[1] Опис лиценци Креативне заједнице доступан је на адреси creativecommons.org.rs/?page_id=74.

UNIVERSITY OF NOVI SAD
FACULTY OF SCIENCES
DEPARTMENT OF MATHEMATICS
AND INFORMATICS

Ana Friedlander **Nataša Krejić**

Nataša Krklec Jerinkić

Lectures on Fundamentals of Numerical Optimization

Novi Sad, 2019

Fundamentals of Numerical Optimization
First edition

Authors
Prof. Dr. Ana Friedlander
University of Campinas, SP, Brasil

Prof. Dr. Nataša Krejić
University of Novi Sad Faculty of Sciences

Dr. Nataša Krklec Jerinkić
Asst. Prof. University of Novi Sad Faculty of Sciences

Reviewers
Academician Prof. Dr. Gradimir Milovanović
Serbian Academy of Sciences and Arts, Mathematical Institute

Dr. Sanja Rapajić
Assoc. Prof. University of Novi Sad Faculty of Sciences

Dr. Dušan Jakovetić
Asst. Prof. University of Novi Sad Faculty of Sciences

Publisher
University of Novi Sad Faculty of Sciences

Editor-in-Chief
Prof. Dr. Milica Pavkov-Hrvojević, Dean of University of Novi Sad Faculty of Sciences

Publication of this textbook is approved by the Council of Faculty of Sciences on January 17, 2019.
ISBN: 978-86-7031-474-0

CIP - Каталогизација у публикацији
Библиотека Матице српске, Нови Сад

519.6(075.8)

FRIEDLANDER, Ana
Lectures on fundamentals of numerical optimization [Elektronski izvor] / Ana Friedlander, Nataša Krejić, Nataša Krklec Jerinkić. - Novi Sad : Faculty of Sciences, 2019

Način dostupa (URL): https://www.pmf.uns.ac.rs/studije/epublikacije/matinf/friedlander_krejic_krklecjerinkic_lec tures_fundamentals_numerical_optimization.pdf . - Opis zasnovan na stanju na dan 14.2.2019. - Nasl. s naslovnog ekrana. - Bibliografija.

ISBN 978-86-7031-474-0
1. Krejić, Nataša [аутор] 2. Krklec-Jerinkić, Nataša [аутор]
а) Нумеричка анализа
COBISS.SR-ID 328208135

Foreword

This book covers the content of Fundamentals of Numerical Optimization course at the Master program Applied Mathematics - Data Science at Department of Mathematics and Informatics, Faculty of Sciences. It provides insights into some of the main methods of nonlinear optimization through a sequence of theoretical considerations, algorithms and exercises. The subject is too wide to be covered within one book and thus the content of the book is mainly determined by the course. Our intention was to present the material at a very accessible level assuming only the undergraduate mathematical background.

The book covers some topics of unconstrained and constrained optimization problems and leans heavily on the book "Elementos de programação não-linear" by Ana Friedlander. The main advantage of the aforementioned book is the set of exercises designed to facilitate understanding of the presented material. The set of carefully designed exercises is almost completely taken from that book and in our opinion presents the fundamental tool for understanding the concepts and ideas presented theoretically.

Novi Sad, January 2019 Ana Friedlander, Nataša Krejić, Nataša Krklec Jerinkić

Contents

Chapter 1

Nonlinear optimization problems

We consider problems of the form

$$\min_{x \in S} f(x), \tag{1.1}$$

where $f : D \to \mathbb{R}$ and $D, S \subseteq \mathbb{R}^n$. Vector x is often referred to as the decision variable and its dimension n represents the dimension of the problem. The function f is called the objective function and D represents its domain. We usually have that $D = \mathbb{R}^n$, but it can be $\mathbb{R}_+$ for example if the objective function is $f(x) = \ln(x)$. On the other hand, S is called the feasible set and it is often a true subset of $\mathbb{R}^n$. This set represents the constraints of the optimization problem stated above. If $S = \mathbb{R}^n$ then we say that the problem (1.1) is an unconstrained optimization problem.

