# Chapter 3

# Optimality conditions

In this chapter we present optimality conditions for the problem

$$\min_{x \in S} f(x)$$

with $S \subseteq R^n$.

A complete treatment can be found in Chapter 1 of [1]

Optimality conditions play a fundamental role not only from a theoretical point of view but also from algorithmic perspective.

In particular we consider the cases

- $S := \mathbb{R}^n$ (unconstrained problems)
- $S$ is a convex set,
- $S$ is a polyhedron.

We start defining a *descent* direction and a *feasible* direction.

## 3.1 Descent Direction

**Definition 3.2 (Descent Direction)** *Let $f : R^n \to R$ and $x \in R^n$. A vector $d \in R^n$, $d \neq 0$ is a descent direction for $f$ in $x$ if there exists a $t^d_{\max} > 0$ such that*

$$f(x+td) < f(x), \quad \textit{for all} \quad t \in (0, t^d_{\max}).$$

The definition of descent direction is a local definition that holds in a neighborhood of the point.

if $f$ is a linear function $f(x) = c^T x$, we get easily

$$c^T(x+td) = c^T x + tc^T d < c^T x \quad \text{for all } t > 0$$

If $f$ is linear, i.e. $f(x) := c^T x$ $d$ is a descent direction if and only if $c^T d < 0$ holds (in any point $x \in \Re^n$).

We derive characterization of descent direction for continuously differentiable functions in section 3.8.

## 3.3 Feasible direction

A vector $d$ is a feasible direction in $\bar{x} \in S$ if it is possible to make "small" movement from $\bar{x}$ along $d$ remaining in $S$. More formally

**Definition 3.4 (Feasible direction)** *Let $S \subseteq R^n$ and $\bar{x} \in S$. A vector $d \neq 0$ is a feasible direction in $\bar{x}$ if there exists $t^f_{\max} > 0$ such that*

$$\bar{x} + td \in S, \quad \textit{for all} \quad t \in [0, t^f_{\max}].$$

**Example 3.5** Consider the polyhedron

$$S = \{x \in R^2 : \ x_1 + x_2 \leq 3, \ x_1 \geq 1, \ x_2 \geq 0\}$$

in Figure 4.1 and the feasible piont $\bar{x} = (1, \ 1)$ (the red point in the figure).

The direction $d = (1, \ 0)$ is feasible in $\bar{x}$. Indeed we have

$$\bar{x} + td = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1+t \\ 1 \end{pmatrix}$$

Putting into the constraints:

1. $x_1 + x_2 = (1+t) + 1 = 2 + t$ which is $\leq 3$ for $t \leq 1$.
2. $x_1 = 1 + t$ which is $\geq 1$ for $t \geq 0$;
3. $x_2 = 1 > 0$.

Hence for $t \in [0, 1]$ ($t^f_{\max} = 1$) $\bar{x} + td \in S$.

The opposite direction $-d = (-1, \ 0)$ is not feasible in $\bar{x}$. We get

$$y = \bar{x} + t(-d) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1-t \\ 1 \end{pmatrix}$$

Putting into the constraints:

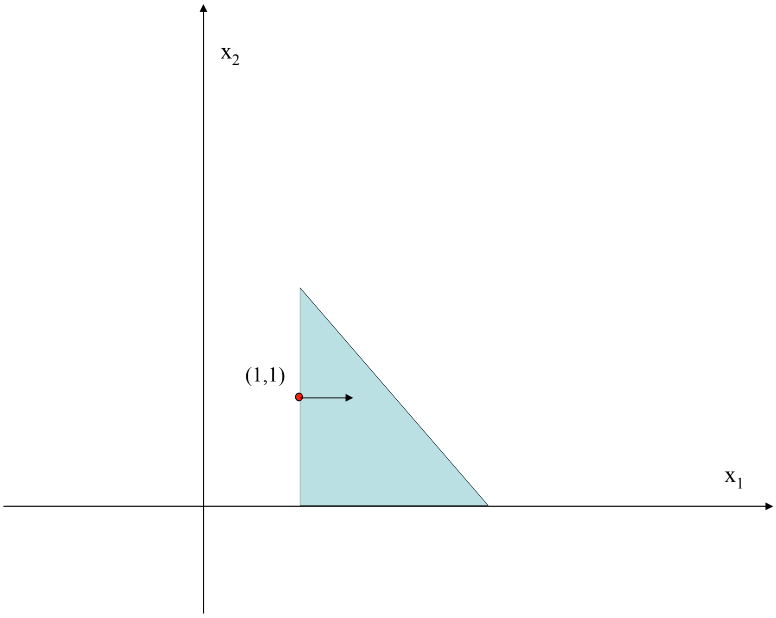


Figure 3.1: Polyhedron Example 3.5.

1. $x_1 + x_2 = (1-t)+1 = 2-t$ which is $\leq 3$ for $t \geq -1$.
2. $x_1 = 1-t$ which is $\geq 1$ for $t \leq 0$;
3. $x_2 = 1 > 0$ for any $t$.

Thus it doen NOT exist $t^f_{\max} > 0$ such that $y$ belongs to $S$.

## 3.6 Optimality condition for constrained problems

Consider the problem

$$\begin{array}{ll} \min & f(x) \\ & x \in S \end{array} \tag{3.1}$$

As an immediate consequence of the definition 3.2 of descent direction and defintion 3.4 of feasibile direction we get the following trivial result.

**Theorem 3.7 (Necessary condition of local minimizer)** *Let $x^* \in S$ be a local minimizer of problem 3.1, then there exists NO feasible direction in $x^*$ which is also descent for $f$.*

**Proof.** By contradiction if such a direction exists $d$ in $x^*$, then in any neighborhood of $x^*$ it is possible to find a sufficiently small $t > 0$ such that $x^* + td \in S$ and $f(x^* + td) < f(x^*)$, which contradicts the assumption that $x^*$ is a local minimizer.

