

Nonlinear Optimization: A Brief Overview



Renato De Leone

Abstract In this chapter some of the most important results for unconstrained and constrained optimization problems are discussed. This chapter does not claim to cover all the aspects in nonlinear optimization that will require more than one complete book. We decided, instead, to concentrate our attention on few fundamental topics that are also at the basis of the new results in nonlinear optimization using *grossone* introduced in the successive chapters.

1 Introduction

The aim of this chapter is to introduce the reader to the most important results and algorithms in unconstrained and constrained optimization. However, it would be impossible to discuss the wide range of results in this area. We decided to concentrate the attention on few fundamental topics that are also at the basis of the new results obtained using ① and introduced in the successive chapters. The interested reader can refer to various classical and recent books to deepen his/her knowledge. We also omitted almost all proofs of the results, but complete references are always provided.

After introducing the concepts of convex set and functions, in Sect. 3 optimality conditions for unconstrained optimization are presented as well as the most important algorithmic techniques: gradient method, conjugate gradient method, Newton's and Quasi-Newton's methods. In the successive Sect. 4 we concentrate the attention on optimality conditions for constrained optimiza-

R. De Leone (✉)

School of Science and Technology, University of Camerino, Camerino (MC), Italy
e-mail: renato.deleone@unicam.it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

37

Y. D. Sergeyev and R. De Leone (eds.), *Numerical Infinities and Infinitesimals in Optimization*, Emergence, Complexity and Computation 43,
https://doi.org/10.1007/978-3-030-93642-6_2

tion and the construction of the dual of nonlinear optimization problems. Finally, some important algorithms for constrained nonlinear optimization problems are presented.

Three important aspects in optimization are not presented here: global optimization, multi-objective optimization, and non-smooth optimization. In this chapter we only discuss first and second order optimality conditions for local optima, and all algorithms will compute a local minimum, or a stationary point (for the unconstrained case) or a Karush–Kuhn–Tucker point (for the constrained case). Global optimization is an important area, theoretically and practically, for which several different approaches have been suggested: space covering methods, trajectory methods, random sampling, random search, etc. A basic reference on various aspects of global optimization is [40], while practical applications are discussed in [51]. Finally, a comprehensive archive of online information can be found at the web page <http://www.globaloptimization.org/>.

Many optimization problems are multi-objective in nature, and different techniques have been proposed to deal with this important aspect of optimization. In general, due to the presence of conflicting objectives, there is no single solution that simultaneously optimizes all the objectives, and, hence, the aim is to determine non-dominated, Pareto optimal solutions [16, 23]. Two general approaches to multiple-objective optimization are present: combine the individual objective functions into a single function, or move all but one objective to the constraint set. More recently algorithms based on different meta-heuristics have also been proposed in literature [41].

Non-smooth optimization problems arise in many important practical applications. Here it is assumed that the function is continuous, but not differentiable. The methods for non-smooth optimization can be roughly divided into two main classes: subgradient methods and bundle methods. Both classes of methods are based on the assumption that, at each point, the objective function value and one subgradient can be computed. A classical book on this topic is [43]. A survey on different numerical methods for non-smooth optimization and the most recent developments presented in [5]. Finally, a recent compact survey on non-smooth optimization can be found in [28].

We briefly describe our notation now. All vectors are column vectors and will be indicated with lower case Latin letter (x, y, \dots). Subscripts indicate components of a vector, while superscripts are used to identify different vectors. Matrices will be indicated with upper case roman letter (A, B, \dots). For a $m \times n$ matrix A , A_{ij} is the element in the i th row, j th column, $A_{.j}$ is the j -th column of A , while A_i is its i -th row. The set of real numbers and the set of nonnegative real numbers will be denoted by \mathbb{R} and \mathbb{R}_+ respectively. The space of the n -dimensional vectors with real components will be indicated

by \mathbb{R}^n and \mathbb{R}_+^n is an abbreviation for the nonnegative orthant in \mathbb{R}^n . The symbol $\|x\|$ indicates the norm of the vector x . In particular, the Euclidean norm is denoted by $\|x\|_2$, and $\|x\|_2^2 = x^T x$. Superscript T indicates transpose. The scalar product of two vectors x and y in \mathbb{R}^n will be denoted by $x^T y$. The space of the $m \times n$ matrices with real components will be indicated by $\mathbb{R}^{m \times n}$. The rank of a matrix A will be indicated by $\text{rank} A$. A square matrix $A \in \mathbb{R}^{n \times n}$ is positive semidefinite if $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$ and positive definite if $x^T A x > 0$ for all $0 \neq x \in \mathbb{R}^n$. If $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\pm\infty\}$, $\text{dom } f$ is the set of points x for which $f(x)$ is defined and $f(x) \in \mathbb{R}$. The gradient $\nabla f(x)$ of a continuously differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ at a point $x \in \mathbb{R}^n$ is a column vector with components $[\nabla f(x)]_j = \frac{\partial f(x)}{\partial x_j}$. For a twice differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, the Hessian $\nabla^2 f(x)$ belongs to $\mathbb{R}^{n \times n}$ and $[\nabla^2 f(x)]_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$. If $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a continuously differentiable vector-valued function, then $\nabla F(x) \in \mathbb{R}^{m \times n}$ denotes the Jacobian matrix of F at $x \in \mathbb{R}^n$. Here and throughout the symbols $:=$ and \equiv denote definition of the term on the left and the right sides of each symbol, respectively.

2 Convex Sets and Functions

Definition 1 ([56]) A set $C \subseteq \mathbb{R}^n$ is a convex set if

$$x^1, x^2 \in C, \quad \lambda \in [0, 1] \quad \implies \quad (1 - \lambda)x^1 + \lambda x^2 \in C.$$

Therefore, for a convex set C the segment joining any two distinct points in C is all contained in C .

Definition 2 A point $x \in \mathbb{R}^n$ is a convex combination of $x^1, x^2, \dots, x^k \in \mathbb{R}^n$ if there exist scalars $\lambda_1, \lambda_2, \dots, \lambda_k$ such that

$$x = \lambda_1 x^1 + \lambda_2 x^2 + \dots + \lambda_k x^k \quad \text{with } \lambda_1 + \lambda_2 + \dots + \lambda_k = 1, \quad \lambda_j \geq 0, \quad j = 1, \dots, k.$$

The concept of convex hull of a set is extremely important in optimization.

Definition 3 ([46, Definition 3.1.16]) Given $S \subseteq \mathbb{R}^n$ the convex hull of S is defined as

$$\text{conv } S := \left\{ \begin{array}{l} x \in \mathbb{R}^n : x = \lambda_1 x^1 + \lambda_2 x^2 + \dots + \lambda_k x^k \\ x^1, \dots, x^k \in S \\ \lambda_j \in [0, 1], \quad j = 1, \dots, k \\ \lambda_1 + \lambda_2 + \dots + \lambda_k = 1 \end{array} \right\}$$

Clearly, if S is a convex set, then $S = \text{conv } S$.

Hyperplanes

$$H := \left\{ x \in \mathbb{R}^n : p^T x = \mu \right\},$$

and halfspaces

$$S := \left\{ x \in \mathbb{R}^n : p^T x \leq \mu \right\}$$

are examples of convex sets (here $0 \neq p \in \mathbb{R}^n$, and $\mu \in \mathbb{R}$). The set of symmetric positive semidefinite matrices is a convex subsets of $\mathbb{R}^{n \times n}$ the set of square matrices of dimension n . Another interesting example of convex set is the norm cone

$$C := \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \in \mathbb{R}^{n+1} : \|x\|_2 \leq t \right\} \subseteq \mathbb{R}^{n+1}.$$

Definition 4 ([46, Definition 4.1.1]) Let $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, f is a convex function if $\text{dom } f$ is a convex set and

$$f\left[(1 - \lambda)x + \lambda y\right] \leq (1 - \lambda)f(x) + \lambda f(y) \quad \forall x, y \in \text{dom } f, \lambda \in [0, 1] \quad (1)$$

The following propositions provide necessary and sufficient conditions for convexity for differentiable and twice differentiable functions.

Proposition 1 ([8, 3.1.3], [46, Theorem 6.1.2]) Let $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ be a differentiable function. The function f is convex if and only if $\text{dom } f$ is convex and

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \quad \forall x, y \in \text{dom } f. \quad (2)$$

Proposition 2 ([8, 3.1.4], [46, Theorem 6.3.1]) Let $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, be twice differentiable. The function f is convex if and only if $\text{dom } f$ is convex and $\nabla^2 f(x)$ is positive semidefinite $\forall x \in \text{dom } f$.

Finally, let introduce two additional weaker classes of functions whose importance will be clear when necessary optimality conditions for unconstrained optimization problems will be introduced.

Definition 5 ([46, Definition 9.3.1]) Let $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ be a differentiable function. The function f is pseudo-convex if

$$x, y \in \text{dom } f, \quad \nabla f(x)^T (y - x) \geq 0 \Rightarrow f(y) \geq f(x)$$

A convex function is also pseudo-convex.

Definition 6 ([8, 3.4.1], [46, Definition 9.1.1]) Let $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$. The function f is quasi-convex if

$$x, y \in \text{dom } f, \quad f((1 - \lambda)x + \lambda y) \leq \max\{f(x), f(y)\} \text{ for } \lambda \in [0, 1]$$

Clearly, a convex function is also quasi-convex. Moreover, a differentiable pseudo-convex function is also quasi-convex. The opposite is not true.

For simplicity, in the sequel, we will only consider functions whose domain is \mathbb{R}^n .

3 Unconstrained Optimization

In this section we will present optimality conditions and algorithms for the unconstrained optimization problems

$$\min_{x \in \mathcal{F}} f(x) \tag{3}$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is the objective function and the feasible set $\mathcal{F} \subseteq \mathbb{R}^n$ is an open subset of \mathbb{R}^n (note that \mathcal{F} may coincide with \mathbb{R}^n).

First, let introduce the following definitions.

Definition 7 ([6, Definition 3.4.1]) Let $\bar{x} \in \mathcal{F}$. The point \bar{x} is a global minimum of Problem (3) if

$$f(\bar{x}) \leq f(x), \quad \forall x \in \mathcal{F}.$$

The point \bar{x} is a strict global minimum of Problem (3) if

$$f(\bar{x}) < f(x), \quad \forall x \in \mathcal{F}, \quad x \neq \bar{x}.$$

Definition 8 ([6, Definition 3.4.1]) Let $\bar{x} \in \mathcal{F}$. The point \bar{x} is a local minimum of Problem (3) if $\exists \gamma > 0$ such that

$$f(\bar{x}) \leq f(x) \quad \forall x \in \mathcal{F} \text{ and } \|x - \bar{x}\| \leq \gamma.$$

The point \bar{x} is a strict local minimum if $\exists \gamma > 0$ such that

$$f(\bar{x}) < f(x) \quad \forall x \in \mathcal{F} : 0 < \|x - \bar{x}\| \leq \gamma.$$

The point \bar{x} is a strong or isolated local minimum if $\exists \gamma > 0$ such that \bar{x} is the only local minimum in $\{x \in \mathbb{R}^n : \|x - \bar{x}\| \leq \gamma\} \cap \mathcal{F}$.

