x, y) = y^3 - x + 1.$$

In this case, the implicit function is

$$f(x) = \sqrt[3]{x - 1}.$$

1.1.4 **Example.** Let $F : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$,

$$F(x, y) = y^5 + 2x^2y^3 + y + 7.$$

In this case, the implicit function cannot be represented explicitly.

Let $x_0, y_0 \in \mathbb{R}$, $U \in \mathcal{V}_{x_0}$, $V \in \mathcal{V}_{y_0}$ and $F : U \times V \to \mathbb{R}$ be a continuous function.

1.1.5 **Theorem.** *If the function F satisfies the conditions:*

(1) $F(x_0, y_0) = 0$;

(2) *there exist* $\dfrac{\partial F}{\partial x}, \dfrac{\partial F}{\partial y}$ *continuous on* $U \times V$;

(3) $\dfrac{\partial F}{\partial y}(x_0, y_0) \neq 0$,

then there exist $U_0 \in \mathcal{V}_{x_0}$, $V_0 \in \mathcal{V}_{y_0}$, and a function $f : U_0 \to V_0$ such that:

(a) $f(x_0) = y_0$;

(b) $F(x, f(x)) = 0$, *for all* $x \in U_0$;

(c) *f possesses continuous derivative on U_0 and*

$$f'(x) = -\frac{\dfrac{\partial F}{\partial x}(x, f(x))}{\dfrac{\partial F}{\partial y}(x, f(x))}.$$

Proof. Since the function $\dfrac{\partial F}{\partial y}$ is continuous, using the relation

$$\frac{\partial F}{\partial y}(x_0, y_0) \neq 0,$$

we deduce that there exist $U_1 \in \mathcal{V}_{x_0}$, $V_0 \in \mathcal{V}_{y_0}$ such that

$$\frac{\partial F}{\partial y}(x, y) \neq 0,$$

for all $x \in U_1$, and for all $y \in V_0$. This implies that, for all $x \in U_1$, the function $F(x, \cdot)$ is strictly monotone on V_0.

Let $\varepsilon > 0$ and choose $\varepsilon_1 \leq \varepsilon$ such that $[y_0 - \varepsilon_1, y_0 + \varepsilon_1] \subseteq V_0$. Using $F(x_0, y_0) = 0$ and the fact that the function $F(x_0, \cdot)$ is strictly monotone we obtain

$$F(x_0, y_0 - \varepsilon_1) \cdot F(x_0, y_0 + \varepsilon_1) < 0.$$

Since the function $F(\cdot, y_0 - \varepsilon_1)F(\cdot, y_0 + \varepsilon_1)$ is continuous there exists a neighborhood $U_\varepsilon \in \mathcal{V}_{x_0}$ such that

$$F(x, y_0 - \varepsilon_1) \cdot F(x, y_0 + \varepsilon_1) < 0,$$

for all $x \in U_\varepsilon$.

Let $U_0 = U_1 \cap U_\varepsilon$. Since the function $F(x, \cdot)$ is strictly monotone for all $x \in U_0$, we have defined a function $x \mapsto f(x) = y$ on U_0 such that

$$F(x, f(x)) = 0, \quad \forall x \in U_0.$$

But $F(x_0, y_0) = 0$, therefore

$$f(x_0) = y_0.$$

Moreover, for all $x \in U_\varepsilon$, we have

$$f(x) = y \in (f(x) - \varepsilon, f(x) + \varepsilon),$$

i.e., f is continuous at the point x_0.

Let $x' \in U_0$. The point $(x', f(x'))$ satisfies conditions (1), (2) and (3) from the statement of the theorem, so f is continuous at the point x', hence it is continuous on U_0.

Let us prove that f possesses a derivative in U_0. Let $x \in U_0$ and $h \neq 0$ such that $x + h \in U_0$. By virtue of Lagrange's mean-value theorem there exists $\theta \in (0, 1)$ such that

$$\begin{aligned} 0 &= F\left(x+h, f(x+h)\right) - F\left(x, f(x)\right) \\ &= h\frac{\partial F}{\partial x}\left(x + \theta h, f(x) + \theta(f(x+h) - f(x))\right) \\ &\quad + (f(x+h) - f(x))\frac{\partial F}{\partial y}\left(x + \theta h, f(x) + \theta(f(x+h) - f(x))\right). \end{aligned}$$

Since the function f is continuous we obtain

$$\lim_{h \to 0}(f(x+h) - f(x)) = 0,$$

and

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = -\frac{\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))},$$

i.e.,

$$f'(x) = -\frac{\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))}.$$

Since $\dfrac{\partial F}{\partial x}$ and $\dfrac{\partial F}{\partial x}$, are continuous in U_0, then f' is continuous on U_0.

Practically, by differentiating the equality $F(x, y) = 0$ with respect to x and taking into account that y is a function of x, we find

$$F'_x + y'F'_y = 0.$$

Differentiating again, we obtain

$$F''_{x^2} + y'F''_{xy} + y''F'_y + y'F''_{yx} + (y')^2 F''_{y^2} = 0,$$

hence, using the equality

$$y' = -\frac{F'_x}{F'_y},$$

we obtain

$$y'' = -\frac{F''_{x^2}(F'_y)^2 - 2F'_xF'_yF''_{xy} + (F'_x)^2F''_{y^2}}{(F'_y)^3}.$$

The higher order derivatives are obtained similarly.

Let $x_0 = (x_1^0, \ldots, x_n^0) \in \mathbb{R}^n$, $z_0 \in \mathbb{R}$, $U \in \mathcal{V}_{x_0}$, $V \in \mathcal{V}_{z_0}$.