Such condition may be useless when there are points in the feasible region such that there are not feasibile directions at all. This may happen e.g. when there are nonlinear equality constraints

As an example consider the problem

$$\begin{aligned} \min \quad & x_2 \\ & x_1^2 + x_2^2 = 1 \end{aligned}$$

which is represented in the picture 3.2. The unique gloabl solution is $(0, \; -1)^T$. In any point

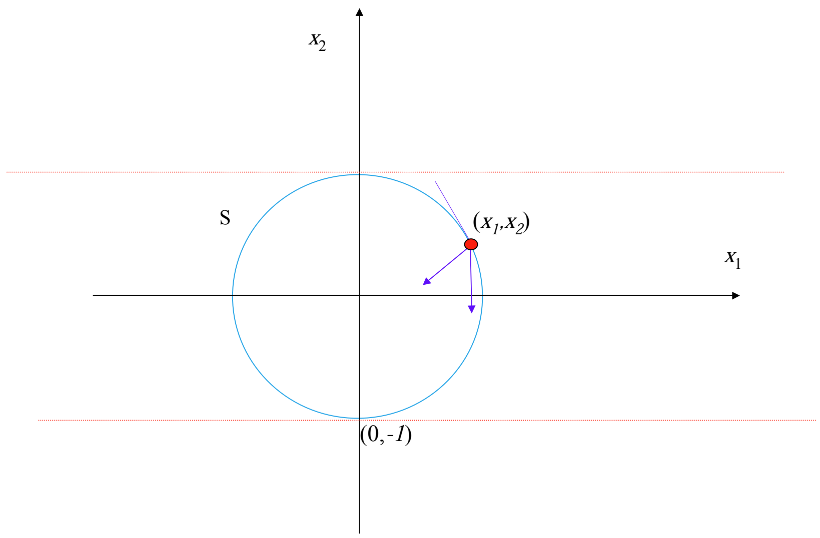


Figure 3.2: In any fesible point there is no feasibile direction.

$(x_1, x_2) \in S$ There is no feasibile direction so that the condition expressed by theorem 3.7 is satisfied for all points in $S$ and we have no additional information.

## 3.8 Characterization of descent directions

We have a sufficient condition

**Theorem 3.9 (First order descent condition)** *Let $f$ be continuously differentiable in a neighborhood of $x \in R^n$ and let $d \neq 0 \in R^n$. If*

$$\nabla f(x)^T d < 0,$$

*then the $d$ is a descent direction for $f$ in $x$.*

**Dimostrazione.** For any continuously differentiable function we have

$$f(x+td) = f(x) + t\nabla f(x)^T d + \alpha(x,t) \tag{3.2}$$

where $\alpha(x,t)$ satisfies $\lim_{t \to 0} \frac{\alpha(x,t)}{t} = 0$. For sufficiently small values of $t$ we have

$$f(x+td) - f(x) \simeq t\nabla f(x)^T d;$$

If $\nabla f(x)^T d < 0$ we get $f(x+td) - f(x) < 0$.

Since $\nabla f(x)^T d = \|\nabla f(x)\| \|d\| \cos\theta$ where $\theta$ is the angle between $\nabla f(x)$ and $d$, geometrically $\nabla f(x)^T d < 0$ means that $d$ is a descent direction if forms an obtuse angle with the gradient of $f$ in $x$.

Let $f$ be a continuously differentiable function and $d \neq 0 \in R^n$. if the angle $\theta$ between $\nabla f(x)$ and $d$ satisfies

$$\theta > 90^o$$

then $d$ is a descent direction for $f$ in $x$.

Among descent directions the most relevant is the *antigradient* $d = -\nabla f(x)$. If $\nabla f(x) \neq 0$, this direction $d = -\nabla f(x)$ is always a descent direction indeed we get

$$\nabla f(x)^T d = -\nabla f(x)^T \nabla f(x) = -\|\nabla f(x)\|^2 < 0.$$

A graphical representation is in figure 3.3.

Analogous condition for an uphill direction.

Let $f$ be continuously differentiable in a neighborhood of $x \in R^n$ and let $d \neq 0 \in R^n$. If

$$\nabla f(x)^T d > 0,$$

then the $d$ is an uphill direction for $f$ in $x$.

When $\nabla f(x) \neq 0$, the gradient direction $d = \nabla f(x)$ is always uphill $x$.

When $f = c^T x$ (linear function), we have $\nabla f(x) = c$, hence (3.2) holds with $\alpha(x,t) := 0$ and the condition of Theorem 3.9 becomes necessary too.

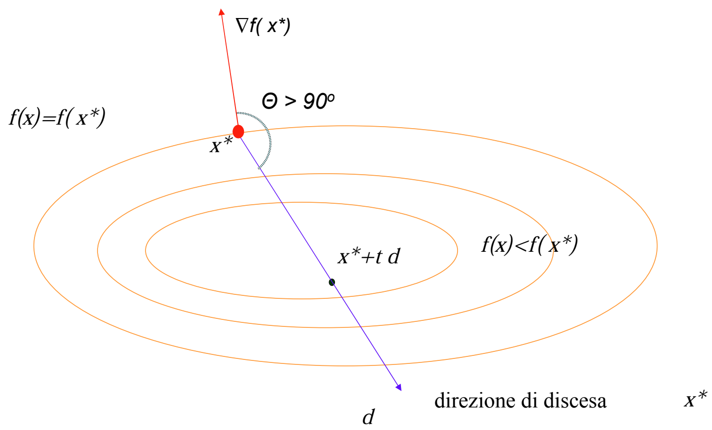


Figure 3.3: A descent direction $d$ in $\Re^2$.

Let $f(x) := c^T x$ then

1. $d$ is a descent direction for $f$ in $x$ if and only if $\nabla f(x)^T d = c^T d < 0$;
2. $d$ is an uphill direction in $x$ if and only if $\nabla f(x)^T d = c^T d > 0$;
3. $d$ è a direction along where the function is maintained constant compared to the value at $x$ if and only if $\nabla f(x)^T d = c^T d = 0$.

For generic $f$ the condition is only sufficient because there may exist descent directions with $\nabla f(x)^T d = 0$.

Let $f$ be continuously differentiable in a neighborhood of $x \in R^n$ and let $d \neq 0 \in R^n$.