If \bar{x} is a strong local minimum, then it is also a strict local minimum. On the other hand, a strict local minimum may not be necessarily a strong local minimum. For example, the function [7, p. 21]

$$f(x) = \begin{cases} x^2 \left(\sqrt{2} - \sin \left(\frac{4}{3}\pi - \sqrt{3} \ln(x^2) \right) \right) & \text{if } x \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

has in $x = 0$ a strict local minimum that is not a strong (isolated) local minimum. In the remaining of the section, without loss of generality, we assume that $\mathcal{F} = \mathbb{R}^n$

3.1 Necessary and Sufficient Optimality Conditions

Consider the nonlinear optimization problem (3) where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a continuously differentiable function. The directional derivatives [3, Chap. 8] of f at \bar{x} along the direction d is given by

$$\lim_{t \rightarrow 0^+} \frac{f(\bar{x} + td) - f(\bar{x})}{t} = \nabla f(\bar{x})^T d.$$

A direction $d \in \mathbb{R}^n$ is a descent direction for $f(x)$ at \bar{x} if

$$\exists \delta > 0 : f(\bar{x} + td) < f(\bar{x}) \quad \forall 0 < t \leq \delta \quad (4)$$

If f is continuously differentiable, then d is a descent direction at \bar{x} if and only if $\nabla f(\bar{x})^T d < 0$.

A point x^* is a local minimum if at x^* there are no descent directions. The following theorems provide first and second order necessary optimality conditions for a point x^* .

Theorem 1 *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a continuously differentiable function and let $x^* \in \mathbb{R}^n$. If x^* is a local minimum then*

$$\nabla f(x^*) = 0. \quad (5)$$

The above result can be easily proved by contradiction; in fact, if $\nabla f(x^*) \neq 0$, then $d = -\nabla f(x^*)$ is a descent direction. This observation is the basis of the most used technique for function minimization: the gradient method.

Theorem 2 ([35, Proposition 2.5]) *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be twice continuously differentiable and let $x^* \in \mathbb{R}^n$. If x^* is a local minimum then*

Algorithm 1: Generic Minimization Algorithm

- 1 **Choose** $x^0 \in \mathbb{R}^n$, **Set** $k = 0$;
 - 2 **if** $x^k \in \Omega$ **then Stop**;
 - 3 **Compute** a search direction $d^k \in \mathbb{R}^n$;
 - 4 **Compute** a stepsize $\alpha_k > 0$ along d^k ;
 - 5 **Set** $x^{k+1} = x^k + \alpha_k d^k$, $k = k + 1$, **return** to 2;
-

$$\nabla f(x^*) = 0 \tag{6}$$

$$\nabla^2 f(x^*) \text{ is positive semidefinite.} \tag{7}$$

The above conditions cannot exactly be reversed to obtain sufficiency optimality conditions. In fact, positive definiteness of the Hessian matrix is required to obtain a strict local minimum.

Theorem 3 ([35, Proposition 2.6]) *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be twice continuously differentiable and let $x^* \in \mathbb{R}^n$. Suppose that:*

$$\nabla f(x^*) = 0 \tag{8}$$

$$\nabla^2 f(x^*) \text{ is positive definite.} \tag{9}$$

Then x^ is a strict local minimum.*

Conditions for global optima are much more complex to obtain. However, if $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a differentiable pseudo-convex function, then a local minimum is also a global minimum [46, Theorem 9.3.7].

3.2 Algorithms for Unconstrained Optimization

A very general minimization algorithm is presented in Algorithm 1 where

$$\Omega := \{x \in \mathbb{R}^n : \nabla f(x) = 0\} \tag{10}$$

is the set of stationary points. The algorithm, starting from the initial point x^0 , constructs either a finite sequence terminating in a stationary point or an infinite sequence which, under adequate conditions, converges to a stationary point, or, has at least an accumulation point that is also a stationary point.

Given the current point x^k and a search direction d^k , the new point x^{k+1} is obtained by moving along d^k with a stepsize α_k .

The following general theorem establishes conditions on the direction d^k and the stepsize α_k ensuring that, in the case an infinite sequence is obtained by Algorithm 1, the sequence has at least one accumulation point and each accumulation point is a stationary point. Before stating the theorem, it is necessary to introduce the concept of forcing function.

Definition 9 ([50, Definition 14.2.1]) A function $\sigma : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a forcing function if

$$\left. \begin{array}{l} \forall \text{ sequence } \{t_k\} \subseteq \mathbb{R}_+ \\ \lim_k \sigma(t_k) = 0 \end{array} \right\} \implies \lim_k t_k = 0.$$

Note that any non-decreasing function $\sigma : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that $\sigma(0) = 0$ and $\sigma(t) > 0$ when $t > 0$ is a forcing function. The functions

$$\sigma(t) = t, \quad \sigma(t) = ct^q, \quad \text{with } c > 0, \quad q > 0$$

are examples of forcing functions.

Theorem 4 Let $\{x^k\}$ be obtained by Algorithm 1 and assume that

(i) the level set

$$\mathcal{L}_0 := \{x \in \mathbb{R}^n : f(x) \leq f(x^0)\}$$

is compact,

(ii) $d^k \neq 0$ when $\nabla f(x^k) \neq 0$,

(iii) $f(x^{k+1}) \leq f(x^k)$,

(iv) if $\nabla f(x^k) \neq 0$ for all k then

$$\lim_k \frac{\nabla f(x^k)^T d^k}{\|d^k\|} = 0,$$

(v) when $d^k \neq 0$

$$\frac{|\nabla f(x^k)^T d^k|}{\|d^k\|} \geq \sigma \left(\|\nabla f(x^k)\| \right)$$

for some forcing function σ .

Then, either the algorithm terminates after a finite number of iterations in a stationary point or an infinite sequence $\{x^k\}$ is generated that satisfies the following properties:

- (a) the sequence $\{x^k\}$ remains in \mathcal{L}_0 and has at least one accumulation point;
- (b) each accumulation point of $\{x^k\}$ belongs to \mathcal{L}_0 ;

- (c) the sequence of real numbers $\{f(x^k)\}$ converges;
- (d) $\lim_k \|\nabla f(x^k)\| = 0$;
- (e) each accumulation point x^* of the sequence $\{x^k\}$ is a stationary point.

With reference to Theorem 4,

- Conditions (iii) and (v) can be guaranteed by choosing an opportune descent direction d^k and a line-search along this direction. In particular, using the Euclidean norm and choosing $\sigma(t) = ct$ (with $c > 0$), Condition (v) can be written as

$$\nabla f(x^k)^T d^k \leq -c \|d^k\|_2 \|\nabla f(x^k)\|_2,$$

that is, when x^k is not a stationary point, a direction d^k must be chosen such that

$$\frac{\nabla f(x^k)^T d^k}{\|d^k\|_2 \|\nabla f(x^k)\|_2} \leq -c.$$

Geometrically, the above condition requires that the cosine of the angle between the direction d^k and the direction $-\nabla f(x^k)$ (the “antigradient”) must be greater than a constant independent of k . This implies that d^k and the gradient direction cannot be orthogonal as k goes to infinity.

If $\sigma(t) = ct^q$ (with $c > 0$), Condition (v) becomes, instead,

$$\frac{\nabla f(x^k)^T d^k}{\|d^k\|} \leq -c \|\nabla f(x^k)\|^q.$$

- Condition (iv) can be guaranteed using specific safeguard rules on the line-search along the direction d^k .

From the above considerations, it is clear the importance of inexact line-search procedures for determining the stepsize α_k .

A simple, but also very effective, line-search procedure is the Armijo Backtracking method [4, 7, p. 29]. Given x^k and a descent search direction d^k , the algorithm starts with a fixed large value of α and decreases it (that is, the new values of α is obtained by multiplying the current value of α by a constant $\delta < 1$) until the stopping criterion

$$f(x^k + \alpha d^k) \leq f(x^k) + \gamma \alpha \nabla f(x^k)^T d^k \tag{11}$$

is satisfied, where $\gamma \in (0, \frac{1}{2})$. In other words, the Armijo procedures chooses as α_k the maximum value of α in the set

$$S = \left\{ \alpha : \alpha = \delta^l \alpha_{\text{init}}, l = 0, 1, \dots \right\}$$

for which Condition (11) holds.

The following proposition demonstrates that, if the search direction d^k is opportunely chosen, and the stepsize is obtained according to the Armijo rule, then conditions (iii) and (iv) in Theorem 4 are automatically satisfied.

Proposition 3 *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be continuously differentiable. Suppose that:*

- (i) \mathcal{L}_0 is a compact set;
- (ii) $\nabla f(x^k)^T d^k < 0, \forall k$;
- (iii) there exists a forcing function $\sigma : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that

$$\|d^k\| \geq \sigma \left(\frac{|\nabla f(x^k)^T d^k|}{\|d^k\|} \right).$$

Then, if α_k is chosen according to the Armijo procedure, and $x^{k+1} = x^k + \alpha_k d^k$

- (a) $f(x^{k+1}) < f(x^k)$;
- (b) $\lim_k \frac{|\nabla f(x^k)^T d^k|}{\|d^k\|} = 0$.

A different line-search stopping criterion was proposed by Goldstein [31] considering two lines

$$f(x^k) + \gamma_1 \alpha \nabla f(x^k)^T d^k = 0,$$

$$f(x^k) + \gamma_2 \alpha \nabla f(x^k)^T d^k = 0$$

where $0 < \gamma_1 < \gamma_2 < \frac{1}{2}$. The chosen stepsize $\alpha_k > 0$ must satisfy, instead of (11), the following conditions:

$$f(x^k + \alpha_k d^k) \leq f(x^k) + \gamma_1 \alpha_k \nabla f(x_k)^T d^k, \quad (12a)$$

$$f(x^k + \alpha_k d^k) \geq f(x^k) + \gamma_2 \alpha_k \nabla f(x_k)^T d^k. \quad (12b)$$

From a geometrical point of view, this corresponds to choose a value for α_k for which $f(x^k + \alpha_k d^k)$ is between the two straight lines with slopes $\gamma_1 \nabla f(x^k)^T d^k$ and $\gamma_2 \nabla f(x_k)^T d^k$, and passing through the point $(0, f(x_k))$.