1.1.6 THEOREM. *If a function $F : U \times V \to \mathbb{R}$ is continuous and satisfies the conditions:*

(1) $F(x_0; z_0) = 0$;

(2) *there exist* $\dfrac{\partial F}{\partial x_i}, \dfrac{\partial F}{\partial z}$ *continuous on* $U \times V$, $i = 1, \ldots, n$;

(3) $\dfrac{\partial F}{\partial z}(x_0; z_0) \neq 0$, *then there exist* $U_0 \in \mathcal{V}_{x_0}$, $V_0 \in \mathcal{V}_{z_0}$, *and a function* $f : U_0 \to V_0$ *such that:*

(a) $f(x_0) = z_0$;

(b) $F(x; f(x)) = 0, \quad \forall x \in U_0$;

(c) *f has partial derivatives continuous on U_0 and*

$$\frac{\partial f}{\partial x_i}(x) = -\frac{\dfrac{\partial F}{\partial x_i}(x; f(x))}{\dfrac{\partial F}{\partial z}(x; f(x))},$$

$i = 1, \ldots, n$.

1.1.7 EXAMPLE. Show that the equation

$$F(x, y) = x + y + \tan y = 0,$$

$x \in \mathbb{R}$, $y \in (-\pi/2, \pi/2)$ defines a strictly decreasing function y on $\mathbb{R}$.
From:

$$F(0,0) = 0, \; \frac{\partial F}{\partial y}(x, y) = 1 + \frac{1}{\cos^2 y} \neq 0, \; \forall x \in \mathbb{R}, \; \forall y \in (-\pi/2, \pi/2),$$

we obtain

$$1 + y' + \frac{y'}{\cos^2 y} = 0, \quad y' = -\frac{\cos^2 y}{1 + \cos^2 y} < 0,$$

so y is a strictly decreasing function on $\mathbb{R}$.

1.2 Existence Theorems for Systems of Implicit Functions

Let $U \subset \mathbb{R}^n$, $a \in U$, $F : U \to \mathbb{R}^m$, $F = (f_1, \ldots, f_m)$. Suppose that the partial derivatives $\dfrac{\partial f_i}{\partial x_j}(a)$, $i = 1, \ldots, m$; $j = 1, \ldots, p$, $p \leq n$, exist.

1.2.1 DEFINITION. *The matrix*

$$J_F(a) := \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \dfrac{\partial f_1}{\partial x_2}(a) & \ldots & \dfrac{\partial f_1}{\partial x_p}(a) \\ \dfrac{\partial f_2}{\partial x_1}(a) & \dfrac{\partial f_2}{\partial x_2}(a) & \ldots & \dfrac{\partial f_2}{\partial x_p}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(a) & \dfrac{\partial f_m}{\partial x_2}(a) & \ldots & \dfrac{\partial f_m}{\partial x_p}(a) \end{bmatrix}.$$

is called the `Jacobian matrix`[1] *of the function F at a point a with respect to the variables $x_1, \ldots, x_p$.*

[1] Karl Gustave Jacob Jacobi (1804–1851), a German mathematician.

Let $m = p$.

1.2.2 DEFINITION. *The determinant of the Jacobian matrix is called the* `Jacobian determinant` *or the* `functional determinant` *of the functions $f_1, \ldots, f_m$ with respect to the variables $x_1, \ldots, x_m$, at the point a, and is denoted by*

$$\frac{D(f_1, \ldots, f_m)}{D(x_1, \ldots, x_m)}(a).$$

```
In[1]:= (* Mathematica *)

In[2]:= (* The Jacobian matrix for a vector function {F1,F2} *)

In[3]:= D[{F1[x1, x2], F2[x1, x2]}, {{x1, x2}}]
```

Out[3]= $\{\{F_1^{(1,0)}[x_1, x_2], F_1^{(0,1)}[x_1, x_2]\}, \{F_2^{(1,0)}[x_1, x_2], F_2^{(0,1)}[x_1, x_2]\}\}$

```
In[4]:= MatrixForm[%]
```

Out[4]//MatrixForm= $\begin{pmatrix} F_1^{(1,0)}[x_1, x_2] & F_1^{(0,1)}[x_1, x_2] \\ F_2^{(1,0)}[x_1, x_2] & F_2^{(0,1)}[x_1, x_2] \end{pmatrix}$

```
In[5]:= (* The Hessian matrix for F *)

In[6]:= D[F[x1, x2], {{x1, x2}, 2}]
```

Out[6]= $\{\{F^{(2,0)}[x_1, x_2], F^{(1,1)}[x_1, x_2]\}, \{F^{(1,1)}[x_1, x_2], F^{(0,2)}[x_1, x_2]\}\}$

```
In[7]:= MatrixForm[%]
```

Out[7]//MatrixForm= $\begin{pmatrix} F^{(2,0)}[x_1, x_2] & F^{(1,1)}[x_1, x_2] \\ F^{(1,1)}[x_1, x_2] & F^{(0,2)}[x_1, x_2] \end{pmatrix}$

1.2.3 THEOREM. *If the functions F, G possess continuous partial derivatives in a neighborhood of a point $(x_0, y_0, u_0, v_0) \in \mathbb{R}^4$ and satisfy the conditions:*

(1) $F(x_0, y_0, u_0, v_0) = 0,$

(2) $G(x_0, y_0, u_0, v_0) = 0,$

(3) $\dfrac{D(F, G)}{D(u, v)}(x_0, y_0, u_0, v_0) \neq 0,$

then the system

$$\begin{cases} F(x, y, u, v) = 0, \\ G(x, y, u, v) = 0, \end{cases}$$

defines two implicit functions

$$u = u(x, y), \qquad v = v(x, y),$$

possessing continuous partial derivatives in a neighborhood V of the point (x_0, y_0), and these functions satisfy the relations:

$$\begin{cases} F(x, y, u(x, y), v(x, y)) = 0, \\ G(x, y, u(x, y), v(x, y)) = 0 \end{cases}$$

$$u(x_0, y_0) = u_0, \qquad v(x_0, y_0) = v_0,$$

for all $(x, y) \in V$.