1. if $\nabla f(x)^T d < 0$ then $d$ is a descent direction for $f$ in $x$;
2. if $\nabla f(x)^T d > 0$ $d$ is a uphill direction for $f$ in $x$;
3. if $\nabla f(x)^T d = 0$ it is non possible to state if $d$ is descent, uphill or constant value direction for $f$ in $x$.

COndition of Theorem 3.9 becomes necessary under convexity.

**Proposition 3.10** *If $f$ is convex then $d$ is a descent direction for $f$ in $x$ if and only if*

$$\nabla f(x)^T d < 0.$$

If $f$ is twice continuosly differentiable we can use second derivatives for further characterization.

**Definition 3.11 (Negative curvature direction)**

*Let $f : R^n \to R$ $f$ be twice continuously differentiable in a neighborhood of $x \in R^n$ and let $d \neq 0 \in R^n$. If $d^T \nabla^2 f(x) d < 0$, the direction $d \in R^n$ is called a negative curvature direction.*

**Theorem 3.12 (Second order descent condition)** *Let $f : R^n \to R$ $f$ be twice continuously differentiable in a neighborhood of $x \in R^n$ and let $d \neq 0 \in R^n$. Assume that*

$$\nabla f(x)^T d = 0,$$

*and $d$ satisfies $d^T \nabla^2 f(x) d < 0$. Then $d$ is a descent direction for $f$ in $x$.*

**Proof.** Since $f$ is twice continuously differentiable we have:

$$f(x+td) = f(x) + t\nabla f(x)^T d + \frac{1}{2}t^2 d^T \nabla^2 f(x) d + \beta(x,td)$$

where $\beta(x,td)/t^2 \to 0$. Since $\nabla f(x)^T d = 0$, we obtain:

$$\frac{f(x+td) - f(x)}{t^2} = \frac{1}{2} d^T \nabla^2 f(x) d + \frac{\beta(x,td)}{t^2}$$

since

$$\lim_{t \to 0} \frac{\beta(x,td)}{t^2} = 0,$$

for sufficiently small values of $t$ we get $f(x+td) - f(x) < 0$.

## 3.13 First ($1^{st}$) order optimality conditions

Let $x^* \in S$ be a local minimizer of problem 3.1, then there exists NO feasible direction $d \in R^n$ in $x^*$ such that $\nabla f(x^*)^T d < 0$.

Alternatively

**Theorem 3.14 (First order necessary condition)** *Let $x^* \in S$ be a local minimizer of problem 3.1, then we have*

$$\nabla f(x^*)^T d \geq 0 \quad \textit{for all the feasible directions } d \in R^n \textit{ in } x^*.$$

Since $\nabla f(x^*)^T d = \|\nabla f(x^*)\| \|d\| \cos\theta$ where $\theta$ is the angle between $\nabla f(x^*)$ and $d$, geometrically we get the condition $\nabla f(x^*)^T d \geq 0$ states that $0^o \leq \theta \leq 90^o$ for all feasible $d$.

Let $x^* \in S$ be a local minimizer of problem 3.1, then for all feasible $d$, the angle $\theta$ formed with $\nabla f(x^*)$ satisfies

$$0^o \leq \theta \leq 90^o.$$

**Example 3.15** Consider the problem

$$\begin{array}{ll} \min & -x_1^2 - x_2^2 \\ & -1 \leq x_1 \leq 1 \\ & -1 \leq x_2 \leq 1 \end{array}$$

represented in the figure 3.4. Let consider the minimum pouint $x^* = (1,\ 1)^T$. Check that the first

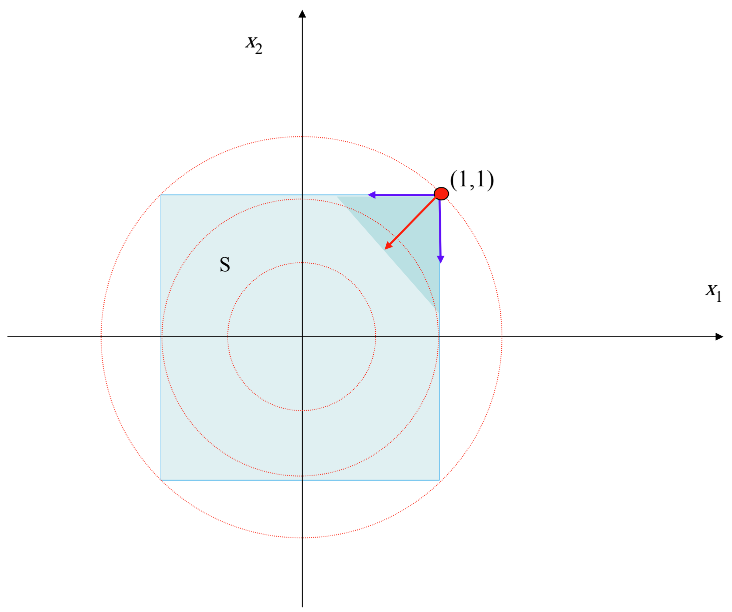


Figure 3.4: Illustrazione Esempio 3.15.

order optimality condition is satisfied.

We can also use the second order characterization of descent direction of Theorem 3.12 to get second order optimality conditions.

**Theorem 3.16 (Second order necessary condition)** *Let $x^* \in S$ be a local minimizer of problem 3.1, then for all the feasible directions in $x^*$ such that $\nabla f(x^*)^T d = 0$ we have*

$$d^T \nabla^2 f(x^*)^T d \geq 0 \quad \textit{for all the feasible directions } d \in R^n \textit{ in } x^*.$$

In the next sections we characterize the feasible directions in the cases of our interest.

## 3.17 Unconstrained optimization

Unconstrained problems are a special class of nonlinear programming problems when $S := R^n$. Thus the problem is

$$\min_{x \in \Re^n} f(x), \tag{3.3}$$

with $f : \Re^n \to \Re$ nonlinear and continuously differentiable.

First of al we derive *existence condition* of a solution. To this aim let $\hat{x} \in \mathbb{R}^n$ and $f(\hat{x})$ the corresponding function value. Let define the non empty set which is called *level set* of the function $f$

$$\mathscr{L}(\hat{x}) = \{x \in \Re^n : \ f(x) \leq f(\hat{x})\}.$$

We can consider the "constrained" problem

$$\begin{array}{ll} \min & f(x) \\ & x \in \mathscr{L}(\hat{x}) \end{array} \tag{3.4}$$

which is equivalent to problem (3.3). Indeed if the global solution $x^*$ of (3.3), if it exists, belongs to the set $\mathscr{L}(\hat{x})$.
We can apply the Weierstrass theorem to problem (3.4).