If S is a true subset of $\mathbb{R}^n$, then the problem (1.1) is called constrained optimization problem. One of the simplest constraints are the so called box constraints where

$$S = \{x \in \mathbb{R}^n \mid x_i \in [l_i, u_i], i = 1, 2, ..., n\}$$

These constraints are often represented as $x \in [l, u]$ and the components of the vectors l and u are in $\mathbb{R} \bigcup \{\pm\infty\}$. Furthermore, S can be a hyperplane

$$S = \{x \in \mathbb{R}^n \mid x^T a \leq 0\}$$

for some $a \in \mathbb{R}^n$ or some ball in $\mathbb{R}^n$

$$S = \{x \in \mathbb{R}^n \mid \|x - b\| \leq c\},$$

where $b \in \mathbb{R}^n, c \in \mathbb{R}$. In general, the constraints are stated as follows

$$S = \{x \in \mathbb{R}^n \mid h(x) = 0,\ g(x) \leq 0\}, \tag{1.2}$$

where $h : \mathbb{R}^n \to \mathbb{R}^m$ represents the equality constraints and $g : \mathbb{R}^n \to \mathbb{R}^p$ determines the inequality constraints. Both of these constraints are explicit constraints. On the other hand, if the domain of the objective function is a true subset of $\mathbb{R}^n$, then we say that D represents implicit constraints. Implicit constraints can be produced by the domain of h and g as well, as discussed in Chapter 9.

Solving the optimization problem means finding the best feasible decision variable - the one that minimizes the objective function on the feasible set. Notice that maximizing f on S is equivalent to minimizing $-f$ on S so, without loss of generality, we can identify optimization with minimization.

Important question is whether the optimization problem has a solution. The answer is not trivial, but the following theorem states the assumptions under which we are certain that a solution exists.

Theorem 1.1 *(Bolzano-Weierstrass) Every real, continuous function attains its global minimum on any compact subset of* $\mathbb{R}^n$.

The consequence of this theorem is that the problem (1.1) has a solution if the objective function f is continuous and the feasible set S is compact, i.e., bounded and closed. Now, let us define two types of solutions of the considered problem.

Definition 1 *A point x^* is a global solution of the problem (1.1) if $f(x^*) \leq f(x)$ for every $x \in S$. If $f(x^*) < f(x)$ for every $x \in S$, $x \neq x^*$, then x^* is a strict global solution.*

Definition 2 *A point x^* is a local solution of the problem (1.1) if there exists $\varepsilon > 0$ such that $f(x^*) \leq f(x)$ for every $x \in S$ such that $\|x - x^*\| \leq \varepsilon$. If $f(x^*) < f(x)$ for every $x \in S$, $x \neq x^*$ such that $\|x - x^*\| \leq \varepsilon$, then we say that x^* is strict local solution.*

Roughly speaking, a local minimizer is the best feasible point in its own vicinity. On the other hand, a global minimizer is the best point of all feasible points. Finding a global minimizer is usually very hard, so local minimizers are of great interest in nonlinear optimization. In the next chapter we will discuss some basic characterizations of global and local solutions for unconstrained optimization problems. Analogous characterizations for the constrained case will be considered latter.

1.1 Exercises

1. Let $A \in \mathbb{R}^{n\times n}$ and $x \in \mathbb{R}^n$. For each of the following statements provide a proof or a counter example.
   (a) There exists $x \neq 0$ such that $Ax = 0$ if $|A| = 0$.
   (b) There exists $x \neq 0$ such that $Ax = 0$ only if $|A| = 0$.
   (c) There exists $x \neq 0$ such that $Ax = 0$ if and only if $|A| = 0$.
2. Let $A \in \mathbb{R}^{m\times n}$ where $m \geq n$ and $rank(A) = n$. Prove that $A^T A$ is nonsingular.
3. Consider the equations

$$\sum_{j=1}^{n} a_{i,j} x_j = b_i, \quad i = 1, 2, ..., n-1, \tag{1.3}$$

or equivalently, $Ax = b$ where $A \in \mathbb{R}^{(n-1)\times n}$, $b \in \mathbb{R}^{n-1}$ and $x \in \mathbb{R}^n$. The set of points which satisfy (1.3) represents a line in $\mathbb{R}^n$. This line can be written as

$$y = x + \lambda d,$$

where $\lambda \in \mathbb{R}$, $d, x \in \mathbb{R}^n$. Discuss the choice of x and d.