The partial derivatives u'_x, u'_y, v'_x, v'_y can be found by differentiating the previous system. Hence,

$$\begin{cases} F'_x + u'_x F'_u + v'_x F'_v &= 0, \\ G'_x + u'_x G'_u + v'_x G'_v &= 0, \end{cases}$$

and

$$u'_x = -\frac{\frac{D(F,G)}{D(x,v)}}{\frac{D(F,G)}{D(u,v)}}, \qquad v'_x = -\frac{\frac{D(F,G)}{D(u,x)}}{\frac{D(F,G)}{D(u,v)}},$$

Likewise, we obtain:

$$u'_y = -\frac{\frac{D(F,G)}{D(y,v)}}{\frac{D(F,G)}{D(u,v)}}, \qquad v'_y = -\frac{\frac{D(F,G)}{D(u,y)}}{\frac{D(F,G)}{D(u,v)}}.$$

1.3 Change of Coordinates

Let $A, B \subseteq \mathbb{R}^n$ be open sets, $n \in \mathbb{N}^*$.

1.3.1 Definition. *A continuous bijective function $F : A \to B$ with a continuous inverse is called a* `homeomorphism`.

1.3.2 Definition. *A homeomorphism $F : A \to B$ possessing a continuous differentiable inverse is called a* `diffeomorphism`.

1.3.3 Theorem. *If $F : A \to B$ is a diffeomorphism, then the Jacobian $\det(J_F(a))$ is different from zero and*

$$(J_F(a))^{-1} = J_{F^{-1}}(F(a)),$$

for all $a \in A$.

1.3.4 Definition. *A mapping $F : A \to \mathbb{R}^n$, such that $F : A \to F(A)$ is a homeomorphism is called a* `change of coordinates` *in A.*

If $F = (f_1, \ldots, f_n)$, then $(f_1(x), \ldots, f_n(x))$ are the `coordinates` of the point x in the coordinate system $(f_1, \ldots, f_n)$.

1.3.5 Example. For $A = \{(x, y) \in \mathbb{R}^2 \mid x > 0,\ y > 0\}$ and $F : A \to \mathbb{R}^2$,

$$F(x, y) = (\rho(x, y), \theta(x, y)),$$

$$\rho = \rho(x, y) = \sqrt{x^2 + y^2},$$

$$\theta = \theta(x, y) = \arctan \frac{y}{x},$$

we have

$$\frac{D(\rho, \theta)}{D(x, y)} = \begin{vmatrix} \frac{\partial \rho}{\partial x} & \frac{\partial \rho}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{vmatrix} = \frac{1}{\sqrt{x^2+y^2}} > 0.$$

The function F changes the Cartesian coordinates into polar coordinates.

1.4 Change of Variables

The study of some sets of points (curves, surfaces, etc.) whose elements are determined by their coordinates with respect to a certain `coordinate system`, can be simplified by a convenient change of the coordinate system, which is called a `change of variables`.

Let $A, B \subset \mathbb{R}^2$ be two open sets and $F : A \to B$ be a diffeomorphism, $F \in C^k(A)$, $k \geq 2$,

$$F(u,v) = (h(u,v), g(u,v)) = (x,y).$$

Consider a curve $(\Gamma) \subseteq B$ which in the "old" coordinate system, (x,y), is specified by the equation

$$y = y(x).$$

In the "new" coordinate system, (u,v), the curve (Γ) is specified by the equation

$$g(u,v) = y(h(u,v)),$$

hence, by theorem (1.1.5) , we obtain a local equality

$$v = v(u).$$

We shall find the relation between the derivatives of the functions

$$y = y(x), \qquad v = v(u).$$

We have:

$$y'(x) = \frac{dy}{dx} = \frac{\frac{dy}{du}}{\frac{dx}{du}} = \frac{\frac{\partial g}{\partial u} + \frac{\partial g}{\partial v}v'(u)}{\frac{\partial h}{\partial u} + \frac{\partial h}{\partial v}v'(u)},$$

$$y''(x) = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{1}{\frac{dx}{du}} \cdot \frac{d}{du}\left(\frac{\frac{dy}{du}}{\frac{dx}{du}}\right) = \frac{\frac{d^2y}{du^2} \cdot \frac{dx}{du} - \frac{d^2x}{du^2} \cdot \frac{dy}{du}}{\left(\frac{dx}{du}\right)^3},$$

etc.

Some particular cases will be discussed.

(I). Interchange of Variables. Considering $F(u,v) = (v,u)$, i.e., $x = v$, $y = u$, we obtain:

$$y'(x) = \frac{dy}{dx} = \frac{1}{\frac{dx}{dy}} = \frac{1}{x'(y)},$$

$$y''(x) = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{d}{dx}\left(\frac{1}{x'(y)}\right) = \frac{1}{\frac{dx}{dy}} \cdot \frac{d}{dy}\left(\frac{1}{x'(y)}\right) = -\frac{x''(y)}{(x'(y))^3}.$$

1.4.1 Example. Transform the equation

$$y(y')^3 + y'' = 0,$$

where $y = y(x)$, by interchange of variables.

We have:

$$y' = \frac{1}{x'}, \qquad y'' = -\frac{x''}{(x')^3}.$$

The transformed equation is $x'' - y = 0$, where $x = x(y)$.