Moreover, let γ_1 and γ_2 such that $0 < \gamma_1 < \gamma_2 < 1$. The step length α_k satisfy the Wolfe conditions [62] if the following conditions are satisfied:

$$f(x^k + \alpha_k d^k) \leq f(x^k) + \gamma_1 \alpha_k \nabla f(x^k)^T d^k, \quad (13a)$$

$$\nabla f(x^k + \alpha_k d^k)^T d^k \geq \gamma_2 \nabla f(x^k)^T d^k. \quad (13b)$$

It is possible to show that, when d^k is an opportune descent direction, then using both the Goldstein and the Wolfe line-search methods, similar results to those in Proposition 3 can be obtained.

Various other line-search procedures have been proposed in literature, including line-search procedures that do not use information on the gradient of the function [15, 44] and/or non-monotone line-search procedures [33, 34, 63] for which the monotonicity of the objective function for the generated sequence is not required.

3.3 The Gradient Method

In the gradient method (also known as the steepest descent method) the search direction is given by

$$d^k = -\nabla f(x^k). \quad (14)$$

For this choice of the search direction, both Condition (iii) and Condition (v) in Theorem 4 are trivially satisfied with $\sigma(t) = t$ and, therefore, the method is globally convergent when one of line-search procedures outlined in the previous subsection is applied. However, the gradient method, even when applied to the minimization of a strictly convex quadratic function

$$\min_x \frac{1}{2} x^T M x + q^T x$$

($M \in \mathbb{R}^{n \times n}$ positive definite, and $q \in \mathbb{R}^n$) and exact line-search is utilized, exhibits linear rate of convergence. The theorem below shows that the gradient method has at least linear rate of convergence.

Theorem 5 ([45, Sect. 7.6, p. 218]) *Let $M \in \mathbb{R}^{n \times n}$ be a symmetric positive definite matrix and let $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ be the eigenvalues of M , let $q \in \mathbb{R}^n$ and let $f(x) := \frac{1}{2} x^T M x + q^T x$. Let $\{x^k\}$ be obtained by the gradient method with exact line-search:*

$$x^{k+1} = x^k - \frac{(Mx^k + q)^T (Mx^k + q)}{(Mx^k + q)^T M (Mx^k + q)} (Mx^k + q).$$

Then¹

¹ Here $\|x\|_M^2 := x^T M x$.

$$\|x^{k+1} - x^*\|_M \leq \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} \|x^k - x^*\|_M \quad (15a)$$

and

$$\|x^{k+1} - x^*\|_2 \leq \left(\frac{\lambda_n}{\lambda_1}\right)^{\frac{1}{2}} \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} \|x^k - x^*\|_2 \quad (15b)$$

where $x^* = -M^{-1}q$ is the unique minimum point.

Moreover, it is not difficult to construct an examples (see [8, Example 9.3.2, p. 469]) of quadratic problems where the inequalities (15) above are satisfied as equality, demonstrating that, in fact, the gradient method has linear rate of convergence.

3.4 The Newton's Method

As stated before the gradient method exhibits a slow rate of convergence. Here we consider a method that, instead, has quadratic rate of convergence at the expenses of a higher cost per iteration, requiring, in addition to calculate the gradient at each point, also the calculation of the Hessian function. Moreover, only local convergence can be established, that is convergence is guaranteed only if the initial point x^0 is chosen sufficiently close to the stationary point x^* .

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be twice continuously differentiable. From Taylor's series expansion we have that

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

where

$$\lim_{s \rightarrow 0} \frac{\beta(x, s)}{\|s\|^2} = 0.$$

In Newton's method, at each iteration a quadratic approximation of the function f around the current point x^k is constructed:

$$q_k(s) := f(x^k) + \nabla f(x^k)^T s + \frac{1}{2} s^T \nabla^2 f(x^k) s,$$

and the new point x^{k+1} is obtained as $x^{k+1} = x^k + s^k$ where

$$s^k \in \underset{s}{\operatorname{argmin}} q_k(s).$$

The following theorem provides conditions for local convergence of the Newton’s method for minimization problems.

Theorem 6 ([25, Theorem 3.1.1] [35, Proposizione 7.2]) *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a twice continuously differentiable function, and assume that*

- (a) *there exists $x^* \in \mathbb{R}^n : \nabla f(x^*) = 0$,*
- (b) *$\nabla^2 f(x^*)$ is non singular,*
- (c) *$\nabla^2 f(x)$ is a Lipschitz–continuous function, i.e.,*

$$\exists L > 0 : \forall x, y \in \mathbb{R}^n \quad \|\nabla^2 f(x) - \nabla^2 f(y)\| \leq L \|x - y\| .$$

Then, there exists $B(x^, \rho) := \{x \in \mathbb{R}^n : \|x - x^*\| \leq \rho\}$ with $\rho > 0$ such that, if $x^0 \in B(x^*, \rho)$,*

- (i) *the Newton’s iterate is well defined,*
- (ii) *the sequence $\{x^k\}$ remains in $B(x^*, \rho)$,*
- (iii) *the sequence $\{x^k\}$ converges to x^* with at least a quadratic rate of convergence*

$$\|x^{k+1} - x^*\| \leq \alpha \|x^k - x^*\|^2, \text{ for some } \alpha > 0.$$

3.5 The Conjugate Gradient Method

The slow convergence of the gradient method and the heavy requirements of the fast converging Newton’s method where, at each iteration, it is necessary to calculate the Hessian matrix, set the stage for new classes of methods that only require to compute the gradient of the function to minimize, and, at the same time, exhibit q –superlinear rate of convergence [50]. In this section the important class of conjugate gradient methods is introduced, while the next section is devoted to Quasi-Newton’s methods.

Definition 10 Let $M \in \mathbb{R}^{n \times n}$ be a symmetric positive definite matrix. The n nonzero vectors d^0, d^1, \dots, d^{n-1} in \mathbb{R}^n are conjugate directions with respect to the matrix M if

$$d^{iT} M d^j = 0 \quad \forall i, j = 0, 1, \dots, n - 1, i \neq j. \tag{16}$$

Consider the quadratic function

$$f(x) = \frac{1}{2} x^T M x + q^T x \tag{17}$$

Algorithm 2: Generic Conjugate Direction Method for Quadratic Problems

```

1 Choose  $x^0 \in \mathbb{R}^n$ ;
2 Choose  $d^0, d^1, \dots, d^{n-1}$  nonzero conjugate directions with respect to  $M$ ;
3 for  $k=0$  to  $n-1$  do
4   Set  $\alpha_k = \frac{-(Mx^k + q)^T d^k}{d^{kT} M d^k}$ ;
5   Set  $x^{k+1} = x^k + \alpha_k d^k$ 
6 end
  
```

where M is a symmetric positive matrix, $M \in \mathbb{R}^{n \times n}$ and $q \in \mathbb{R}^n$; the following algorithm utilizes n conjugate directions to determine the unique point minimizing $f(x)$.

One important result for this algorithm is that for quadratic strictly convex problems, when conjugate directions are utilized as search directions, then convergence is achieved in a finite number of iterations.

Theorem 7 Consider the quadratic function $f(x) = \frac{1}{2}x^T Mx + q^T x$ where M is a symmetric positive matrix, $M \in \mathbb{R}^{n \times n}$ and $q \in \mathbb{R}^n$. Starting from any $x^0 \in \mathbb{R}^n$ apply Algorithm 2. Then x^{k+1} is the minimizer of $f(x)$ over the affine set

$$S_k := \left\{ x \in \mathbb{R}^n : x = x^0 + \sum_{j=0}^k \lambda_j d^j, \lambda_j \in \mathbb{R} \right\}$$

for $k = 0, 1, \dots, n-1$. Moreover, $x^n = -M^{-1}q$ is the unique minimizers of $f(x)$ over \mathbb{R}^n .

The above Algorithm 2 allows different choices for the set of conjugate directions. A interesting choice is to link these conjugate directions to the gradient of the function $f(x)$. Algorithm 3 explores this possibility.

Theorem 8 ([25, Theorem 4.1.1]) Consider the quadratic function $f(x) = \frac{1}{2}x^T Mx + q^T x$ where $M \in \mathbb{R}^{n \times n}$ is a symmetric positive matrix, and $q \in \mathbb{R}^n$. The Conjugate Gradient Algorithm 3 terminates after $m < n$ iterations and for all $i = 0, \dots, m$

$$d^i{}^T M d^j = 0, \quad j = 0, \dots, i-1, \quad (18a)$$

Algorithm 3: Conjugate Gradient Algorithm

```

1 Choose  $x^0 \in \mathbb{R}^n$ , Set  $k = 0$ ;
2 while  $x^k \notin \Omega$  do
3   if  $k = 0$  then
4     Set  $d^k = -\nabla f(x^k)$ ;
5   else
6     Set  $\beta_{k-1} = \frac{\nabla f(x^k)^T M d^{k-1}}{d^{k-1T} M d^{k-1}}$ ;
7     Set  $d^k = -\nabla f(x^k) + \beta_{k-1} d^{k-1}$ ;
8   end
9   Set  $\alpha_k = \frac{-\nabla f(x^k)^T d^k}{d^{kT} M d^k}$ ;
10  Set  $x^{k+1} = x^k + \alpha_k d^k$ ;
11  Set  $k = k + 1$ ;
12 end

```

$$\nabla f(x^i)^T \nabla f(x^j) = 0, \quad j = 0, \dots, i - 1, \quad (18b)$$

$$\nabla f(x^i)^T d^i = -\nabla f(x^i)^T \nabla f(x^i). \quad (18c)$$

The theorem above shows that the directions generated by Algorithm 3 are conjugate directions with respect to the matrix M , and, therefore, finite termination of algorithm follows from Theorem 7.

It is important to note that

$$\alpha_k = -\frac{\nabla f(x^k)^T d^k}{d^{kT} M d^k} = \frac{\nabla f(x^k)^T \nabla f(x^k)}{d^{kT} M d^k} > 0. \quad (19)$$

Moreover, from $x^k = x^{k-1} + \alpha_{k-1} d^{k-1}$ it follows that

$$\alpha_{k-1} M d^{k-1} = M x^k - M x^{k-1} = \nabla f(x^k) - \nabla f(x^{k-1})$$

and hence

$$\nabla f(x^k)^T M d^{k-1} = \frac{1}{\alpha_{k-1}} \nabla f(x^k)^T (\nabla f(x^k) - \nabla f(x^{k-1})) = \frac{1}{\alpha_{k-1}} \nabla f(x^k)^T \nabla f(x^k).$$

Now, from (19), $\alpha_{k-1} d^{k-1T} M d^{k-1} = \nabla f(x^{k-1})^T \nabla f(x^{k-1})$, and hence

$$\begin{aligned}
\beta_{k-1} &= \frac{\nabla f(x^k)^T M d^{k-1}}{d^{k-1T} M d^{k-1}} = \frac{1}{\alpha_{k-1}} \frac{\nabla f(x^k)^T \nabla f(x^k)}{d^{k-1T} M d^{k-1}} \\
&= \frac{\nabla f(x^k)^T \nabla f(x^k)}{\nabla f(x^{k-1})^T \nabla f(x^{k-1})} =: \beta_{k-1}^{\text{FR}}
\end{aligned} \tag{20}$$

This formula was first proposed by Fletcher and Reeves [27]. Alternative choices, all equivalent in the case of strictly convex quadratic problems, for β_{k-1} are

$$\beta_{k-1}^{\text{PRP}} = \frac{\nabla f(x^k)^T [\nabla f(x^k) - \nabla f(x^{k-1})]}{\nabla f(x^{k-1})^T \nabla f(x^{k-1})}. \tag{21a}$$

$$\beta_{k-1}^{\text{HS}} = -\frac{\nabla f(x^k)^T [\nabla f(x^k) - \nabla f(x^{k-1})]}{d^{k-1T} [\nabla f(x^k) - \nabla f(x^{k-1})]} \tag{21b}$$

proposed, respectively, by Polak and Ribière [52] and Polyak [53] and Hestenes and Stiefel [39].