**Existence condition for unconstrained problems.**
if $\hat{x} \in \Re^n$ exists such that $\mathscr{L}(\hat{x})$ is compact (bounded and closed) then problem (3.3) admits a global solution.

We derive optimality condition by simplifying conditions of the preceding section.

We observe that in the unconstrained case, feasibility does not play any role because it is possible to move along any vector $d \in R^n$. hence we get the following results.

**First order unconstrained necessary condition.**
Let $x^* \in S$ be a local minimizer of problem 3.1, then there exists NO descent direction for $f$ in $x^*$.

Condition above is no more than the definition of local solution. However we can use characterization of descent direction.

Let $x^* \in S$ be a local minimizer of problem 3.1, then there exists NO direction $d \in R^n$ suxh that $\nabla f(x^*)^T d < 0$.

We get the following condition.

**Theorem 3.18 (First order unconstrained necessary condition)** *Let $f : R^n \to R$ $f$ be continuously differentiable on $R^n$ and let $x^\star \in R^n$. If $x^* \in S$ in an unconstrained local minimizer of problem 3.1 then*

$$\nabla f(x^\star) = 0.$$

**Proof.** (By contradiction). Suppose that $\nabla f(x^*) \neq 0$. Taking $d = -\nabla f(x^*) \neq 0$ yields $\nabla f(x^*)^T d = -\nabla f(x^*)^T \nabla f(x^*) < 0$, so $d$ is a descent direction which contradicts the assumption that $x^*$ is a local minimizer.

**Definition 3.19 (Stationary point)** *Let $f : R^n \to R$ $f$ be continuously differentiable on $R^n$. A point $\bar{x}$ satisfying $\nabla f(\bar{x}) = 0$ is called* **stationary point** *of $f$.*

Candidates to be local minimizers can be found by solving the nonlinear square system $\nabla f(x) = 0$.

**Example 3.20** Consider the function

$$f(x_1, x_2) = {x_1}^4 + {x_2}^4 - 3x_1x_2.$$

Find the candidates to be minimizers applying the first order necessary condition.

When $\nabla f(x^*) \neq 0$ we have also an uphill direction $d = \nabla f(x^*)$ in $x^*$, so that $\nabla f(x^*) = 0$ is a necessary condition of local maximizer too.

Let $f : R^n \to R$ $f$ be continuously differentiable on $R^n$. and let $x^\star \in R^n$ such that $\nabla f(x^\star) = 0$ then one of the following assertions is true

1. $x^\star$ is a local minimizer;
2. $x^\star$ is a local maximizer;
3. $x^\star$ is a saddle point (in $x^\star$ there are both descent and uphill directions). See Figure 3.5.

When the function is convex the condition becomes also sufficient, so that we get.

**Theorem 3.21 (Necessary and sufficient condition convex function)** *Let $f : R^n \to R$ $f$ be continuously differentiable on $R^n$. Assume that $f$ is convex. The point $x^\star$ id a global minimizer of $f$ on $R^n$ if and only if $\nabla f(x^\star) = 0$. Furthermore if $f$ is strictly convex on $R^n$ and $\nabla f(x^\star) = 0$, the $x^\star$ is the unique global mimizer of $f$.*

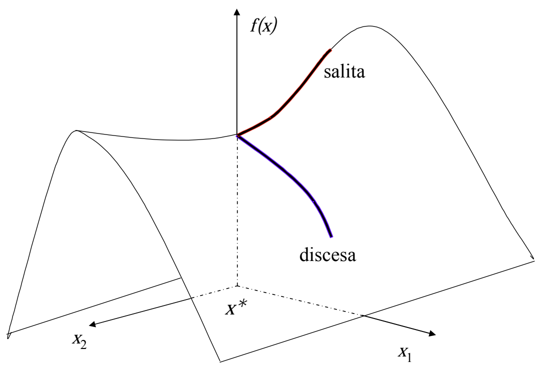


Figure 3.5: Esempio di punto di sella in $\Re^2$.

We can also use characterization of the secondo order given by Theorem 3.12.

**Theorem 3.22 (Second order unconstrained necessary condition.)** *Let $f : R^n \to R$ $f$ be twice continuously differentiable on $R^n$. Let $x^* \in S$ be a local minimizer of problem 3.1, then for any $d$ such that $\nabla f(x^*)^T d = 0$ it holds*

$$d^T \nabla^2 f(x^*)^T d \geq 0 \quad \textit{for all } d \in R^n$$

**Theorem 3.23 (Second order unconstrained necessary condition.)** *Let $f : R^n \to R$ $f$ be twice continuously differentiable in a neighborhood of $x^*$. Let $x^* \in S$ be a local minimizer of problem 3.1, then we have:*

(a) $\nabla f(x^\star) = 0$;

(b) $\nabla^2 f(x^*)$ *is positive semidefinite* $\quad y' \nabla^2 f(x^\star) y \geq 0$, for all $y \in R^n$.

When $\nabla^2 f(x^\star)$ is positive definite, we get a stronger result.

**Theorem 3.24 (Second order sufficient condition)** *Let $f : R^n \to R$ $f$ be twice continuously differentiable in a neighborhood of $x^*$.*

*Assume that the following conditions hold:*

(a) $\nabla f(x^\star) = 0$

(b) *the hessian matrix is positive definite in $x^\star$, i.e.:*

$$y^T \nabla^2 f(x^\star) y > 0 \quad \text{for all} \quad y \in R^n, \quad y \neq 0.$$

*Then $x^\star$ is a strict local minimizer.*

From the preceding results we can get easily optimality conditions for a point to be a mazimizer.

In particular, the condition $\nabla f(x^\star) = 0$ is a necessary condition for $x^\star$ to be a local maximizer, too.