(II). Change of the Independent Variable. For $x = \varphi(u)$, choosing $Y = y \circ \varphi$, we obtain:

$$y'(x) = \frac{dy}{dx} = \frac{\frac{dY}{du}}{\frac{dx}{du}} = \frac{Y'(u)}{\varphi'(u)},$$

$$y''(x) = \frac{d}{dx}\left(\frac{Y'(u)}{\varphi'(u)}\right) = \frac{1}{\varphi'(u)} \cdot \frac{d}{du}\left(\frac{Y'(u)}{\varphi'(u)}\right)$$
$$= \frac{Y''(u)\varphi'(u) - Y'(u)\varphi''(u)}{(\varphi'(u))^3}.$$

1.4.2 EXAMPLE. Transform the equation

$$x^2y'' + xy' - y = 0,$$

where $y = y(x)$, by change of the independent variable

$$x = \mathsf{e}^u.$$

We have:

$$y'(x) = \mathsf{e}^{-u}Y'(u),$$
$$y''(x) = \mathsf{e}^{-2u}\left(Y''(u) - Y'(u)\right).$$

The transformed equation is

$$Y'' - Y = 0,$$

where $Y = Y(u)$.

(III). Two Independent Variables. Let $A, B \subset \mathbb{R}^2$ be open sets and $T : A \to B$ be a diffeomorphism,

$$T(u, v) = (x(u, v), y(u, v)).$$

Consider the change of variables

$$\begin{cases} x = x(u, v), \\ y = y(u, v), \end{cases} \quad (u, v) \in A.$$

We will find a relation between the operators $\dfrac{\partial}{\partial x}, \dfrac{\partial}{\partial y}$, with respect to the "old" variables x, y and the operators $\dfrac{\partial}{\partial u}, \dfrac{\partial}{\partial v}$, with respect to the "new" variables u, v.

Let $z : B \to \mathbb{R}$ be an arbitrary function $z \in C^1(B)$. We define the function $Z : A \to \mathbb{R}$,

$$Z = z \circ T.$$

We have:

$$\begin{cases} \dfrac{\partial Z}{\partial u} = \dfrac{\partial x}{\partial u} \cdot \dfrac{\partial z}{\partial x} + \dfrac{\partial y}{\partial u} \cdot \dfrac{\partial z}{\partial y} \\ \dfrac{\partial Z}{\partial v} = \dfrac{\partial x}{\partial v} \cdot \dfrac{\partial z}{\partial x} + \dfrac{\partial y}{\partial v} \cdot \dfrac{\partial z}{\partial y} \end{cases},$$

hence:

1.4.3
$$\begin{cases} \dfrac{\partial z}{\partial x} = \dfrac{1}{\dfrac{D(x,y)}{D(u,v)}}\left(\dfrac{\partial y}{\partial v} \cdot \dfrac{\partial Z}{\partial u} - \dfrac{\partial y}{\partial u} \cdot \dfrac{\partial Z}{\partial v}\right) \\ \dfrac{\partial z}{\partial y} = \dfrac{1}{\dfrac{D(x,y)}{D(u,v)}}\left(-\dfrac{\partial x}{\partial v} \cdot \dfrac{\partial Z}{\partial u} + \dfrac{\partial x}{\partial u} \cdot \dfrac{\partial Z}{\partial v}\right) \end{cases},$$

so:

$$\begin{cases} \dfrac{\partial}{\partial x} = \dfrac{1}{\dfrac{D(x,y)}{D(u,v)}}\left(\dfrac{\partial y}{\partial v} \cdot \dfrac{\partial}{\partial u} - \dfrac{\partial y}{\partial u} \cdot \dfrac{\partial}{\partial v}\right) \\ \dfrac{\partial}{\partial y} = \dfrac{1}{\dfrac{D(x,y)}{D(u,v)}}\left(-\dfrac{\partial x}{\partial v} \cdot \dfrac{\partial}{\partial u} + \dfrac{\partial x}{\partial u} \cdot \dfrac{\partial}{\partial v}\right) \end{cases},$$

Since the domains of the the operators $\dfrac{\partial}{\partial x}$ and $\dfrac{\partial}{\partial y}$ are different from the domains of the operators $\dfrac{\partial}{\partial u}$ and $\dfrac{\partial}{\partial v}$, the previous equalities must be taken in the sense of formulas (1.4.3).

1.4.4 EXAMPLE. In the case of the change of variables

$$\begin{cases} x = \rho\cos\theta \\ y = \rho\sin\theta \end{cases},$$

$\rho \in (0,\infty)$, $\theta \in [0,2\pi)$, we obtain:

$$\begin{cases} \dfrac{\partial}{\partial x} = \cos\theta\cdot\dfrac{\partial}{\partial\rho} - \dfrac{\sin\theta}{\rho}\cdot\dfrac{\partial}{\partial\theta}, \\ \dfrac{\partial}{\partial y} = \sin\theta\cdot\dfrac{\partial}{\partial\rho} + \dfrac{\cos\theta}{\rho}\cdot\dfrac{\partial}{\partial\theta}. \end{cases}$$

2.1 Local Extremum of a Function

Let $D \subset \mathbb{R}^n$ and $f : D \to \mathbb{R}$.

2.1.1 DEFINITION. *The function f has a* `local maximum` *at a point $x_0 \in D$ if there exists a neighborhood $U \in \mathcal{V}_{x_0}$ such that*

$$f(x) \leq f(x_0), \quad \forall x \in U.$$

On the other hand, f has a `local minimum` *at the point x_0 if*

$$f(x_0) \leq f(x), \quad \forall x \in U.$$

Local minima and local maxima are referred to as `local extrema`. A point at which a function has an extremum is called a `point of extremum`.

In what follows, assume that f possesses partial derivatives at the point $x_0 \in D$.