4. Find the eigenvalues and eigenvectors of $A = uu^T$ where $u \in \mathbb{R}^n$.

5. Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. For $x \in \mathbb{R}^n$ we define $q(x) = f(Ax + b)$, where $f : \mathbb{R}^m \to \mathbb{R}$. Calculate the gradient and the Hessian of q.

6. Draw the level sets of the following quadratic functions $f : \mathbb{R}^2 \to \mathbb{R}$.

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

   (b) $f(x_1, x_2) = x_1^2 + x_2^2 + 2x_1x_2$.

   (c) $f(x_1, x_2) = x_1^2 + x_2^2 - x_1x_2$.

   (d) $f(x_1, x_2) = x_1x_2$.

7. Explain the geometry of the level sets for

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

where $A = A^T \in \mathbb{R}^{2\times 2}$, $b \in \mathbb{R}^2$ and $c \in \mathbb{R}$ given that:

   (a) A is positive definite, i.e., $A \succ 0$.

   (b) A is positive semidefinite ($A \succeq 0$) and there exists x such that $Ax + b = 0$.

   (c) $A \succeq 0$ and there is no x such that $Ax + b = 0$.

(d) A is indefinite and nonsingular.

8. Prove that the eigenvalues of a symmetric matrix are positive if and only if the matrix is positive definite.

9. Prove that a symmetric matrix is singular if and only if zero is its eigenvalue.

10. Prove that if λ is an eigenvalue of a nonsingular symmetric matrix A, then λ^{-1} is an eigenvalue of A^{-1}.

11. Let $A \in \mathbb{R}^{m \times n}$ where $m \leq n$ and $rank(A) = k$. Let us denote the null space of A by $Null(A)$ and the subspace of images by $Im(A)$.

    (a) Prove that $Null(A) \perp Im(A^T)$.

    (b) Prove that $dim(Null(A)) = n - k$.

    (c) Prove that $\mathbb{R}^n = Null(A) \oplus Im(A^T)$.

Chapter 2

Optimality conditions for unconstrained problems

Let us consider the unconstrained optimization problem

$$\min_{x \in \mathbb{R}^n} f(x), \tag{2.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$. In this chapter, we are stating necessary and sufficient conditions for a solution of problem (2.1). These conditions are important not only theoretically, but also from practical point of view. Vast majority of methods for solving unconstrained optimization problems are constructed to converge to points that satisfy necessary optimality conditions stated below. On the other hand, once we find a point that satisfies the necessary conditions the sufficient conditions help us to decide whether this candidate is a true minimizer. Depending on the order of derivatives, we distinguish first and second order necessary conditions, while the sufficient conditions are of the second order.

Recall that if the function is one dimensional, i.e., $f : \mathbb{R} \to \mathbb{R}$, then the solution x^* of (2.1) satisfies $f'(x^*) = 0$. Moreover, if the function is twice continuously differentiable, then $f''(x^*) \geq 0$.

Let us state the first order necessary conditions.

Theorem 2.1 *Suppose that $f \in C^1(\mathbb{R}^n)$. If x^* is a local solution of (2.1), then $\nabla f(x^*) = 0$.*

Proof. Suppose that x^* is a local solution of (2.1). Let us consider an arbitrary but fixed direction $d \in \mathbb{R}^n$. Furthermore, define

$$\phi(\lambda) := f(x^* + \lambda d).$$

Since x^* is a local minimizer of f, we know that there exists $\varepsilon > 0$ such that $f(x^*) \leq f(x^* + \lambda d)$ for every λ such that $|\lambda| \leq \varepsilon$. This implies that

$$\phi(0) \leq \phi(\lambda)$$

for every $|\lambda| \leq \varepsilon$. Thus, $\lambda = 0$ is a local minimizer of the function $\phi : \mathbb{R} \to \mathbb{R}$ and we know that $\phi'(0) = 0$. On the other hand,

$$\phi'(\lambda) = \nabla^T f(x^* + \lambda d)d,$$

so we conclude that

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

Since this equality holds for an arbitrary $d \in \mathbb{R}^n$, we conclude that $\nabla f(x^*)$ is orthogonal to every vector $d \in \mathbb{R}^n$ and therefore $\nabla f(x^*) = 0$. ∎

The point which satisfy the first order necessary conditions are called stationary points.