For general non-quadratic functions, there is no guarantee that the conjugate gradient method will terminate in a finite number of steps. Moreover, it is extremely difficult to exactly solve the one-dimensional subproblem and inexact line-search methods such as Armijo or Goldstein or Wolfe methods must be utilized. Moreover, each n iterations or when a non-descent direction is generated, a reset step should be performed using as search direction the negative gradient direction. Computational results, however, show that the use of a restarting procedure is not convenient and is better to opportunely modify the formula for β_{k-1} and to choose specific line-search procedure to globalize the overall scheme.

The first global convergence result of the Fletcher-Reeves method with inexact line search was given by Al-Baali [2]. Under strong Wolfe conditions, he demonstrated that the method generates sufficient descent directions and, therefore, global convergence can be established.

For the Polak-Ribière-Polyak method the convergence for general non-linear function is uncertain. While global convergence can be proved in the case of strongly convex functions, there are examples of not strongly convex functions, for which the method may not converge, even with an exact line search. Instead, convergence can be proved when

$$\beta_{k-1}^{\text{PRP}^+} = \min\{\beta_{k-1}^{\text{PRP}}, 0\}$$

is utilized [29]. Additional information on conjugate gradient method for general nonlinear functions can be found in [35, 38, 54].

3.6 Quasi-Newton's Methods

As already noted earlier, the Newton's method is locally convergent at quadratic rate and requires, at each iteration, to compute the Hessian function. On the other hand, the gradient method for which global convergence can be established, only requires to compute the gradient of the function; however, its convergence can be quite slow (at most linear, in some cases).

The aim of Quasi-Newton's methods is to define procedures that

- will not require to compute the Hessian function,
- exhibits q -superlinear rate of convergence.

Two classes of Quasi-Newton methods have been studied in literature:

- direct methods for which

$$\begin{cases} x^{k+1} = x^k - [B^k]^{-1} \nabla f(x^k), \\ B^{k+1} = B^k + \Delta B^k, \end{cases} \quad B^{k+1} \approx \nabla^2 f(x^{k+1});$$

- inverse methods where

$$\begin{cases} x^{k+1} = x^k - H^k \nabla f(x^k), \\ H^{k+1} = H^k + \Delta H^k, \end{cases} \quad H^{k+1} \approx [\nabla^2 f(x^{k+1})]^{-1}.$$

In order to derive updating formulas for B^k and H^k , consider again the quadratic function

$$f(x) = \frac{1}{2}x^T Mx + q^T x$$

where $M \in \mathbb{R}^{n \times n}$ is a symmetric positive matrix, and $q \in \mathbb{R}^n$. Let now $x, y \in \mathbb{R}^n$, then

$$\nabla f(y) = \nabla f(x) + M(y - x)$$

or equivalently

$$M^{-1} (\nabla f(y) - \nabla f(x)) = y - x.$$

Therefore, since B^{k+1} (resp. H^{k+1}) must be an approximation of M (resp. M^{-1}), it is reasonable to require that

Algorithm 4: Generic (Direct) Quasi–Newton Method

```

1 Choose  $x^0 \in \mathbb{R}^n$ , Choose  $B^0 \in \mathbb{R}^{n \times n}$ , Set  $k = 0$ ;
2 while  $x^k \notin \Omega$  do
3   Set  $d^k = -(B^k)^{-1} \nabla f(x^k)$ ;
4   Compute  $\alpha_k$  using a line–search procedure;
5   Set  $x^{k+1} = x^k + \alpha_k d^k$ ;
6   Compute  $B^{k+1} = B^k + \Delta B^k$  Set  $k = k + 1$ ;
7 end
  
```

$$\nabla f(x^{k+1}) - \nabla f(x^k) = B^{k+1} (x^{k+1} - x^k) \quad (22a)$$

or

$$H^{k+1} (\nabla f(x^{k+1}) - \nabla f(x^k)) = x^{k+1} - x^k. \quad (22b)$$

After defining $\gamma^k := \nabla f(x^{k+1}) - \nabla f(x^k)$ and $\delta^k := x^{k+1} - x^k$ the above conditions become

$$\gamma^k = B^{k+1} \delta^k \quad (23a)$$

and

$$H^{k+1} \gamma^k = \delta^k. \quad (23b)$$

These conditions are known as “secant conditions”. Furthermore, it is reasonable to require that B^{k+1} (resp. H^{k+1}) be symmetric and as close as possible to B^k (resp. H^k) imposing that B^{k+1} (resp. H^{k+1}) differs from B^k (resp. H^k) by a matrix of rank 1 or 2 and a minimality condition in some norm, for example in Frobenius norm²[18]. The generic direct Quasi–Newton method is reported in Algorithm 4.

A similar algorithm can be easily constructed for a Quasi–Newton generic inverse method.

² For a matrix $A \in \mathbb{R}^{m \times n}$, the Frobenius norm $\|A\|_F$ of A is defined as [48]

$$\begin{aligned} \|A\|_F &:= \sqrt{\sum_{i=1}^m \sum_{j=1}^n A_{ij}^2} = \sqrt{\sum_{i=1}^m \|A_{i \cdot}\|_2^2} = \sqrt{\sum_{j=1}^n \|A_{\cdot j}\|_2^2} \\ &= \sqrt{\text{trace}(A^T A)} = \sqrt{\text{trace}(A A^T)}. \end{aligned}$$

In the sequel, unless strictly necessary we will drop the superscripts k and $k + 1$; the current matrix will be indicated simply by B (resp. H) while the updated matrix will be indicated by \bar{B} (resp. \bar{H}). A similar rule will be used for γ^k and δ^k .

The simplest updating direct formula for B is a rank-1 updating formula

$$\bar{B} = B + \rho uv^T \tag{24}$$

where $\rho \in \mathbb{R}$ and $u, v \in \mathbb{R}^n$, $u, v \neq 0$.

For the symmetric case (i.e., $u = v$), we have

$$\bar{B} = B + \rho uu^T \tag{25}$$

and, imposing the secant condition $\bar{B}\delta = \gamma$ either $\bar{B} = B$ (which happens if already $B\delta = \gamma$) or, under the hypothesis that $(\gamma - B\delta)^T \delta \neq 0$, the following formula, known as Symmetric Rank-1 updating formula (SR1), is obtained

$$\bar{B} = B + \frac{(\gamma - B\delta)(\gamma - B\delta)^T}{(\gamma - B\delta)^T \delta}. \tag{26}$$

This formula was first proposed by Davidon [14] and later rediscovered by Broyden [10].

Using $H = B^{-1}$ and the Sherman–Morrison–Woodbury [32, 37] formula, it is possible to derive the inverse Symmetric Rank-1 updating scheme:

$$\bar{H} = H + \frac{(\delta - H\gamma)(\delta - H\gamma)^T}{(\delta - H\gamma)^T \gamma}. \tag{27}$$

Note that in this case the secant condition completely determines the updating formula under the hypothesis that $(\gamma - B\delta)^T \delta \neq 0$ (resp. $(H\gamma - \delta)^T \gamma \neq 0$). However, it must be noticed that there is no guarantee that the search direction that is obtained is a descent direction.

The symmetric rank-1 updating formula (27) has a very interesting behaviour when applied to a quadratic function.

Theorem 9 ([17, Theorem 7.1]) *Let $A \in \mathbb{R}^{n \times n}$ be symmetric and non-singular. Let $\{s^0, s^1, \dots, s^m\}$ be m vectors spanning \mathbb{R}^n and let $y^k = As^k$, $k = 0, \dots, m$. Again, let H^0 be a symmetric $n \times n$ real matrix and for $k = 0, \dots, m$ let*

$$H^{k+1} = H^k + \frac{(s^k - H^k y^k)(s^k - H^k y^k)^T}{(s^k - H^k y^k)^T y^k}$$

assuming that $(s^k - H^k y^k)^T y^k \neq 0$. Then

$$H^{m+1} = A^{-1}.$$

The theorem above shows that, given a nonsingular matrix A , the inverse A^{-1} can be calculated using the SR1 formula which only involves matrix–vector multiplications

For general nonlinear functions, under specific assumptions, including that

$$\left| (\gamma^k - B^k \delta^k)^T \delta^k \right| \geq c \left\| \gamma^k - B^k \delta^k \right\|_2 \left\| \delta^k \right\|_2$$

and that the sequence $\{x^k\}$ has at least one accumulation point x^* , then [13, Theorem 2]

$$\lim_k \left\| B^k - \nabla^2 f(x^*) \right\| = 0.$$

In the non–symmetric case (mostly utilized for solving system of nonlinear equations) the updating formula require to choose both vectors u and v and the secant condition does not define uniquely both vectors. Therefore, the new matrix \bar{B} is required to be as close as possible to the current matrix B . The following lemma shows that a rank-1 updating formula is obtained when the closest matrix (in Frobenius norm) to the current matrix B is calculated among all matrices for which the secant condition is satisfied.

Theorem 10 ([17, Theorem 4.1], [18, Lemma 8.1.1]) *Let $B \in \mathbb{R}^{n \times n}$ and let $\gamma, \delta \in \mathbb{R}^n$. Then, the solution of the minimization problem*

$$\begin{aligned} \min_A \quad & \|A - B\|_F \\ \text{*subject to} \quad & \gamma = A\delta \end{aligned} \tag{28}$$

is given by

$$\bar{B} = B + \frac{(\gamma - B\delta) \delta^T}{\delta^T \delta}. \tag{29}$$

Using again the Sherman–Morrison–Woodbury formula, it possible (under the hypothesis that $\gamma^T H \delta \neq 0$) to derive the inverse updating formula:

$$\bar{H} = H + \frac{(\delta - H\gamma) \delta^T H}{\delta^T H \gamma} \tag{30}$$

The updating formula (30) was first proposed by Broyden [9] in the context of solving systems of nonlinear equations. Locally q –superlinear con-

vergence for this method was then proved by Broyden, Dennis and Morè in [12], see also [19, Theorem 3.3.3] and [19, Theorem 8.2.2].