A second order condition for local maximizer is $\nabla f(x^\star) = 0$ and $\nabla^2 f(x^\star)$ *negative semidefinite*. A secondo order sufficient condition for $x^\star$ to be a strict local maximizer is $\nabla f(x^\star) = 0$ and $\nabla^2 f(x^\star)$ *negative definite*.

When $\nabla f(x^\star) = 0$ and the hessian $\nabla^2 f(x^\star)$ is *indefinite* (there exists vectors $d$ such that $d^T \nabla^2 f(x^\star) d > 0$ and others such that $d^T \nabla^2 f(x^\star) d < 0$) then $x^\star$ is neither a local minimizer nor a local maximizer and $x^\star$ is usually called *saddle point*.

A saddle point is a stationary point where there are *descent* direction (when $d^T \nabla^2 f(x^\star) d < 0$) and *uphill* directions (when $d^T \nabla^2 f(x^\star) d > 0$).

If $\nabla f(x^\star) = 0$ and $\nabla^2 f(x^\star)$ id only *semidefinite* (negative o positive) we cannot conclude on the nature of $x^\star$ without further information.

### 3.24.1 Quadratic functions

A special class of nonlinear functions are the quadratic

$$f(x) = \frac{1}{2} x^T Q x + c^T x$$

with $Q$ symmetric $n \times n$ and $c \in \Re^n$.

We get the following result.

**Theorem 3.25 (Minimization of a quadratic function)** *Let $q(x) = \frac{1}{2} x^T Q x + c^T x$, with $Q$ symmetric and $c \in R^n$. Then:*

*(i) $q(x)$ admits a global solution if and only if $Q \succeq 0$ and $x^\star$ exists such that $Qx^\star + c = 0$;*

*(ii)* $q(x)$ *admits a unique global solution if and only if* $Q \succ 0$*;*

*(iii) if* $Q$ *is positive semidefinite any point* $x^\star$ *such that* $Qx^\star + c = 0$ *is a global minimizer of* $q(x)$.

**Proof.** We have that $\nabla q(x) = Qx + c$ e $\nabla^2 f(x) = Q$. The function $q$ is convex if and only if $Q$ is positive semidefinite ($Q \succeq 0$), and strictly convex if and only if $Q$ is positive definite ($Q \succ 0$).
Let $x = x^\star + s$ we get:

$$q(x) = q(x^\star + s) = q(x^\star) + (Qx^\star + c)^T s + \frac{1}{2} s^T Q s. \tag{3.5}$$

Assume that $Qx^\star + c = 0$ and $Q \succeq 0$ ; from (3.5) we get $q(x) \geq q(x^\star)$ for all $x \in R^n$.
ON the otherhand, if $q$ admits a minimizer $x^\star$, from Theorem 3.18 we have $\nabla q(x^\star) = 0$ and $q(x) \geq q(x^\star)$ for all $x \in \Re^n$. From (3.5) we get $0 \leq q(x) - q(x^\star) = \frac{1}{2} s^T Q s$ with $s = x - x^\star$, namely $Q \succeq 0$. So that we have proved (i).
Condition (ii) follows from (i) and $s^T Q s > 0$ for all $s$ if and only if $Q$ è definita positiva.
Finally, la (iii) follows again from (3.5) because for each $x^\star$ such that $\nabla q(x^\star) = 0$ we have $q(x) \geq q(x^\star)$.

See next example.

**Example 3.26** Consider the function

$$q(x_1, x_2) = \frac{1}{2}(\alpha x_1^2 + \beta x_2^2) - x_1.$$

Study the existence and the nature of stationary point varying the parameters $\alpha$ e $\beta$. Gradient and hessian of $q$:

$$\nabla q = \begin{pmatrix} \alpha x_1 - 1 \\ \beta x_2 \end{pmatrix} \qquad \nabla^2 q = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$$

When $\alpha > 0$ and $\beta > 0$, there exists a unique solution to the system $\nabla q = 0$ which is $\left(\frac{1}{\alpha}, 0\right)^T$ and $Q \succ 0$. Hence is the unique global minimizer.
When $\alpha = 0$ and $\beta \in \mathbb{R}$ there exists no solution to the system $\nabla q = 0$. If $\beta \geq 0$ $Q \succeq 0$ and the function is convex
When $\alpha > 0$ and $\beta = 0$, we have an infinite number of solution to the system $\nabla q = 0$ and $\left(\frac{1}{\alpha}, \xi\right)^T$ with $\xi \in \mathbb{R}$ Furthermore $\nabla^2 q \succeq 0$ hence all these point are global minimizers.

When $\alpha < 0$ and $\beta > 0$ we have a unique solution $(\frac{1}{\alpha}, 0)$ but $Q$ is inedfinite. It is a saddle point.

When $\alpha < 0$ and $\beta < 0$, There is aunique solution $\left(\frac{1}{\alpha}, 0\right)^T$ with $Q$ negative definite.it is a global maximizer.

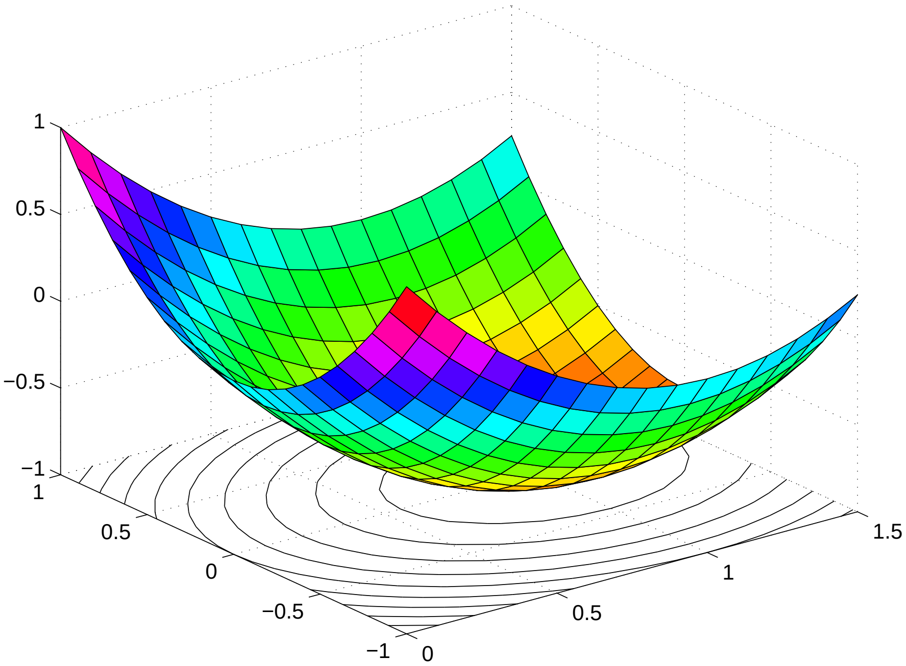


Figure 3.6: Grafico di $q(x_1, x_2)$ per $\alpha = \beta = 1$.