The second order necessary conditions are the following.

Theorem 2.2 *Suppose that $f \in C^2(\mathbb{R}^n)$. If x^* is a local solution of (2.1), then*

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

b) $\nabla^2 f(x^*) \succeq 0$.

Proof. Suppose that x^* is a local solution of (2.1). The first part of the statement follows from the previous theorem. Now, let us consider an arbitrary direction $d \in \mathbb{R}^n$ and define $\phi(\lambda) := f(x^* + \lambda d)$ as in the previous proof. Following the same arguments, we conclude that $\lambda = 0$ is a local minimizer of ϕ. This implies $\phi'(0) = 0$, but it also implies $\phi''(0) \geq 0$. Since the second derivative of ϕ is

$$\phi''(\lambda) = d^T \nabla^2 f(x^* + \lambda d) d,$$

there follows that $d^T \nabla^2 f(x^*) d \geq 0$. As d is an arbitrary vector, we conclude that $d^T \nabla^2 f(x^*) d \geq 0$ for every $d \in \mathbb{R}^n$, i.e., the matrix $\nabla^2 f(x^*)$ is positive semidefinite, which completes the proof. ■

Finally, let us state the second order sufficient conditions. Notice that the second condition is stronger than in the previous theorem since the Hessian is assumed to be positive definite instead of just positive semidefinite. Also, notice that the following theorem states that the local minimizer is strict.

Theorem 2.3 *Suppose that* $f \in C^2(\mathbb{R}^n)$. *If*

1. $\nabla f(x^*) = 0$ *and*
2. $\nabla^2 f(x^*) \succ 0$,

then x^* *is a strict local solution of (2.1).*

Proof. The second condition states that the Hessian of the objective function is positive definite at x^*. Moreover, $f \in C^2(\mathbb{R}^n)$ and the Hessian $\nabla^2 f(x)$ is continuous. Thus, the Hessian matrix remains positive definite in some neighborhood of the point x^*. More precisely, there exists $\varepsilon > 0$ such that $\nabla^2 f(x) \succ 0$ for every $x \in V = \{y \in$

$\mathbb{R}^n \mid \|y - x^*\| < \varepsilon\}$. Hence for any $x \neq x^*$ such that $\|x - x^*\| \leq \varepsilon_1$, with $\varepsilon_1 < \varepsilon$ the Taylor expansion yields

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

for some $t \in (0, 1)$. Notice that $\nabla^2 f(x^* + t(x - x^*)) \succ 0$. Moreover, $\nabla f(x^*) = 0$ and the Taylor's expansion implies that $f(x) > f(x^*)$. Since this inequality holds for any x such that $\|x - x^*\| \leq \varepsilon_1$, we conclude that x^* must be a strict local minimizer of function f. ∎

2.1 Exercises

1. Let $g : \mathbb{R} \to \mathbb{R}$ be strictly increasing and $f : \mathbb{R}^n \to \mathbb{R}$. Prove that minimizing $f(x)$ is equivalent to minimizing $g(f(x))$.

2. Solve
$$\min_{x \in \mathbb{R}^n} \|Ax - b\|,$$
where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Consider all possible cases and give geometrical representation.

3. Let $a_1 \leq a_2 \leq ... \leq a_n$. Solve the following problems

   (a) $\min_{x \in \mathbb{R}^n} \sum_{i=1}^n |x - a_i|$.

   (b) $\min_{x \in \mathbb{R}^n} \max_{i=1,\ldots,n} |x - a_i|$.

   (c) $\min_{x \in \mathbb{R}^n} \sum_{i=1}^n |x - a_i|^2$.

   (d) $\max_{x \in \mathbb{R}^n} \prod_{i=1}^n |x - a_i|$.

4. Consider the Rosenbrock function
$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

(a) Find the first and the second order derivative of f.

(b) Show that $x^* = (1,1)^T$ is a local minimizer.

(c) Prove that $\nabla^2 f(x)$ is singular if and only if $x_2 - x_1^2 = 0.005$.