The generic rank–2 updating formula is given by

$$\bar{B} = B + auu^T + bvv^T$$

where $a, b \in \mathbb{R}$ and $u, v \in \mathbb{R}^n$; similar rank–2 updating formulas can be defined for \bar{H} .

The use of a rank–2 updating formula allows to imposes important conditions, in addition to the secant property, on the updated matrix such as symmetry and positive definiteness. Theorem 11 show that the closest matrix to B for which symmetry and secant condition are satisfied can be obtained via a rank–2 modification. Here a weighted Frobenius norm is utilized, where W is the weight matrix:

$$\|A - B\|_{W,F} := \|W(A - B)W\|_F.$$

Theorem 11 ([17, Theorem 7.3]) *Let $B \in \mathbb{R}^{n \times n}$ be a symmetric matrix and let $c, \delta, \gamma \in \mathbb{R}^n$ with $c^T \delta > 0$. Let $W \in \mathbb{R}^{n \times n}$ be a symmetric nonsingular matrix such that $Wc = W^{-1}\delta$. Then, the unique solution of*

$$\begin{aligned} \min_A \quad & \|A - B\|_{W,F} \\ \text{*subject to} \quad & A\delta = \gamma \\ & A \text{ symmetric} \end{aligned} \tag{31}$$

is given by

$$\bar{B} = B + \frac{(\gamma - B\delta)c^T + c(\gamma - B\delta)^T}{c^T \delta} - \frac{(\gamma - B\delta)^T \delta}{(c^T \delta)^2} cc^T. \tag{32}$$

The above theorem leaves space for opportune choices of the matrix W and the vector c , which can be utilized to impose positive definiteness of the updating formula, i.e., conditions that ensure that the matrix \bar{B} be positive definite, provided that the same property holds for B . Positive definiteness of the matrix \bar{B} ensures that the search direction utilized in the Quasi–Newton method (see Algorithm 4) is a descent direction.

In view of the fact that $\bar{B}\delta = \gamma$, if \bar{B} is positive definite then $0 < \delta^T \bar{B}\delta = \delta^T \gamma$. It is not difficult to show that this condition is also necessary for the positive definiteness of \bar{B} [17, Theorem 7.5].

A simple, but extremely interesting, choice for the vector c is $c = \gamma$. This specific choice leads to the updating formula

$$\bar{B} = B + \frac{(\gamma - B\delta)\gamma^T + \gamma(\gamma - B\delta)^T}{\gamma^T\delta} - \frac{(\gamma - B\delta)^T\delta}{(\gamma^T\delta)^2}\gamma\gamma^T. \quad (33)$$

Again, using the Sherman–Morrison–Woodbury formula, it is possible to compute the corresponding inverse updating formula (here $H = B^{-1}$ and $\bar{H} = \bar{B}^{-1}$)

$$\bar{H} = H + \frac{\delta\delta^T}{\gamma^T\delta} - \frac{H\gamma\gamma^T H}{\gamma^T H \gamma}. \quad (34)$$

These formulas, known as direct and inverse DFP formulas, were firstly proposed by Davidon [14] and later rediscovered by Fletcher and Powell that also clarified and improved them [26].

Moreover, it is also possible to construct updating formulas directly for H similarly requiring that \bar{H} be symmetric, that the secant condition be satisfied and \bar{H} be as close as possible to H in a specific norm as shown in the theorem below, in a similar way as done for the direct updating formula.

Theorem 12 *Let $H \in \mathbb{R}^{n \times n}$ be a symmetric matrix and let $d, \delta, \gamma \in \mathbb{R}^n$ with $d^T \gamma > 0$. Let $W \in \mathbb{R}^{n \times n}$ be a symmetric nonsingular matrix such that $Wd = W^{-1}\gamma$. Then, the unique solution of*

$$\begin{aligned} \min_A \quad & \|W(A - H)W\|_F \\ \text{*subject to} \quad & A\gamma = \delta \\ & A \text{ symmetric} \end{aligned} \quad (35)$$

is given by

$$\bar{H} = H + \frac{(\delta - H\gamma)d^T + d(\delta - H\gamma)^T}{d^T\gamma} - \frac{(\delta - H\gamma)^T\gamma}{(d^T\gamma)^2}dd^T. \quad (36)$$

Similarly to what done in the direct case, it is possible to impose that the new matrix \bar{H} be positive definite. Moreover, the choice $d = \delta$ brings to the following updating formula

$$\bar{H} = H + \frac{(\delta - H\gamma)\delta^T + \delta(\delta - H\gamma)^T}{\delta^T\gamma} - \frac{(\delta - H\gamma)^T\gamma}{(\delta^T\gamma)^2}\delta\delta^T. \quad (37)$$

Using the Sherman–Morrison–Woodbury formula the corresponding direct updating formula can be obtained

Algorithm 5: Quasi–Newton method (Broyden class updating scheme)

```

1 Choose  $x^0 \in \mathbb{R}^n$ , Choose  $H^0 \in \mathbb{R}^{n \times n}$ , Set  $k = 0$ ;
2 while  $x^k \notin \Omega$  do
3   Set  $d^k = -H^k \nabla f(x^k)$ ;
4   Compute  $\alpha_k$  using a line–search procedure;
5   Set  $x^{k+1} = x^k + \alpha_k d^k$ ;
6   Set  $\delta = x^{k+1} - x^k$ ,  $\gamma = \nabla f(x^{k+1}) - \nabla f(x^k)$ ,
       $v = \left(\gamma^T H^k \gamma\right)^{\frac{1}{2}} \left(\frac{\delta}{\gamma^T \delta} - \frac{H^k \gamma}{\gamma^T H^k \gamma}\right)$ ;
7   Choose  $\phi_k \geq 0$ ;
8   Compute  $H^{k+1} = H^k + \frac{\delta \delta^T}{\gamma^T \delta} - \frac{H^k \gamma \gamma^T H^k}{\gamma^T H^k \gamma} + \phi_k v v^T$ ;
9   Set  $k = k + 1$ ;
10 end
```

$$\bar{B} = B + \frac{\gamma \gamma^T}{\delta^T \gamma} - \frac{B \delta \delta^T B}{\delta^T B \delta}. \quad (38)$$

The above formulas are known as (inverse and direct) BFGS updating formulas from Broyden [11], Fletcher [24], Goldfarb [30] and Shanno [57] that discovered them.

Starting from the DFP and BFGS updating formulas a whole class (known as Broyden class) of updating methods can be constructed:

$$\bar{H}^\phi = (1 - \phi) \bar{H}^{\text{DFP}} + \phi \bar{H}^{\text{BFGS}}$$

where \bar{H}^{DFP} and \bar{H}^{BFGS} are given by (34) and (37) respectively, and $\phi \in \mathbb{R}$.

Note that \bar{H}^ϕ can be easily rewritten as

$$\bar{H}^\phi = \bar{H}^{\text{DFP}} + \phi v v^T$$

where

$$v = \left(\gamma^T H \gamma\right)^{\frac{1}{2}} \left(\frac{\delta}{\gamma^T \delta} - \frac{H \gamma}{\gamma^T H \gamma}\right)$$

clearly showing the relationship between \bar{H}^ϕ and \bar{H}^{DFP} .

The generic Quasi–Newton algorithm using the Broyden updating scheme is reported in Algorithm 5 (here we return to use again the index k).

For quadratic strictly convex problems

$$\min f(x) := \frac{1}{2}x^T Mx + q^T x$$

where $M \in \mathbb{R}^{n \times n}$ is a positive definite matrix and $q \in \mathbb{R}^n$, finite termination is guaranteed when $\phi_k \geq 0$ and α_k is obtained via exact line-search [17, Theorem 8.1], [59, Theorem 5.2.1]. Finite termination follows from the observation that for all $k = 0, \dots, n$

$$\left(x^{k+1} - x^k\right)^T M \left(x^{i+1} - x^i\right) = 0, \quad \forall i < k$$

i.e., the directions $\delta^0 = x^1 - x^0, \delta^1 = x^2 - x^1, \dots, \delta^k = x^{k+1} - x^k$ are M -conjugate and x^{k+1} minimizes $f(x)$ in the affine space the affine set

$$S_k := \left\{ x \in \mathbb{R}^n : x = x^0 + \sum_{j=0}^k \sigma_j \delta^j, \sigma_j \in \mathbb{R} \right\}$$

The following theorem, originally due to Powell, shows global convergence for rank-2 symmetric updating schemes for generic convex optimization problems.

Theorem 13 ([17, Theorem 8.2]) *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be twice differentiable and convex. Let $x^0 \in \mathbb{R}^n$ and assume the level set \mathcal{L}_0 be bounded. Let H^0 be symmetric and positive definite and let $\{x^k\}$ be generated by Algorithm 5 with either*

- $\phi_k = 0$ (DFP updating formula) and exact line-search, or
- $\phi_k = 1$ (BFGS updating formula) and inexact line-search satisfying the Wolfe condition.

Then, for any $\epsilon > 0$ there exists k such that $\|\nabla f(x^k)\| < \epsilon$.

Dixon [22] demonstrated that in case of exact line-search, the sequence $\{x^k\}$ is independent of $\phi_k \geq 0$. Furthermore, in [24], it is shown that, for stability reasons, it is better to choose $\phi_k \in [0, 1]$. However, from a computational point of view the choice $\phi_k = 1$ appears better than $\phi_k = 0$.

Finally, q -superlinear convergence of rank-2 updating methods is obtained in the case $\phi_k = 0$ or $\phi_k = 1$ [17, Theorem 8.9] if the stepsize α_k is chosen according to the Wolfe rule (in which case it is possible to show that $\alpha_k = 1$

for $k \geq k_0$) provided that $\sum_{k=1}^{+\infty} \|x^k - x^*\| < +\infty$.

We refer the interested reader to [35, 59] for additional insight on Quasi-Newton methods.

4 Constrained Optimization

In this section we concentrate our attention on the constrained optimization problem

$$\begin{aligned} \min_x \quad & f(x) \\ \text{*subject to} \quad & c_i(x) \leq 0 \quad i = 1, \dots, m \\ & c_i(x) = 0 \quad i = m + 1, \dots, m + h \end{aligned} \tag{39}$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is the objective function and $c_i : \mathbb{R}^n \rightarrow \mathbb{R}, i = 1, \dots, m + h$ are the constraints defining the feasible region X

$$X := \{x \in \mathbb{R}^n : c_i(x) \leq 0, i = 1, \dots, m \text{ and } c_i(x) = 0, i = m + 1, \dots, m + h\}.$$

For simplicity, we will always assume that all the functions are at least twice continuously differentiable.