## 3.27 Algorithmic use of unconstrained optimality condition

Consider the unconstrained problem

$$\min_{x \in R^n} f(x) \tag{3.6}$$

where $f : R^n \to R$ and assume:

**Assumption 3.28** *The function $f : R^n \to R$ is continuously differentiable and the level set $\mathscr{L}_{x^0}$ is compact for some $x^0 \in R^n$.*

Algorithms for the solution of problem (3.6) are usually able to find stationary points of $f$, namely points belonging to the target set

$$\Omega := \{\omega \in R^n : \nabla f(\omega) = 0\}.$$

In general it is possible to guarantee to reach points with objective function value less than the initial value, when $x^0$ is not stationary and this satisfactory for many applications.
When the objective function is convex, finding a stationary point solve the problem because every stationary point is a global solution.

We consider algorithms with the following basic scheme:

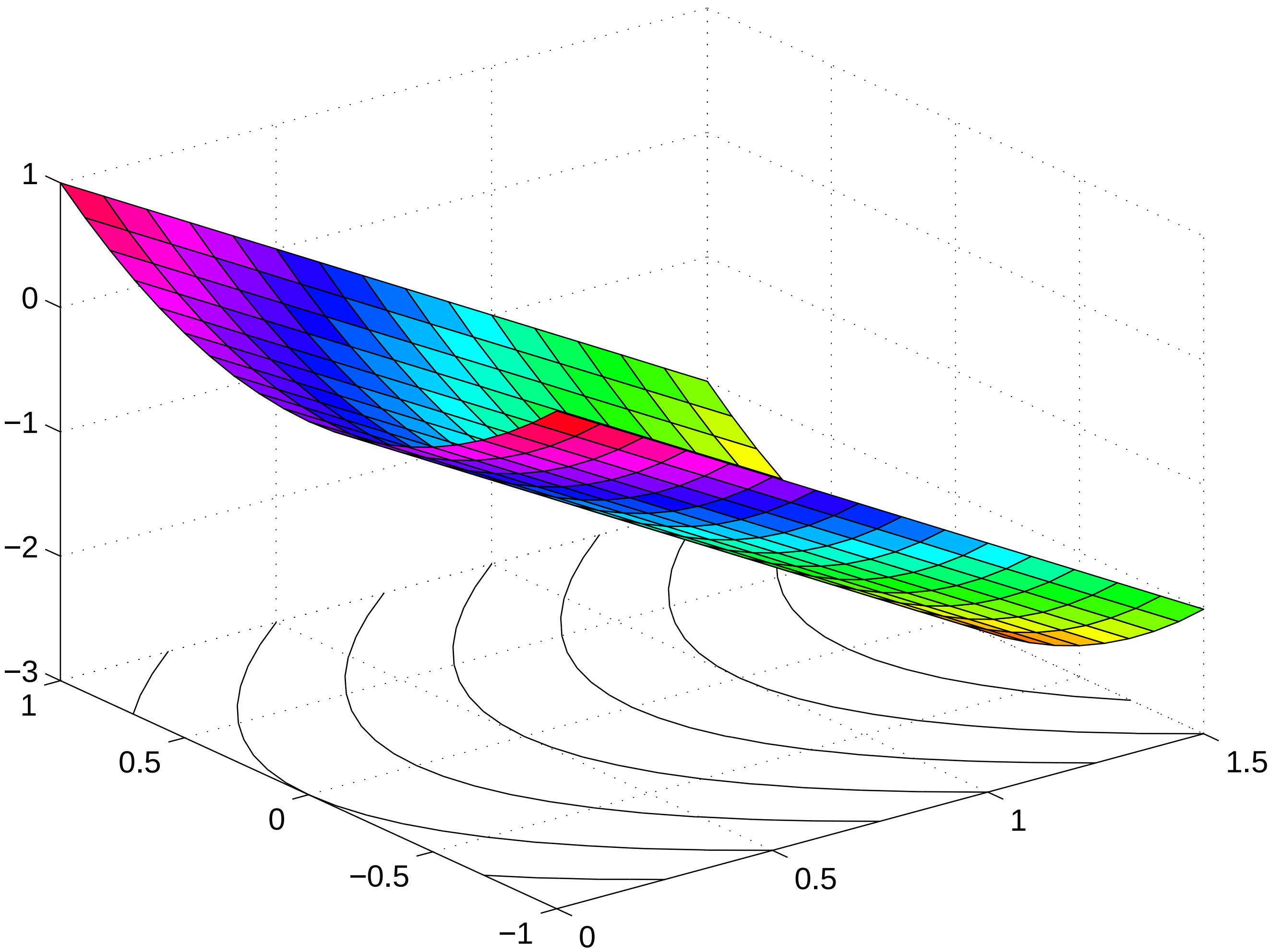


Figure 3.7: Plot of $q(x_1,x_2)$ when $\alpha = 0$ $\beta = 1$.

**Unconstrained optimization scheme**

1. Fixed a *starting point* $x^0 \in R^n$ and set $k = 0$.
2. If $x^k \in \Omega$ stop.
3. Evaluate a *descent direction* $d^k \in R^n$.
4. Evaluate a *stepsize* $\alpha^k \in R$ along $d^k$;
5. Find a new point $x^{k+1} = x^k + \alpha^k d^k$. Set $k = k+1$ and go to Step 2.