2.1.2 DEFINITION. *A point x_0 at which*

$$\nabla f(x_0) = 0$$

is called a `stationary` *or* `critical point` *of f.*

2.1.3 THEOREM. *Any interior point of local extremum of f is a critical point.*

In what follows assume that the function f possesses continuous derivatives of order $p \geq 2$ in D.

2.1.4 DEFINITION. *The differential $d^2 f(x)$ is said to be:*

`positive definite` *if $d^2 f(x)(h) > 0$;*
`negative definite` *if $d^2 f(x)(h) < 0$;*
`positive semidefinite` *if $d^2 f(x)(h) \geq 0$;*
`negative semidefinite` *if $d^2 f(x)(h) \leq 0$; for all $h \in \mathbb{R}^n$, $h \neq 0$.*
Otherwise, it is `indefinite`.

2.1.5 THEOREM. **(Sylvester)** *Using the notations*

$$a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x), \qquad i, j = 1, \ldots, n,$$

we have:

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \ \ldots, \ \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} > 0,$$

if and only if $d^2 f(x)$ is positive definite;

$$a_{11} < 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \ \ldots, \ (-1)^n \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} > 0,$$

if and only if $d^2 f(x)$ is negative definite.

The Sylvester theorem is proved in the theory of quadratic forms.

2.1.6 THEOREM. *Let $x_0 \in D$ be a critical point of the function f.*
If $d^2 f(x_0)$ is positive definite, then x_0 is a point of local minimum.
If $d^2 f(x_0)$ is negative definite, then x_0 is a point of local maximum.

Proof. Taylor's formula gives

$$f(x) = f(x_0) + df(x_0)(x - x_0) + \frac{1}{2}\, d^2 f(x_0 + \theta(x - x_0))(x - x_0),$$

where $\theta \in (0,1)$. But $df(x_0) = 0$, then the previous equality can be written as

$$f(x) - f(x_0) = \frac{1}{2}\, d^2 f(x_0 + \theta(x - x_0))(x - x_0).$$

For x sufficiently close to x_0, the sign of the expression $f(x) - f(x_0)$ coincides with the sign of $d^2 f(x_0)(x - x_0)$.

For definiteness, let V be a neighborhood of the point x_0 such that $d^2 f(x_0)(x - x_0) > 0$, for all $x \in V \setminus \{x_0\}$. It follows that

$$f(x) - f(x_0) > 0, \quad \forall x \in V \setminus \{x_0\},$$

hence x_0 is a point of local minimum of f. Similarly we treat the case when $d^2 f(x_0)$ is negative definite.

2.1.7 REMARK. *If $d^2 f(x_0)$ is semidefinite then the question whether f has an extremum at x_0 remains open because there are examples where f has an extremum at x_0 or has no extremum.*
If $d^2 f(x_0)$ is indefinite, then it is sure that f has no extremum at x_0.

In order to investigate a function f for an extremum, we can run through the following steps:

- find the interior critical points of f by solving the system of equations

$$\left\{ \frac{\partial f}{\partial x_i}(x) = 0\,, \quad i = 1, \ldots, n; \right.$$

- find the sign of the second differential of f at the critical points using, e.g., Sylvester's criterion;

- study the local extrema of f on the boundary of the domain.

2.1.8 EXAMPLE. Determine the local extrema of the function $f : \mathbb{R}^2 \to \mathbb{R}$, such that

$$f(x, y) = (x + y + 1)^3 - 27xy.$$

The critical points are the solutions of the system

$$\begin{cases} \dfrac{\partial f}{\partial x}(x, y) = 0, \\[2ex] \dfrac{\partial f}{\partial y}(x, y) = 0, \end{cases} \iff \begin{cases} 3(x + y + 1)^2 - 27y = 0, \\ 3(x + y + 1)^2 - 27x = 0, \end{cases}$$

i.e.,

$$(1, 1) \quad \text{and} \quad \left(\frac{1}{4}, \frac{1}{4}\right).$$

We have:

$$\begin{cases} \dfrac{\partial^2 f}{\partial x^2}(x,y) & = & 6(x+y+1), \\ \dfrac{\partial^2 f}{\partial x \partial y}(x,y) & = & 6(x+y+1)-27, \\ \dfrac{\partial^2 f}{\partial y^2}(x,y) & = & 6(x+y+1), \end{cases}$$

$$\begin{cases} \dfrac{\partial^2 f}{\partial x^2}(1,1) & = & 18, \\ \dfrac{\partial^2 f}{\partial x \partial y}(1,1) & = & -9, \\ \dfrac{\partial^2 f}{\partial y^2}(1,1) & = & 18, \end{cases} \qquad \begin{cases} \dfrac{\partial^2 f}{\partial x^2}\left(\dfrac{1}{4},\dfrac{1}{4}\right) & = & 9, \\ \dfrac{\partial^2 f}{\partial x \partial y}\left(\dfrac{1}{4},\dfrac{1}{4}\right) & = & -18, \\ \dfrac{\partial^2 f}{\partial y^2}\left(\dfrac{1}{4},\dfrac{1}{4}\right) & = & 9. \end{cases}$$

hence we deduce that the differential

$$d^2 f(1,1)(dx,dy) = 18\,dx^2 - 18\,dxdy + 18\,dy^2$$

is positive definite, so $(1,1)$ is a point of local minimum.