5. Find the stationary points of

$$f(x) = 2x_1^3 - 3x_1^2 - 6x_1x_2(x_1 - x_2 - 1).$$

Which are minimizers and which maximizers? Which of them are local?

6. Prove that the function

$$f(x) = (x_2 - x_1^2)^2 + x_1^5$$

has one stationary point which is not local minimizer nor maximizer.

7. In order to approximate a function g on $[0,1]$ with a polynomial $p(x) = a_0 + a_1x + ... + a_nx^n$, one minimizes the function

$$f(a) = \int_0^1 (g(x) - p(x))^2 dx,$$

where $a = (a_0, ..., a_n)^T$. Find the optimality conditions.

8. Consider the problem

$$\min_{x \in \mathbb{R}^n} f(x) = x_1^2 - x_1x_2 + 2x_2^2 - 2x_1 + e^{x_1+x_2}.$$

(a) Find the first order optimality conditions. Are they the sufficient conditions as well? Why?

(b) The point $\tilde{x} = (0,0)^T$ is optimal?

(c) Find a direction $\tilde{d} \in \mathbb{R}^2$ such that $\nabla^T f(\tilde{x})\tilde{d} < 0$.

(d) Minimize the function f along the direction $\tilde{d}$ starting from $\tilde{x}$.

9. Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be smooth. Moreover, let $f(x) = \|F(x)\|^2$. Assume that x^* is a local minimizer of f such that the Jacobian $\nabla F(x^*)$ is non-singular. Prove that x^* is a solution of $F(x) = 0$.

10. Let
$$f(x) = (x_1^3 + x_2)^2 + 2(x_2 - x_1 - 4)^4.$$
Given $x \in \mathbb{R}^2$ and $d \in \mathbb{R}^2 \backslash \{0\}$, define $g(\lambda) = f(x + \lambda d)$.

(a) Find $g(\lambda)$ explicitly.

(b) For $x = (0, 0)^T$ and $d = (1, 1)^T$ find a minimizer of g.

11. Assume that
$$f(x) = (x_1 - 1)^2 x_2.$$
Consider the points of the form $\hat{x} = (1, x_2)^T$.

(a) Analyze the optimality conditions for these points.

(b) What can we say about $\hat{x}$ based on (a)?

12. Let
$$f(x) = \frac{1}{2} x^T Q x - b^T x,$$
where $Q = Q^T \in \mathbb{R}^{n \times n}$, $Q > 0$ and $b \in \mathbb{R}^n$. Let $x^0, x^1, ..., x^n \in \mathbb{R}^n$ and define $\delta^j = x^j - x^0$ and $\gamma^j = \nabla f(x^j) - \nabla f(x^0)$ for $j = 1, ..., n$. Prove that if the vectors $\delta^1, ..., \delta^n$ are linearly independent then
$$x^* = x^n - [\delta^1 \mid ... \mid \delta^n] \left([\gamma^1 \mid ... \mid \gamma^n]\right)^{-1} \nabla f(x^n)$$
is a global minimizer of f.

13. The Frobenius norm of matrix $A = [a_{i,j}] \in \mathbb{R}^{m \times n}$ is defined as

$$\|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j}^2}.$$

For given $A \in \mathbb{R}^{n \times n}$ find B which solves

$$\min_{B \in \mathbb{R}^{n \times n}, B^T = B} \|A - B\|_F.$$

14. If possible, find coefficients a and b such that

$$f(x) = x^3 + ax^2 + bx$$

has a local maximum at $x = 0$ and local minimum at $x = 1$.

Chapter 3

Convexity

The concept of convexity is very important in numerical optimization. We distinguish two types of convexity - convexity of function and convexity of set. One of the most important features of convex functions is that every local minimizer is also a global minimizer of that function. On the other hand, if the set is convex then every line segment connecting two points of the set remains within which is essential in vast number of algorithms. We state the definitions below.

Definition 3 *A set $S \subseteq \mathbb{R}^n$ is convex if for any $x, y \in S$ and any $\lambda \in [0,1]$ there holds $\lambda x + (1-\lambda)y \in S$.*

One of the most representative convex set is ball. Box is also an example of convex set. On the other hand, an example of a nonconvex set would be a banana-shaped set. In that case, there is at least one couple of points from that set such that the line segment determined by the points is not entirely within the set, i.e., there exists $\lambda \in (0,1)$ such that $\lambda x + (1-\lambda)y \notin S$. Next, we state the definition of convex function.