Definition 11 A feasible point \bar{x} is a local minimum if $\exists \gamma > 0$ such that

$$f(\bar{x}) \leq f(x) \quad \forall x \in X \text{ and } \|x - \bar{x}\| \leq \gamma.$$

More generally, a point \bar{x} is a local minimum if $\exists \mathcal{I}$ neighborhood of \bar{x} such that

$$f(\bar{x}) \leq f(x) \quad \forall x \in X \cap \mathcal{I}.$$

A point \bar{x} is a strict local minimum if $\exists \mathcal{I}$ neighborhood of \bar{x} such that

$$f(\bar{x}) < f(x) \quad \forall x \in X \cap \mathcal{I}, x \neq \bar{x}.$$

A point \bar{x} is a strong or isolated local minimum if $\exists \mathcal{I}$ neighborhood of \bar{x} such that \bar{x} is the only local minimum in $X \cap \mathcal{I}$.

For the above minimization problem, the Lagrangian function is defined as follows:

$$L(x, \lambda) := f(x) + \sum_{i=1}^{m+h} \lambda_i c_i(x) \tag{40}$$

where $\lambda \in \mathbb{R}^{m+h}$. First and second order necessary and sufficient optimality conditions will be expressed in terms of the above Lagrangian function.

Another important concept in constrained optimization is the concept of active constraints.

Definition 12 Let \bar{x} be a feasible point for the constrained optimization problem (39), i.e., $\bar{x} \in X$. The set of active constraints at \bar{x} is

$$\mathcal{A}(\bar{x}) := \left\{ i : 1 \leq i \leq m : c_i(\bar{x}) = 0 \right\} \cup \left\{ m+1, \dots, m+h \right\}.$$

4.1 Necessary and Sufficient Optimality Conditions

In order to derive optimality conditions for the constrained optimization problem (39), two sets must be introduced: the tangent cone and the set of linearized feasible directions.

Definition 13 Let $\bar{x} \in X$. A direction $d \in \mathbb{R}^n$ is tangent to the set X at \bar{x} if there exists a sequence $\{z^k\} \subseteq X$ with $\lim_k z^k = \bar{x}$ and a sequence $\{\theta_k\}$ of positive scalars with $\lim_k \theta_k = 0$ such that

$$\lim_k \frac{z^k - \bar{x}}{\theta_k} = d. \quad (41)$$

The set of all tangent vectors to X at \bar{x} is called the tangent cone at \bar{x} and is denoted by $\mathcal{T}_X(\bar{x})$.

Definition 14 Let $\bar{x} \in X$ and let $\mathcal{A}(\bar{x})$ be the set of indices of active constraints. The set of linearized feasible directions $\mathcal{F}(\bar{x})$ is the set

$$\mathcal{F}(\bar{x}) := \left\{ d \in \mathbb{R}^n : \begin{array}{l} \nabla c_i(\bar{x})^T d \leq 0, \quad i = 1, \dots, m, i \in \mathcal{A}(\bar{x}), \\ \nabla c_i(\bar{x})^T d = 0, \quad i = m+1, \dots, m+h \end{array} \right\}.$$

Clearly, if $d \in \mathcal{F}(\bar{x})$ also $\alpha d \in \mathcal{F}(\bar{x})$ for all nonnegative α and hence also the set $\mathcal{F}(\bar{x})$ is a cone.

These two sets can be viewed as local approximations of the feasible region X around \bar{x} .

Note that, while the tangent cone only depends on the geometry of the feasible region, the set of linearized feasible directions depends on the specific formulation of the constraints utilized to define the feasible region. Moreover, for all $\bar{x} \in X$, $\mathcal{T}_X(\bar{x}) \subseteq \mathcal{F}(\bar{x})$.

A fundamental question is under which conditions the two sets coincide, that is $\mathcal{T}_X(\bar{x}) = \mathcal{F}(\bar{x})$ or, to be more precise, under which conditions the dual cones of the two sets coincide.³ Since the seminal work of Kuhn and Tucker [42], a number of different (inter-related) Constraint Qualification

³ Give a cone K , the dual cone of K is the set

$$K^* := \left\{ d \in \mathbb{R}^n : d^T x \geq 0, \forall x \in K \right\}$$

conditions have been proposed in literature both for the case of only inequality constraints and for the more general case of equality of inequality constraints (see [60] and references therein, including the two schemes showing the relationships between the different Constraint Qualification conditions, and [46]).

In [60] the various Constraint Qualification conditions are partitioned in 4 different levels with weakest (less stringent) conditions being at level 1 and strongest (more stringent) conditions at level 4. In the following, we state some of the most common Constraint Qualification conditions.

Definition 15 (*Guignard’s Constraint Qualification, GCQ*) [36] Let $\bar{x} \in X$. The Guignard’s Constraint Qualification conditions are satisfied at \bar{x} if $\mathcal{F}(\bar{x}) = \text{cl}(\text{conv}(\mathcal{T}_X(\bar{x})))$.⁴

The GCQ is considered the weakest possible Constraint Qualification.

Definition 16 (*Abadie’s Constraint Qualification, ACQ*) [1] Let $\bar{x} \in X$. The Abadie’s Constraint Qualification conditions are satisfied at \bar{x} if $\mathcal{T}_X(\bar{x}) = \mathcal{F}(\bar{x})$.

The Abadie’s Constraint Qualification requires that $\mathcal{T}_X(\bar{x})$ be a convex cone. Hence this condition is stronger than Guignard’s Constraint Qualification.

Definition 17 (*Slater’s Constraint Qualification, SCQ*) [58] Let $\bar{x} \in X$. The Slater’s Constraint Qualification conditions are satisfied at \bar{x} if

- $c_i(x)$ is pseudo-convex at \bar{x} for all $1 \leq i \leq m$ and $i \in \mathcal{A}(\bar{x})$,
- $c_i(x)$ is both quasi-convex and quasi-concave⁵ at \bar{x} for all $m + 1 \leq i \leq m + h$,
- the vectors $\{\nabla c_i(\bar{x}), m + 1 \leq i \leq m + h\}$ are linearly independent,
- there exists \hat{x} such that $c_i(\hat{x}) < 0$, for all $1 \leq i \leq m$ and $i \in \mathcal{A}(\bar{x})$ and $c_i(\hat{x}) = 0$ for all $m + 1 \leq i \leq m + h$.

Definition 18 (*Mangasarian–Fromovitz’s Constraint Qualification, MFCQ*) [47, 3.4–3.6] Let $\bar{x} \in X$. The Mangasarian–Fromovitz’s Constraint Qualification conditions are satisfied at \bar{x} if

- the vectors $\{\nabla c_i(\bar{x}), m + 1 \leq i \leq m + h\}$ are linearly independent,
- there exists d such that $\nabla c_i(\bar{x})^T d < 0$, for all $1 \leq i \leq m$ and $i \in \mathcal{A}(\bar{x})$ and $\nabla c_i(\bar{x})^T d = 0$ for all $m + 1 \leq i \leq m + h$.

⁴ For a set S , $\text{cl}(S)$ indicates its closure.

⁵ A generic function g is quasi-concave if and only if $-g$ is quasi-convex.

Definition 19 (*Linear Independence Constraint Qualification, LICQ*) [60, 5.4.6] Let $\bar{x} \in X$. Linear Independence Constraint Qualification (LICQ) conditions are satisfied at \bar{x} if the vectors

$$\left\{ \nabla c_i(\bar{x}), 1 \leq i \leq m, \quad i \in \mathcal{A}(\bar{x}) \right\} \cup \left\{ \nabla c_i(\bar{x}), m+1 \leq i \leq m+h \right\}$$

are linearly independent.

Both LICQ and SCQ imply MFCQ. Moreover, MFCQ implies ACQ. Therefore, if any of these Constraint Qualification conditions hold at a feasible point \bar{x} , then $\mathcal{T}_X(\bar{x}) = \mathcal{F}(\bar{x})$.

The following lemma provides necessary conditions for a local optimal solution.

Lemma 1 *Let x^* be a local optimal solution of Problem (39). Then*

$$\nabla f(x^*)^T d \geq 0 \quad \forall d \in \mathcal{T}_X(x^*) \quad (42)$$

Proof Suppose, by contradiction, that at x^* there is a direction $d \in \mathcal{T}_X(x^*)$ such that

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

Then, there exists a sequence $\{d^k\}$ converging to d and a sequence $\{\theta_k\}$ of positive scalars converging to zero with $z^k = x^* + \theta_k d^k \in X$ for all k . Therefore,

$$\begin{aligned} f(z^k) &= f(x^* + \theta_k d^k) = f(x^*) + \nabla f(x^*)^T (z^k - x^*) + o(\|z^k - x^*\|) \\ &= f(x^*) + \theta_k \nabla f(x^*)^T d^k + o(\theta_k). \end{aligned}$$

But

$$\lim_k \nabla f(x^*)^T d^k = \nabla f(x^*)^T d < 0,$$

and hence

$$f(z^k) - f(x^*) = \theta_k \nabla f(x^*)^T d^k + o(\theta_k) < 0$$

for k sufficiently large. Therefore, for each neighborhood of x^* , there exists a sufficiently large index k such that the point z^k belongs to such neighborhood and, hence, x^* is not a local minimum.

In case that Constraint Qualification conditions are satisfied, the necessary optimality conditions can be expressed in a more convenient way, in terms of the Karush–Kuhn–Tucker conditions. This results utilizes Motzkin's theorem of the alternative, see [46, Chap. 2].

Theorem 14 ([46, 7.3.7][49, Theorem 12.1]) *Let x^* be a local minimum for the constrained optimization Problem (39) and assume that at x^* some Constraint Qualification conditions are satisfied. Then there exists a Lagrange multiplier vector λ^* such that the following conditions, known as Karush–Kuhn–Tucker (or KKT) conditions, hold:*

$$\nabla_x L(x^*, \lambda^*) = 0, \tag{43a}$$

$$c_i(x^*) \leq 0 \quad \forall i = 1, \dots, m, \tag{43b}$$

$$c_i(x^*) = 0 \quad \forall i = m + 1, \dots, m + h, \tag{43c}$$

$$\lambda_i^* \geq 0 \quad \forall i = 1, \dots, m, \tag{43d}$$

$$\lambda_i^* c_i(x^*) = 0 \quad \forall i = 1, \dots, m, \tag{43e}$$

where $L(., .)$ is the Lagrangian function (40).

Note that Condition 43a can be rewritten as

$$\nabla f(x^*) + \sum_{i=1}^{m+h} \lambda_i^* \nabla c_i(x^*) = 0.$$

Conditions (43e) are called complementarity conditions and they impose that for all $i = 1, \dots, m$ either $\lambda_i^* = 0$ or $c_i(x^*) = 0$ or both.

In order to derive second order necessary and sufficient conditions, we need to introduce first the concept of critical cone.