The differential

$$d^2 f\left(\frac{1}{4},\frac{1}{4}\right)(dx,dy) = 9\,dx^2 - 36\,dxdy + 9\,dy^2$$

is indefinite, therefore $\left(\dfrac{1}{4},\dfrac{1}{4}\right)$ is not a point of local extremum of f.

```
In[1]:= (* Mathematica *)

In[2]:= f = (x + y + 1)^3 - 27 x y;

In[3]:= FindMinimum[(x + y + 1)^3 - 27 x y, {{x, 0.5, -2, 2}, {y, 0.6, -2, 2}}]

Out[3]= {0., {x → 1., y → 1.}}
```

2.1.9 EXAMPLE. Let us investigate the function $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = x^2 + y^3,$$

for an extremum.

Solving the system

$$\begin{cases} 2x & = & 0, \\ 3y^2 & = & 0, \end{cases}$$

we find the critical points. The point $(0,0)$ is the only critical point. We have:

$$df(x,y) = 2x\,dx + 3y^2\,dy, \qquad d^2 f(x,y) = 2\,dx^2 + 6y\,dy^2,$$

$$d^2 f(0,0)(dx,dy) = 2dx^2 \geq 0,$$

so $d^2 f(0,0)$ is positive semidefinite and therefore we can say nothing about the point $(0,0)$. But,

$$f(0,y) = y^3, \quad f(0,0) = 0,$$

hence $(0,0)$ is not a point of extremum.

2.2 Conditional Extrema

Let $U \subseteq \mathbb{R}^{n+m}$ be an open set and $f \in C^1(U)$.

Consider the functions $g_i \in C^1(U)$, $i = 1, \ldots, m$. Let $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$. Suppose that the variables x and y are related by

2.2.1
$$\begin{cases} g_1(x;y) = 0, \\ g_2(x;y) = 0, \\ \ldots\ldots\ldots\ldots \\ g_m(x;y) = 0, \end{cases}$$

which are called `constraints`, or `constraint equations`.

Let A be the set of points $(x;y) \in U$ for which the constraints (2.2.1), hold simultaneously, i.e.,

$$A = \{(x;y) \in U \mid g_i(x;y) = 0,\ i = 1, \ldots, m\}.$$

2.2.2 DEFINITION. *A point $(x_0;y_0) \in A$ is called a* `point of local extremum` *of a function f under the constraints (2.2.1) if there exists a neighborhood $W \in \mathcal{V}_{(x_0;y_0)}$ such that the difference $f(x;y) - f(x_0;y_0)$ does not change its sign on $A \cap W$.*

Notice that the point $(x_0;y_0)$ is a point of local extremum of the function f/A.

Let us present the `method of Lagrange multipliers` for determining the critical points.

2.2.3 THEOREM. **(Lagrange)** *If $(x_0;y_0)$ is a point of local extremum of f under the constraints (2.2.1) and*

$$\frac{D(g_1, \ldots, g_m)}{D(y_1, \ldots, y_m)}(x_0;y_0) \neq 0,$$

then there exists m real numbers $\lambda_1, \ldots, \lambda_m$, called `Lagrange multipliers,` *such that the auxiliary function*

$$\Phi = f + \sum_{i=1}^{m} \lambda_i g_i,$$

called `Lagrangian,` *satisfies the following conditions:*

$$\begin{cases} \dfrac{\partial \Phi}{\partial x_j}(x_0;y_0) = 0, \\ \dfrac{\partial \Phi}{\partial y_k}(x_0;y_0) = 0, \end{cases}$$

$j = 1, \ldots, n; \quad k = 1, \ldots, m.$

In order to solve a problem of local extremum under constraints we can run through the following steps:

(1) Solve the system

$$\begin{cases} \dfrac{\partial \Phi}{\partial x_j}(x;y) & = & 0 \\ \dfrac{\partial \Phi}{\partial y_k}(x;y) & = & 0 \\ g_k(x;y) & = & 0, \end{cases}$$

$j = 1, \ldots, n; \quad k = 1, \ldots, m$, with the unknowns:

$$x_1, \ldots, x_n, \quad y_1, \ldots, y_m, \quad \lambda_1, \ldots, \lambda_m;$$

(2) Let $(x_0;y_0;\lambda_0)$ be a solution of the previous system. Find the constraints between the differentials $dx_1, \ldots, dx_n$, and $dy_1, \ldots, dy_m$ at the point $(x_0;y_0)$.

(3) Find the sign of $d^2\Phi(x_0; y_0)$ taking into account the relations between the differentials $dx_1, \dots, dx_n$.

2.2.4 EXAMPLE. Determine the local conditional extrema of the function

$$H(x_1, \dots, x_n) = \sum_{i=1}^{n} x_i \ln x_i,$$

$x_i > 0,\ i = 1, \dots, n$, under the constraints

$$x_1 + \dots + x_n = 1.$$

Consider the function

$$\Phi = \sum_{i=1}^{n} x_i \ln x_i + \lambda \left(\sum_{i=1}^{n} x_i - 1 \right).$$

The system

$$\begin{cases} \dfrac{\partial \Phi}{\partial x_1} & = & 0 \\ \dfrac{\partial \Phi}{\partial x_2} & = & 0 \\ \dots\dots\dots\dots & \dots & \dots \\ \dfrac{\partial \Phi}{\partial x_n} & = & 0 \\ x_1 + \dots + x_n & = & 0 \end{cases}$$

can be written in the form

$$\begin{cases} \ln x_1 + 1 + \lambda & = & 0 \\ \ln x_2 + 1 + \lambda & = & 0 \\ \dots\dots\dots\dots & \dots & \dots \\ \ln x_n + 1 + \lambda & = & 0 \\ x_1 + \dots + x_n & = & 0. \end{cases}$$