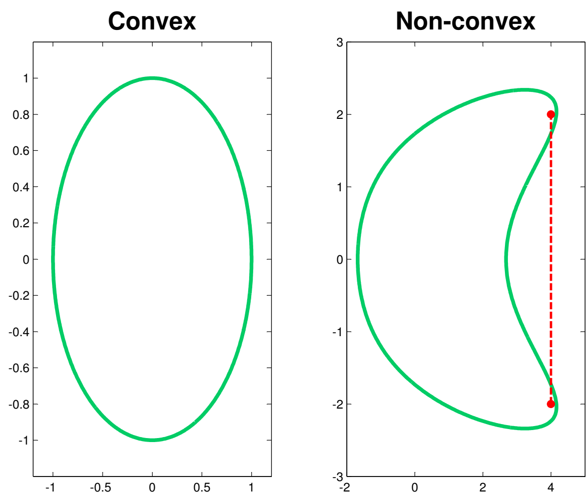

Figure 3.1: Convex and non-convex sets.

Definition 4 *Let S be a convex set. A function $f : S \to \mathbb{R}$ is convex on S if for any $x, y \in S$ and any $\lambda \in [0, 1]$ there holds*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Moreover, we say that the function is strictly convex if the previous inequality is strict for all $x \neq y$ and $\lambda \in (0, 1)$.

For example, the function $f(x) = x^2$ is strictly convex. On the other hand, the function defined as $f(x) = x^2$ for $|x| > 1$ and $f(x) = 1$ for $|x| \leq 1$ is convex, but not strictly convex. We say that a function is concave if the inequalities in the previous definition are opposite, i.e., if we put $\geq$ ($>$) instead of $\leq$ ($<$). In other words, the function f is concave if the function $-f$ is convex and vice versa. Notice that the linear function is both convex and concave at the same time.

One should also notice that the previous definition does not require differentiability of the function f. Moreover, convex function can even be discontinuous. However, if the function f is differentiable, we have the following widely used characterization of convexity. Recall that the directional derivative of function f at point x in direction d is given by

$$\nabla^T f(x)d = \lim_{h \to 0^+} \frac{f(x + hd) - f(x)}{h}.$$

Theorem 3.1 *Suppose that $f \in C^1(S)$ where $S \subseteq \mathbb{R}^n$ is a convex set. Then, the function f is convex on S if and only if the following inequality holds for all $x, y \in S$*

$$f(y) \geq f(x) + \nabla^T f(x)(y - x). \tag{3.1}$$

Proof. First, let us assume that the function f is convex on S. Take arbitrary $x, y \in S$ and $\lambda \in (0, 1)$. Define $z(\lambda) := \lambda y + (1 - \lambda)x$. Since the set S is convex, we know that $z(\lambda) \in S$ for every $\lambda \in [0, 1]$. Now, convexity of the function f implies that

$$f(z(\lambda)) \leq \lambda f(y) + (1 - \lambda)f(x) \tag{3.2}$$

and subtracting $f(x)$ from both sides of the previous inequality we obtain

$$f(z(\lambda)) - f(x) \leq \lambda(f(y) - f(x)).$$

Notice that $z(\lambda) = x + \lambda(y - x)$. Now, dividing the previous inequality by λ we get

$$\frac{f(x + \lambda(y - x)) - f(x)}{\lambda} \leq f(y) - f(x).$$

Taking the limit when $\lambda \to 0^+$ we obtain

$$lim_{\lambda \to 0^+} \frac{f(x + \lambda(y - x)) - f(x)}{\lambda} \leq lim_{\lambda \to 0^+}(f(y) - f(x)).$$

and

$$\nabla^T f(x)(y - x) \leq f(y) - f(x)$$

which is equivalent to (3.1).