Definition 20 Let x^* be a local minimum for the constrained optimization Problem (39) and suppose that the pair (x^*, λ^*) satisfies the Karush–Kuhn–Tucker conditions (43) for some $\lambda^* \in \mathbb{R}^{m+h}$. The Critical Cone is defined as:

$$C(x^*, \lambda^*) := \left\{ d \in \mathcal{F}(x^*) : \nabla c_i(x^*)^T d = 0, \right. \\ \left. \text{for all } i = 1, \dots, m, i \in \mathcal{A}(\bar{x}) \text{ with } \lambda_i^* > 0 \right\}. \tag{44}$$

Equivalently, $d \in C(x^*, \lambda^*)$ if only if

$$\nabla c_i(x^*)^T d = 0 \quad i = 1, \dots, m, i \in \mathcal{A}(\bar{x}) \text{ with } \lambda_i^* > 0 \tag{45a}$$

$$\nabla c_i(x^*)^T d \leq 0 \quad i = 1, \dots, m, i \in \mathcal{A}(\bar{x}) \text{ with } \lambda_i^* = 0 \tag{45b}$$

$$\nabla c_i(x^*)^T d = 0 \quad i = m + 1, \dots, m + h \tag{45c}$$

Note that the Hessian of the Lagrangian function $L(x, \lambda)$ is given by

$$\nabla_{xx}^2 L(x, \lambda) = \nabla^2 f(x) + \sum_{i=1}^{m+h} \lambda_i \nabla^2 c_i(x).$$

This Hessian is of fundamental importance in the second order necessary and sufficient conditions.

Theorem 15 ([25, Theorem 9.3.1] [49, Theorem 12.5]) *Let x^* be a local minimum for the constrained optimization Problem (39) and assume that at x^* some Constrained Qualification Conditions are satisfied. Then, there exists a Lagrange multiplier vector λ^* such that the Karush–Kuhn–Tucker conditions (43) are satisfied and*

$$d^T \nabla_{xx}^2 L(x^*, \lambda^*) d \geq 0, \text{ for all } d \in C(x^*, \lambda^*). \quad (46)$$

Theorem 16 ([25, Theorem 9.3.2] [49, Theorem 12.6]) *Let x^* be a feasible point for the constrained optimization Problem (39) and assume that there exists a vector λ^* such that the pair (x^*, λ^*) satisfies the Karush–Kuhn–Tucker conditions (43). Furthermore, suppose that*

$$d^T \nabla_{xx}^2 L(x^*, \lambda^*) d > 0, \text{ for all } 0 \neq d \in C(x^*, \lambda^*). \quad (47)$$

Then, x^ is a strict local minimum of the constrained optimization Problem (39)*

As for the unconstrained case, also here there is a significative difference between necessary and sufficient optimality conditions. In fact, for necessary optimality condition the Hessian of the Lagrangian must be positive semidefinite in the Critical Cone. Instead, for sufficient optimality condition, the Hessian of the Lagrangian must be positive definite in the same Critical Cone.

4.2 Duality in Constrained Optimization

The concept of duality is central in constrained optimization as well as in other fields of Operations Research (e.g., in Linear programming) and, in general, in mathematics. Let

$$f^* = \inf_{x \in X} f(x). \quad (48)$$

The dual function is given by

$$q(\lambda) := \inf_x L(x, \lambda) \tag{49}$$

where $L(x, \lambda)$ is the Lagrangian function (40). Note that this dual function is a concave function (i.e., $-q(\lambda)$ is a convex function) and its domain \mathcal{D} is a convex set. The dual problem is defined as

$$\begin{aligned} & \max_{\lambda} && q(\lambda) \\ & \text{subject to} && \lambda_i \geq 0, \quad i = 1, \dots, m. \end{aligned} \tag{50}$$

Moreover, let

$$q^* = \sup_{\substack{\lambda \\ \lambda_i \geq 0, i = 1, \dots, m.}} q(\lambda) \tag{51}$$

It is not difficult to show that (Weak Duality), if $\bar{\lambda} \in \mathbb{R}^{m+h}$ with $\bar{\lambda}_i \geq 0$, $i = 1, \dots, m$ (that is, $\bar{\lambda}$ is dual feasible), and $\bar{x} \in X$ (that is, \bar{x} is primal feasible), then

$$\begin{aligned} q(\bar{\lambda}) &= \inf_x L(x, \bar{\lambda}) \leq L(\bar{x}, \bar{\lambda}) = f(\bar{x}) + \sum_{i=1}^{m+h} \bar{\lambda}_i c_i(\bar{x}) \\ &= f(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i c_i(\bar{x}) \leq f(\bar{x}) \end{aligned}$$

and hence

$$q^* \leq f^*. \tag{52}$$

A more convenient form for the dual problem can be derived for convex problems. The theorem below shows the relationship between points satisfying Karush–Kuhn–Tucker conditions and optimal solutions of the dual problem. Moreover, it shows that in convex optimization there is no duality gap.

Theorem 17 ([49, Theorem 12.12]) *Let x^* be an optimal solution of the convex optimization problem (39) where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $c_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, m$ are convex functions and $c_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = m + 1, \dots, m + h$ are linear functions. Assume, further, that at x^* some Constraint Qualification conditions are satisfied. Then, there exists λ^* such that the pair (x^*, λ^*) satisfies the Karush–Kuhn–Tucker conditions (43) and λ^* solves the dual problem (50), i.e.,*

$$q(\lambda^*) = \max_{\lambda} q(\lambda) \quad (53)$$

$$\lambda_i \geq 0, i = 1, \dots, m.$$

Moreover,

$$f^* = f(x^*) = L(x^*, \lambda^*) = q(\lambda^*) = q^*$$

In order to revert, at least partially, this condition, strict convexity is required.

Theorem 18 ([49, Theorem 12.13]) *Let x^* be an optimal solution of the convex optimization problem (39) where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $c_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, m$ are convex functions and $c_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = m + 1, \dots, m + h$ are linear functions. Assume, further, that at x^* some Constraint Qualification conditions are satisfied. Moreover, suppose that $\bar{\lambda}$ is a solution of the dual problem (50) and that the function $L(x, \bar{\lambda})$ is strict convex in x and*

$$L(\bar{x}, \bar{\lambda}) = \inf_x L(x, \bar{\lambda}).$$

Then $x^* = \bar{x}$ and $f(x^*) = L(x^*, \bar{\lambda})$.

The two theorems above are the basis for a different, more convenient, form for the dual problem, known as Wolfe dual [61]:

$$\begin{aligned} & \max_{x, \lambda} && L(x, \lambda) \\ & \text{subject to} && \nabla_x L(x, \lambda) = 0 \\ & && \lambda_i \geq 0, i = 1, \dots, m. \end{aligned} \quad (54)$$

For this dual optimization problem it is possible to show that, if x^* is a local minimum at which Constraint Qualification conditions are satisfied, there exists $\lambda^* \in \mathbb{R}^{m+h}$ such that the pair (x^*, λ^*) satisfies the Karush–Kuhn–Tucker conditions (43) and, furthermore, solves the Wolfe dual problem (54).

4.3 Penalty and Augmented Lagrangian Methods

An important and well studied class of methods for solving nonlinear optimization problems is based on the idea of replacing the original constrained problem by a single or a sequence of unconstrained problems. For these problems the new objective function will contain terms penalizing the violation of the original constraints.

An important issue, both from a theoretical and a practical point of view, is to determine if the minimizer of the penalty function and the solution of

the original optimization problem coincide. This property is called exactness of the penalty function [20].

The simplest approach is to use a quadratic term to penalize the violation of the constraints. For the constrained problem (39), the quadratic penalty function is given by

$$\psi(x, \mu) := f(x) + \mu \left[\sum_{i=1}^m \max\{0, c_i(x)\}^2 + \sum_{i=m+1}^{m+h} c_i(x)^2 \right], \quad (55)$$

where $\mu > 0$ is the penalty parameter. If the original functions $f(x)$ and $c_i(x)$, $i = 1, \dots, m + h$ are sufficiently smooth, the function $\psi(x, \mu)$ is differentiable (and continuously differentiable, if $m = 0$, i.e., there are only equality constraints) and, hence, standard algorithms for unconstrained minimization can be applied to calculate its minimizer.

However, when quadratic penalty terms are utilized, it is necessary to drive the penalty term to $+\infty$. A general scheme require to:

- choose a priori a sequence $\{\mu_k\} \rightarrow +\infty$,
- for each value of μ_k , calculate x^{μ_k} , a minimizer of $\psi(x, \mu_k)$.

The procedure terminates when the violation of the constraints at x^{μ_k} is sufficiently small.

For this simple scheme it is possible to show [25, Theorem 12.1.1] [49, Theorem 17.1] that

1. $\{\psi(x^{\mu_k}, \mu_k)\}$ is non-decreasing,
2. $\{f(x^{\mu_k})\}$ is non-increasing,
3. the constraints violation is non-increasing,
4. every accumulation point x^* of the sequence $\{x^{\mu_k}\}$ is a solution of Problem (39).

Another widely used penalty function is the 1-norm penalty function defined as

$$\phi(x, \mu) := f(x) + \mu \left[\sum_{i=1}^m \max\{0, c_i(x)\} + \sum_{i=m+1}^{m+h} |c_i(x)| \right]. \quad (56)$$

This penalty function is exact in the sense that there exists a finite $\mu^* > 0$ such that for all values of $\mu \geq \mu^*$, if x^* is a strict local solution of the nonlinear problem (39) at which first-order necessary optimality conditions are satisfied, then x^* is a local minimizer of $\phi(x, \mu)$ [49, Theorem 17.3]. However, the 1-norm penalty function is non-smooth and, therefore, specific algorithms must be utilized and convergence is in many cases quite slow.

Continuously differentiable exact penalty functions can also be constructed [20] by using an Augmented Lagrangian function that includes additional terms penalizing the violation of the Karush–Kuhn–Tucker conditions. Under specific assumptions, stationary point, local and global minimizers of the Augmented Lagrangian function exactly correspond to Karush–Kuhn–Tucker points, local and global solutions of the constrained problem [21].

4.4 Sequential Quadratic Programming

A very effective method for solving constrained optimization problems is based on sequentially solving quadratic subproblems (Sequential Quadratic Programming, SQP). The idea behind the SQP approach is to construct, at each iteration, a quadratic approximation of Problem (39) around the current point x^k , and then use the minimizer of this subproblem as the new iterate x^{k+1} [49, Chap. 18].

More specifically, at the current point x^k we construct a quadratic approximation of the problem by a quadratic approximation for the objective function using the Hessian of the Lagrangian function with respect to the x variables and linear approximation of the constraints:

$$\begin{aligned} \min_d \quad & f(x^k) + \nabla f(x^k)^T d + \frac{1}{2} d^T \nabla_{xx}^2 L(x^k, \lambda^k) d \\ \text{*subject to} \quad & \nabla c_i(x^k)^T d + c_i(x^k) \leq 0, \quad i = 1, \dots, m \\ & \nabla c_i(x^k)^T d + c_i(x^k) = 0, \quad i = m + 1, \dots, m + h. \end{aligned} \tag{57}$$

Very fast and efficient algorithms exist for solving the above problem that produce the vector d^k , multipliers associated to the linearized constraints λ^{k+1} , and an estimate of the active constraints.⁶ Then the new point is obtained as $x^{k+1} = x^k + d^k$.