It has the solution $\left(\frac{1}{n}, \dots, \frac{1}{n}, \ln n - 1\right)$. We have:

$$\Phi(x_1, \dots, x_n) = x_1 \ln x_1 + \dots + x_n \ln x_n + (\ln n - 1)(x_1 + \dots + x_n - 1),$$

$$\begin{aligned} d\Phi(x_1, \dots, x_n) &= \sum_{i=1}^{n} (\ln x_i + \ln n)\, dx_i, \\ d^2\Phi(x_1, \dots, x_n) &= \sum_{i=1}^{n} \frac{1}{x_i}\, dx_i^2, \\ d^2\Phi(\tfrac{1}{n}, \dots, \tfrac{1}{n}) &= n \sum_{i=1}^{n} dx_i^2, \end{aligned}$$

which is positive definite independent of the constraint

$$dx_1 + \dots + dx_n = 0.$$

It follows that the point $\left(\frac{1}{n}, \dots, \frac{1}{n}\right)$ is a point of local minimum of the function H under the constraint

$$x_1 + \dots + x_n = 1.$$

2.3 Exercises: Extrema of Functions

2.3.1 Prove that the function

$$f(x,y) = (1 + \mathrm{e}^y)\,\cos x - y\,\mathrm{e}^y, \qquad (x,y) \in \mathbb{R}^2,$$

has an infinite number of maxima and no minimum.

2.3.2 Find the distance from the point $M_0(x_0, y_0, z_0)$ to the plane

$$Ax + By + Cz + D = 0.$$

2.3.3 Find the distance between the straight lines

$$2(x-1) = y = 2z, \qquad x = y = z.$$

2.3.4 Among all rectangular parallelepipeds having the given volume 1, find the parallelepiped with the least area.

2.3.5 Find the greatest and the least values of the function

$$f(x,y) = x^2 + y^2$$

in the disc

$$D: \;(x-2)^2 + (y-2)^2 \leq 2.$$

3.1 Exercises: Extrema of a Function (Solutions)

2.3.1 We have:

$$\begin{cases} \frac{\partial f}{\partial x} = 0, \\ \frac{\partial f}{\partial y} = 0, \end{cases} \iff \begin{cases} -(1+\mathrm{e}^y)\sin x = 0, \\ \mathrm{e}^y(\cos x - 1 - y) = 0. \end{cases}$$

Solving the system, we find the stationary points:

$$(k\pi, (-1)^k - 1), \quad k \in \mathbb{Z},$$

i.e.,

$$(2n\pi, 0), \quad ((2n+1)\pi, -2), \qquad n \in \mathbb{Z}.$$

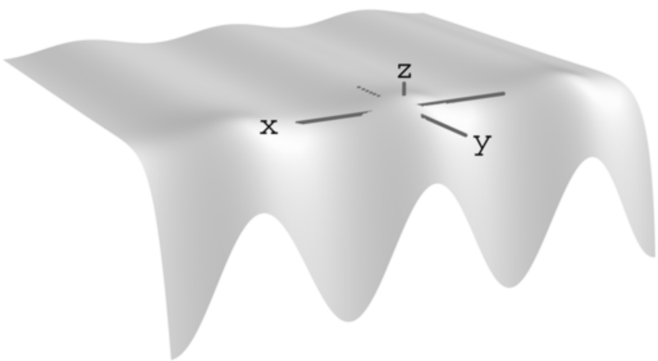
Further, we find the partial derivatives of second order:

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= -(1+\mathrm{e}^y)\cos x, \\ \frac{\partial^2 f}{\partial x \partial y} &= -\mathrm{e}^y \sin x, \\ \frac{\partial^2 f}{\partial y^2} &= \mathrm{e}^y(\cos x - 2 - y). \end{aligned}$$

We calculate the second differential at the stationary points:

$$\begin{aligned} d^2 f(2n\pi, 0) &= -2\,\mathrm{d}x^2 - \mathrm{d}y^2, \quad \text{negative definite}, \\ d^2 f((2n+1)\pi, -2) &= (1+\mathrm{e}^{-2})\,\mathrm{d}x^2 - \mathrm{e}^{-2}\,\mathrm{d}y^2, \quad \text{indefinite}. \end{aligned}$$

Consequently, $(2n\pi, 0)$ are points of maximum and $((2n+1)\pi, -2)$ are not points of extremum.



2.3.2 We find the minimum of the function

$$f(x,y,z) = (x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2$$

with the constraint

$$Ax + By + Cz + D = 0.$$

The problem is reduced to examine the Lagrange function

$$L(x,y,z) = f(x,y,z) + \lambda(Ax + By + Cz + D)$$

for an ordinary extremum. Consider the system

$$\begin{cases} \frac{\partial L}{\partial x} & = 0, \\ \frac{\partial L}{\partial y} & = 0, \\ \frac{\partial L}{\partial z} & = 0, \\ F(x,y,z) & = 0, \end{cases} \iff \begin{cases} 2(x-x_0)+\lambda A & = 0, \\ 2(y-y_0)+\lambda B & = 0, \\ 2(z-z_0)+\lambda C & = 0, \\ Ax+By+Cz+D & = 0. \end{cases}$$

We obtain

$$\lambda = 2\frac{Ax_0+By_0+Cz_0+D}{A^2+B^2+C^2}.$$

The stationary point is

$$\left(x_0-\frac{\lambda A}{2}, y_0-\frac{\lambda B}{2}, z_0-\frac{\lambda C}{2}\right).$$

The second differential

$$d^2L = 2(\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2)$$

is positive definite, hence the stationary point is a point of minimum.

$$f_{min} = \left(\frac{Ax_0+By_0+Cz_0+D}{A^2+B^2+C^2}\right)^2,$$

hence, the distance from the point $M_0(x_0,y_0,z_0)$ to the plane $Ax+By+Cz+D=0$ is

$$\frac{|Ax_0+By_0+Cz_0+D|}{\sqrt{A^2+B^2+C^2}}.$$

2.3.3 The parametric equations of the straight lines are:

$$\begin{cases} x & = u+1, \\ y & = 2u, \\ z & = u, \end{cases} \quad \text{and} \quad \begin{cases} x & = v, \\ y & = v, \\ z & = v. \end{cases}$$