**Stopping criteria.** At Step 2, algorithm checks if $x^k$ belongs to the set $\Omega$ which is equivalent to check if $\nabla f(x^k) = 0$. In practice, when using finite arithmetic we need to specify *a stopping*

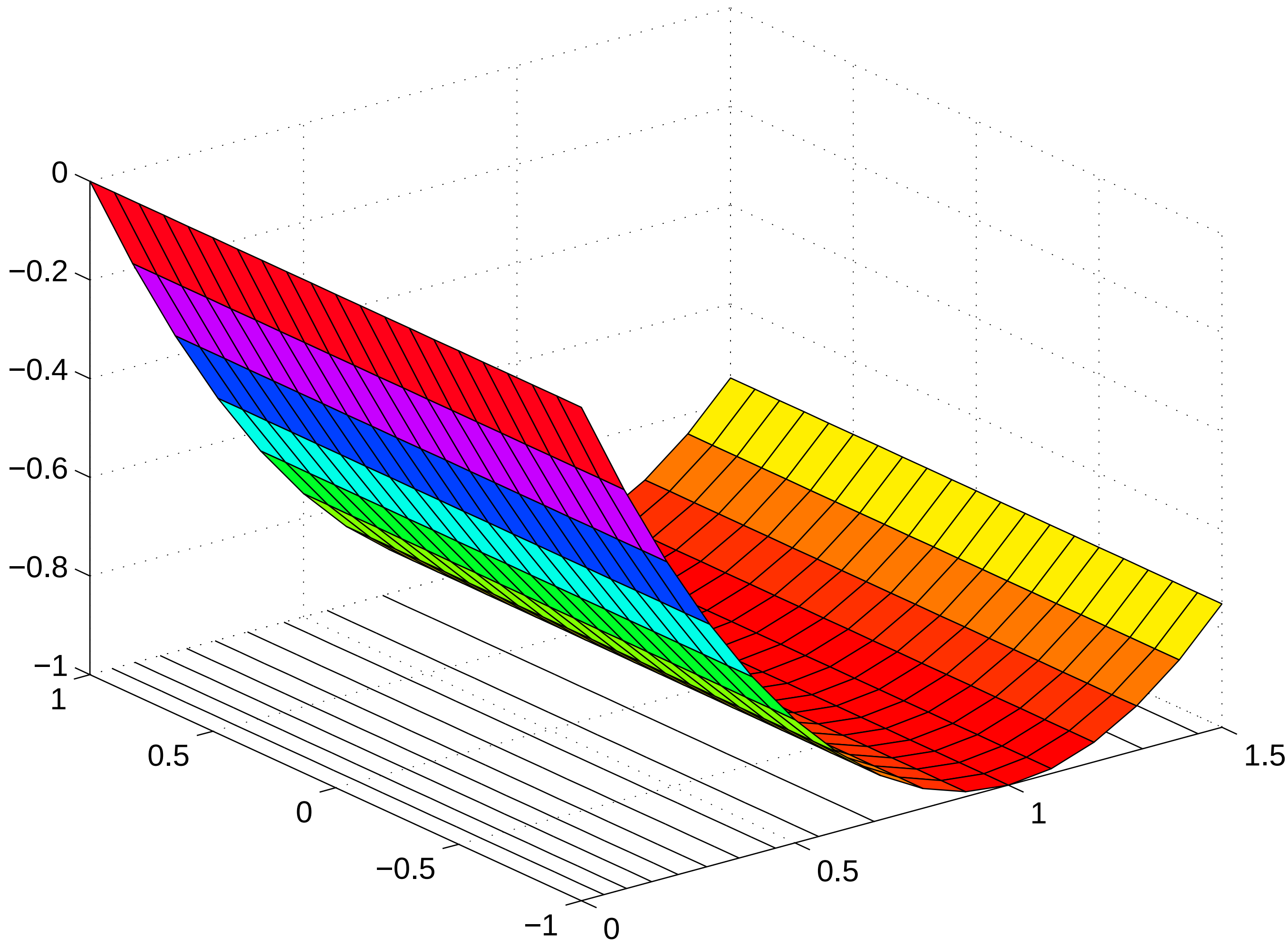


Figure 3.8: Plot of $q(x_1,x_2)$ when $\alpha = 1$ $\beta = 0$.

*criterion*. A first criterion is

$$\|\nabla f(x^k)\| \leq \varepsilon \tag{3.7}$$

with $\varepsilon > 0$ is a sufficiently small values ($\varepsilon \in [10^{-8}, 10^{-3}]$). We can use different norms among the most used we have $\|\cdot\|_2$ and $\|\cdot\|_\infty$. In particular $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$ and $\|x\|_\infty = \max_{1\leq i\leq n}\{|x_i|\}$.

**Choice of the direction.** Among the possibile unconstrained methods we have *gradient method* (or *steepest descent method*)which uses the direction $d^k = -\nabla f(x^k)$;

**Choice of the stepsixe.** The choise of the stepsize $\alpha^k > 0$ is called *linesearch* and it is performed by evaluatig the objective function along the direction $d^k$. The simplest scheme requires to find the minimum of $f(x^k + \alpha d^k)$ with respect to $\alpha$ (*exact linesearch*) or to find an acceptable interval of values for $\alpha^k$ for getting convergence. (*inexact linesearch*).