Under specific assumptions, it is possible to show that if x^* is a local solution of Problem (39), if, at x^* and some λ^* , Karush–Kuhn–Tucker conditions are satisfied, and if (x^k, λ^k) is sufficiently close to (x^*, λ^*) , then there is a local solution of the subproblem (57) whose active set is the same as the active set of the nonlinear optimization problem (39) at x^* [55].

⁶ One of the most important technique for solving convex quadratic programming problems with equality and inequality constraints is based on Active Set strategy where, at each iteration, some of the inequality constraints, and all the equality constraints, are imposed as equalities (the “Working Set”) and a simpler quadratic problem with only equality constraints is solved. Then the Working Set is update and a new iteration is performed. For further details refer to [25, 10.3] and [49, 16.5].

The correct identification of the active constraints when x^k is sufficiently close to x^* is at the basis of the proof of local convergence for the SQP method.

Acknowledgements The author wants to express his gratitude to prof. Nadaniela Egidi for reading a first version of the manuscript and providing many useful suggestions.

References

1. Abadie, J.: On the Kuhn-Tucker Theorem. Operations Research Center University of Calif Berkeley, Technical report (1966)
2. Al-Baali, M.: Descent property and global convergence of the Fletcher' Reeves method with inexact line search. IMA J. Numer. Anal. **5**(1), 121–124 (1985)
3. Apostol, T.M.: Calculus, 2nd edn. Wiley (1967)
4. Armijo, L.: Minimization of functions having Lipschitz continuous first partial derivatives. Pac. J. Math. **16**(1), 1–3 (1966)
5. Bagirov, A.M., Gaudioso, M., Karmita, N., Mäkelä, M.M., Taheri, S.: Numerical Nonsmooth Optimization: State of the Art Algorithms. Springer Nature (2020)
6. Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear Programming: Theory and Algorithms, 3rd edn. Wiley, Newark, NJ (2006)
7. Bertsekas, D.P.: Nonlinear Programming. Athena Scientific (1999)
8. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, New York, NY, USA (2004)
9. Broyden, C.G.: A class of methods for solving nonlinear simultaneous equations. Math. Comput. **19**, 577–593 (1965)
10. Broyden, C.G.: Quasi-Newton methods and their application to function minimization. Math. Comput. **21**, 368–381 (1967)
11. Broyden, C.G.: The convergence of single-rank quasi-Newton methods. Math. Comput. **24**, 365–382 (1970)
12. Broyden, C.G., Dennis, J.E., Moré, J.J.: On the local and superlinear convergence of quasi-Newton methods. IMA J. Appl. Math. **12**(3), 223–245 (1973)
13. Conn, A.R., Gould, N.I.M., Toint, P.L.: Convergence of Quasi-Newton matrices generated by the symmetric rank one update. Math. Program. **50**(2), 177–195 (1991)
14. Davidon, W.C.: Variable metric method for minimization. AEC Research and Development Report ANL-5990, Argonne National Laboratory (1959)
15. De Leone, R., Gaudioso, M., Grippo, L.: Stopping criteria for linesearch methods without derivatives. Math. Program. **30**(3), 285–300 (1984)
16. Deb, K., Deb, K.: Multi-objective optimization. In: Burke, E., Kendall, G. (eds.) Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques, pp. 403–449. Springer, US, Boston, MA (2014)
17. Dennis, J.E., Moré, J.J.: Quasi-Newton methods, motivations and theory. SIAM Rev. **19**, 46–89 (1977)
18. Dennis, J.E., Schnabel, R.B.: Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Prentice-Hall, Englewood Cliffs, New Jersey (1983)
19. Dennis, J.E., Schnabel, R.B.: Chapter I A view of unconstrained optimization. In: Optimization. Handbooks in Operations Research and Management Science, vol. 1, pp. 1–72. Elsevier (1989)

20. Di Pillo, G., Grippo, L.: Exact penalty functions in constrained optimization. *SIAM J. Control Optim.* **27**(6), 1333–1360 (1989)
21. Di Pillo, G., Liuzzi, G., Lucidi, S., Palagi, L.: An exact augmented Lagrangian function for nonlinear programming with two-sided constraints. *Comput. Optim. Appl.* **25**(1), 57–83 (2003)
22. Dixon, L.C.W.: Variable metric algorithms: necessary and sufficient conditions for identical behavior of nonquadratic functions. *J. Optim. Theory Appl.* **10**, 34–40 (1972)
23. Ehrgott, M.: *Multicriteria Optimization*, vol. 491. Springer Science & Business Media (2005)
24. Fletcher, R.: A new approach to variable metric algorithms. *Comput. J.* **13**, 317–322 (1970)
25. Fletcher, R.: *Practical Methods of Optimization*, 2nd edn. Wiley (1990)
26. Fletcher, R., Powell, M.J.D.: A rapidly convergent descent method for minimization. *Comput. J.* **6**(2), 163–168 (1963)
27. Fletcher, R., Reeves, C.M.: Function minimization by conjugate gradients. *Comput. J.* **7**(2), 149–154 (1964)
28. Gaudioso, M., Giallombardo, G., Miglionico, G.: Essentials of numerical nonsmooth optimization. *4OR* **18**(1), 1–47 (2020)
29. Gilbert, J.C., Nocedal, J.: Global convergence properties of conjugate gradient methods for optimization. *SIAM J. Optim.* **2**(1), 21–42 (1992)
30. Goldfarb, D.: A family of variable metric updates derived by variational means. *Math. Comput.* **24**, 23–26 (1970)
31. Goldstein, A.A.: *Constructive Real Analysis*. Harper and Row, London (1967)
32. Golub, G.H., Van Loan, C.F.: *Matrix Computations*. Johns Hopkins University Press, Baltimore, Maryland (1983)
33. Grippo, L., Lampariello, F., Lucidi, S.: A nonmonotone line search technique for Newton’s method. *SIAM J. Numer. Anal.* **23**(4), 707–716 (1986)
34. Grippo, L., Sciandrone, M.: Nonmonotone globalization techniques for the Barzilai-Borwein gradient method. *Comput. Optim. Appl.* **23**(2), 143–169 (2002)
35. Grippo, L., Sciandrone, M.: *Methods of unconstrained optimization. (Metodi di ottimizzazione non vincolata.)*. Unitext 53. La Matematica per il 3+2. Springer, Italia (2011)
36. Guignard, M.: Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space. *SIAM J. Control* **7**(2), 232–241 (1969)
37. Hager, W.W.: Updating the inverse of a matrix. *SIAM Rev.* **31**(2), 221–239 (1989)
38. Hager, W.W., Zhang, H.C.: A survey of nonlinear conjugate gradient methods. *Pac. J. Optim.* **2**(1), 35–58 (2006)
39. Hestenes, M.R., Stiefel, E.: Methods of conjugate gradients for solving linear systems. *J. Res. Natl. Bur. Stand.* **49**, 409–436 (1952)
40. Horst, R., Pardalos, P.M.: *Handbook of Global Optimization*, vol. 2. Springer Science & Business Media (2013)
41. Konak, A., Coit, D.W., Smith, A.E.: Multi-objective optimization using genetic algorithms: a tutorial. *Reliab. Eng. Syst. Saf.* **91**(9), 992–1007 (2006)
42. Kuhn, H.W., Tucker, A.W.: *Nonlinear Programming*, pp. 481–492. University of California Press, Berkeley, California (1951)
43. Lemarechal, C., Mifflin, R.: *Nonsmooth optimization: proceedings of a IIASA workshop, March 28–April 8, 1977*. Elsevier (2014)

44. Li, D.H., Fukushima, M.: A derivative-free line search and global convergence of Broyden-like method for nonlinear equations. *Optim. Methods Softw.* **13**(3), 181–201 (2000)
45. Luenberger, D.G.: *Linear and Nonlinear Programming*, 2nd edn. Addison–Wesley (1984)
46. Mangasarian, O.L.: *Nonlinear Programming*. McGraw-Hill, New York (1969)
47. Mangasarian, O.L., Fromovitz, S.: The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. *J. Math. Anal. Appl.* **17**, 37–47 (1967)
48. Meyer, C.D.: *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA (2000)
49. Nocedal, J., Wright, S.: *Numerical Optimization*. Springer Science & Business Media (2006)
50. Ortega, J.M., Rheinboldt, W.C.: *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press (1970)
51. Pinter, J.: Continuous global optimization: applications. In: C. Floudas, E.P.M. Pardalos (eds.) *Encyclopedia of Optimization*, pp. 482–486. Springer (2008)
52. Polak, E., Ribière, G.: Note sur la convergence de méthodes de directions conjuguées. *ESAIM: Math. Model. Numer. Anal.-Modélisation Mathématique et Anal. Numérique* **3**(R1), 35–43 (1969)
53. Polyak, B.T.: The conjugate gradient method in extremal problems. *USSR Comput. Math. Math. Phys.* **9**(4), 94–112 (1969)
54. Pytlak, R.: *Conjugate Gradient Algorithms in Nonconvex Optimization*, vol. 89. Springer Science & Business Media (2008)
55. Robinson, S.M.: Perturbed Kuhn-Tucker points and rates of convergence for a class of nonlinear-programming algorithms. *Math. Program.* **7**(1), 1–16 (1974)
56. Rockafellar, R.T.: *Convex Analysis*. Princeton Mathematical Series, Princeton University Press, Princeton, N. J. (1970)
57. Shanno, D.F.: Conditioning of Quasi-Newton methods for function minimization. *Math. Comput.* **24**, 647–656 (1970)
58. Slater, M.: Lagrange multipliers revisited. Technical report, Cowles Foundation Discussion Paper No. 80, Cowles Foundation for Research in Economics, Yale University (1950)
59. Sun, W., Yuan, Y.X.: *Optimization Theory and Methods: Nonlinear Programming*, vol. 1. Springer Science & Business Media (2006)
60. Wang, Z., Fang, S.C., Xing, W.: On constraint qualifications: motivation, design and inter-relations. *J. Ind. Manag. Optim.* **9**, 983–1001 (2013)
61. Wolfe, P.: A duality theorem for non-linear programming. *Q. Appl. Math.* **19**, 239–244 (1961)
62. Wolfe, P.: Convergence conditions for ascent methods. *SIAM Rev.* **11**(2), 226–235 (1969)
63. Zhang, H.C., Hager, W.W.: A nonmonotone line search technique and its application to unconstrained optimization. *SIAM J. Optim.* **14**(4), 1043–1056 (2004)