We examine the function $f(u,v) = (u+1-v)^2+(2u-v)^2(u-v)^2$ for a minimum. Consider the system

$$\begin{cases} \frac{\partial f}{\partial u} & = 0, \\ \frac{\partial f}{\partial v} & = 0, \end{cases} \iff \begin{cases} 12u-8v+2 & = 0, \\ -8u+6v-2 & = 0. \end{cases}$$

The stationary point is $(\frac{1}{2},1)$. Further, we find the partial derivatives of second order:

$$\begin{cases} \frac{\partial^2 f}{\partial x^2} & = 12, \\ \frac{\partial^2 f}{\partial x \partial y} & = -8, \\ \frac{\partial^2 f}{\partial y^2} & = 6. \end{cases}$$

We calculate the second differential at the stationary point:

$$d^2f\left(\frac{1}{2},1\right) = 12\,du^2-16\,du\,dv+6\,dv^2.$$

We have

$$\begin{vmatrix} 12 & -8 \\ -8 & 6 \end{vmatrix} = 8 > 0,$$

hence, by Sylvester criterion, the second differential is positive definite; therefore $(\frac{1}{2}, 1)$ is a point of minimum. We obtain

$$f_{min} = f\left(\frac{1}{2}, 1\right) = \frac{1}{2}.$$

The distance is $\frac{\sqrt{2}}{2}$.

2.3.4 We find the minimum of the function

$$f(x, y, z) = xy + xz + yz$$

with the constraint

$$xyz - 1 = 0, \quad x > 0, \quad y > 0, \quad z > 0.$$

The problem is reduced to examine the Lagrange function

$$L(x, y, z) = f(x, y, z) + \lambda(xyz - 1)$$

for an ordinary extremum. Consider the system

$$\begin{cases} \frac{\partial L}{\partial x} & = 0, \\ \frac{\partial L}{\partial y} & = 0, \\ \frac{\partial L}{\partial z} & = 0, \\ xyz - 1 & = 0, \end{cases} \iff \begin{cases} y + z + \lambda yz & = 0, \\ x + z + \lambda xz & = 0, \\ x + y + \lambda xy & = 0, \\ xyz - 1 & = 0. \end{cases}$$

We obtain $\lambda = -2$. The stationary point is $(1, 1, 1)$. From $xyz = 1$ we deduce

$$yz\mathrm{d}x + xz\mathrm{d}y + xy\mathrm{d}z = 0.$$

At the point $(1, 1, 1)$ we obtain

$$\mathrm{d}x + \mathrm{d}y + \mathrm{d}z = 0,$$

hence,

$$\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 + 2(\mathrm{d}x\mathrm{d}y + \mathrm{d}x\mathrm{d}z + \mathrm{d}y\mathrm{d}z) = 0. \tag{$\diamond$}$$

The second differential at the stationary point is

$$d^2L(1, 1, 1) = -2(\mathrm{d}x\mathrm{d}y + \mathrm{d}x\mathrm{d}z + \mathrm{d}y\mathrm{d}z).$$

Using $(\diamond)$, we obtain

$$d^2L(1, 1, 1) = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2$$

which is positive definite; consequently $(1, 1, 1)$ is a point of minimum and the smallest area is equal to 3.

2.3.5 First we find the ordinary extrema of the function f on the interior of the disc. Consider the system

$$\begin{cases} \frac{\partial f}{\partial x} & = 0, \\ \frac{\partial f}{\partial y} & = 0, \end{cases} \iff \begin{cases} 2x & = 0, \\ 2y & = 0. \end{cases}$$

The stationary point $(0, 0)$ does not belong to D. Furthermore, we find the extrema of the function f with the constraint

$$(x - 2)^2 + (y - 2)^2 - 2 = 0.$$

We examine the Lagrange function

$$L_\lambda(x, y) = f(x, y) + \lambda((x - 2)^2 + (y - 2)^2 - 2)$$

for an ordinary extremum. Consider the system

$$\begin{cases} \frac{\partial L}{\partial x} &= 0, \\ \frac{\partial L}{\partial y} &= 0, \\ (x-2)^2 + (y-2)^2 &= 2, \end{cases} \iff \begin{cases} 2(x + \lambda(x-2)) &= 0, \\ 2(y + \lambda(y-2)) &= 0, \\ (x-2)^2 + (y-2)^2 &= 2. \end{cases}$$

We obtain:
$\lambda_1 = -3$; the first stationary point is $(3, 3)$;
$\lambda_2 = 1$; the second stationary point is $(1, 1)$.
We have:

$$\begin{aligned} \mathrm{d}^2 L(x, y) &= \ (2 + 2\lambda)\,\mathrm{d}x^2 + (2 + 2\lambda)\,\mathrm{d}y^2, \\ \mathrm{d}^2 L(3, 3) &= -4(\mathrm{d}x^2 + \mathrm{d}y^2), \quad \text{positive definite}, \\ \mathrm{d}^2 L(1, 1) &= \ 4(\mathrm{d}x^2 + \mathrm{d}y^2), \quad \text{negative definite}. \end{aligned}$$

We obtain:

$$f(3, 3) = 18,$$
$$f(1, 1) = 2.$$

Comparing all the obtained values of the given function, we conclude that

$$f_{greatest} = 18 \quad \text{at} \quad (3, 3),$$

$$f_{least} = 2 \quad \text{at} \quad (1, 1).$$