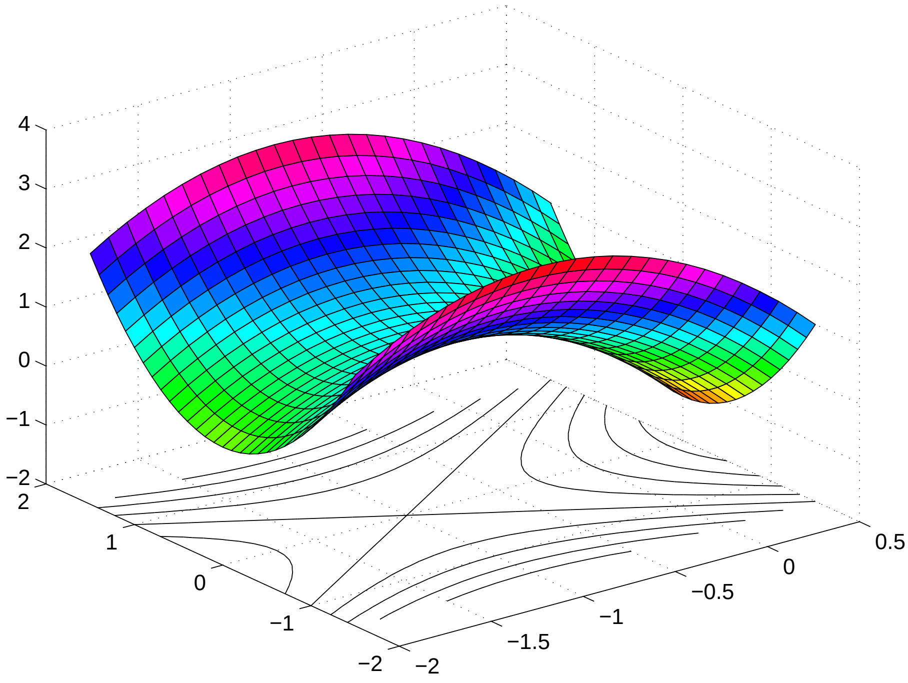


Figure 3.9: Plot of $q(x_1, x_2)$ when $\alpha = -1$ $\beta = 1$.

## 3.29 Unconstrained problems: examples

### 3.29.1 Price discrimination

Price discrimination is a microeconomic pricing strategy where identical or largely similar goods or services are transacted at different prices by the same provider in different markets. Price differentiation essentially relies on the variation in the customers' willingness to pay and in the elasticity of their demand. As a matter of example consider two different market (national and foreign markets) each with its own demand function. Denote by $x_i$ the supply on the each market $i = 1, 2$ and with $P_i = f_i(x_i) = a_i - m_i x$ the inverse demand function on market $i$ (price that customer are willing to pay on market $i$).

Cost production is proportional with constant $c$ to the full supply $x_1 + x_2$.

What quantities will the monopolist sell in the two markets to maximize the revenue?

We get

$$min \quad x^T Mx - (a-c)^T x$$

$$x \geq 0.$$

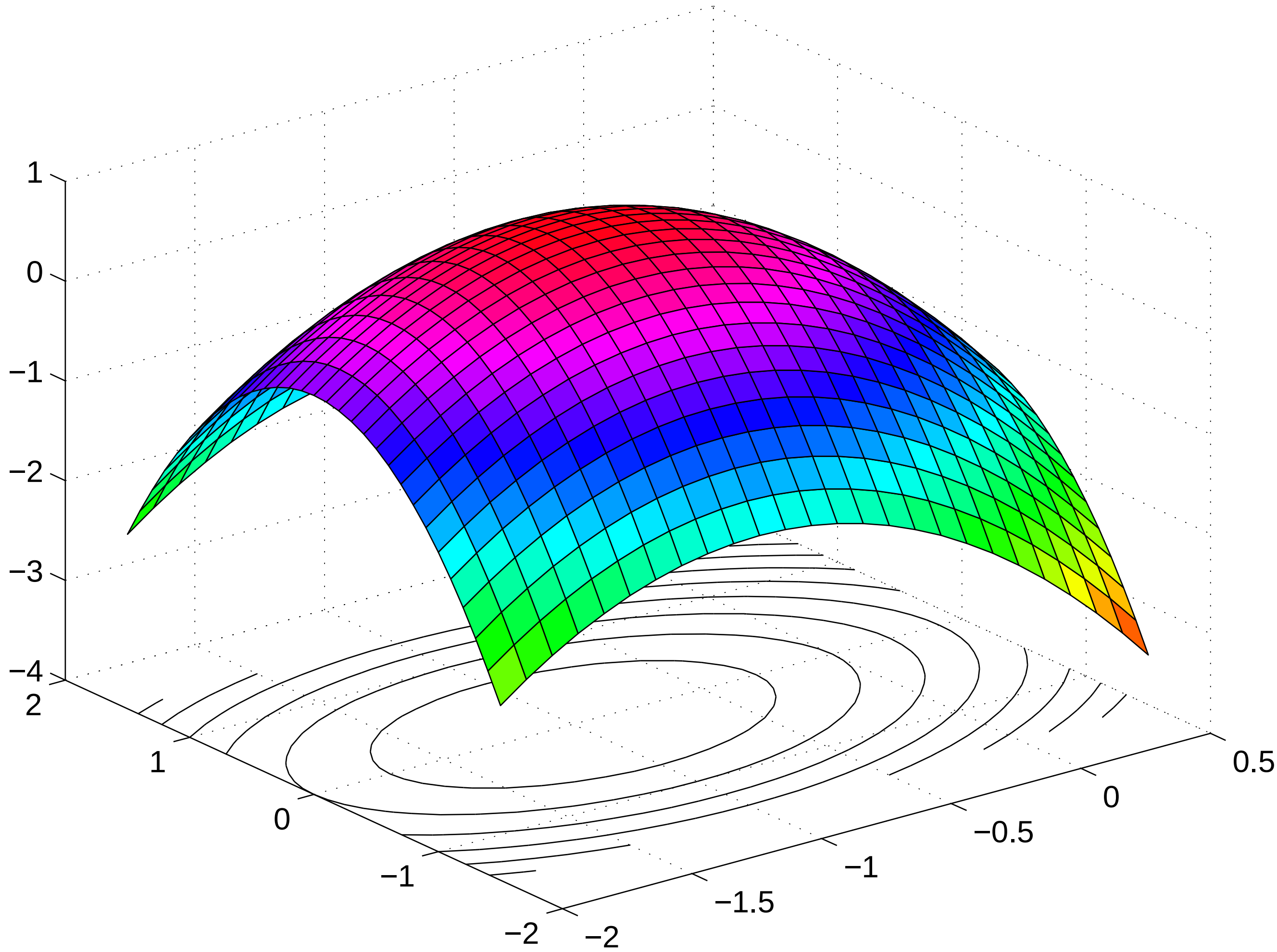


Figure 3.10: Plot of $q(x_1, x_2)$ when $\alpha = \beta = -1$.

# Bibliography

[1] Bertsekas, D.P. (2016), Nonlinear Programming 3rd edition, Athena