Let us assume that (3.1) holds. Again, we take arbitrary $x, y \in S$ and define $z(\lambda)$ as above. Applying (3.1) two times we obtain

$$f(x) \geq f(z(\lambda)) + \nabla^T f(z(\lambda))(x - z(\lambda)) \tag{3.3}$$

and

$$f(y) \geq f(z(\lambda)) + \nabla^T f(z(\lambda))(y - z(\lambda)). \tag{3.4}$$

Notice that $x - z(\lambda) = \lambda(x - y)$ and $y - z(\lambda) = (1 - \lambda)(y - x)$. Now, multiplying (3.3) with $1 - \lambda$ and (3.4) with λ and adding them together we obtain (3.2) which completes the proof. ∎

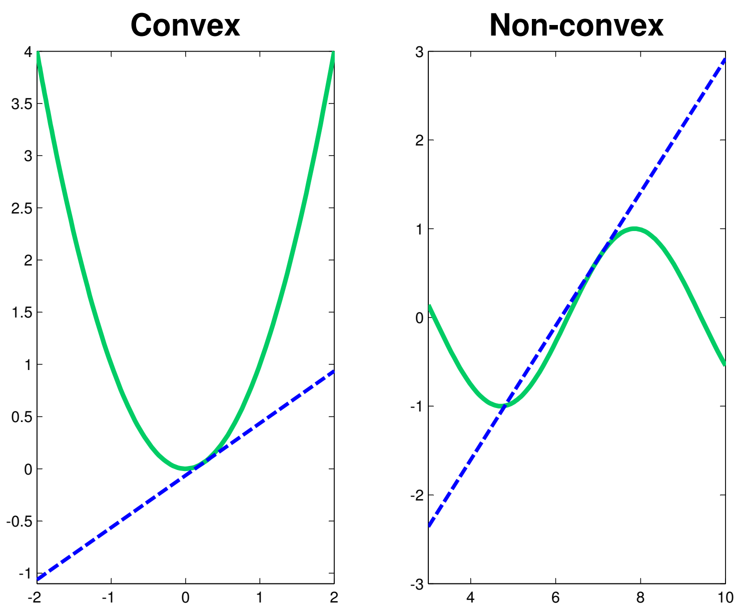

Figure 3.2: Convex and non-convex functions.

The previous characterization implies, for $f : \mathbb{R} \to \mathbb{R}$, that the function is always above the tangent. Following the same ideas as in the second part of the proof of Theorem 3.1, we can prove the following characterization of strictly convex functions.

Theorem 3.2 *Suppose that $f \in C^1(S)$ where $S \subseteq \mathbb{R}^n$ is a convex set. If*

$$f(y) > f(x) + \nabla^T f(x)(y - x)$$

for all $x, y \in S$, $x \neq y$, the function f is strictly convex on S.

Now, we give another important characterization of convex functions. Recall that the Landau small o is defined as

$$\lim_{h \to 0} \frac{o(h)}{h} = 0.$$

Theorem 3.3 *Suppose that $f \in C^2(S)$ where $S \subseteq \mathbb{R}^n$ is a convex set. Then, the following statements hold.*

a) If $\nabla^2 f(x) \succeq 0$ for every $x \in S$, then f is convex on S.

b) If $\nabla^2 f(x) \succ 0$ for every $x \in S$, then f is strictly convex on S.

c) If S is open and f is convex on S, then $\nabla^2 f(x) \succeq 0$ for every $x \in S$.

Proof.

a) Assume that $\nabla^2 f(x) \succeq 0$ for every $x \in S$. Let us take arbitrary $x, y \in S$. Then, due to Taylor's expansion, there exists $z \in S$ such that

$$f(y) = f(x) + \nabla^T f(x)(y - x) + \frac{1}{2}(y - x)^T \nabla^2 f(z)(y - x).$$

Since $\nabla^2 f(z) \succeq 0$, we obtain $f(y) \geq f(x) + \nabla^T f(x)(y - x)$ and the function f is convex according to Theorem 3.1.

b) Using the same arguments as in part 1. we obtain $f(y) > f(x) + \nabla^T f(x)(y-x)$ and Theorem 3.2 implies that the function f is strictly convex.

c) Let us take an arbitrary $x \in S$ and an arbitrary $d \in \mathbb{R}^n$. Since S is open, there exists $\bar{h} > 0$ such that $x + hd$ remains in S for every $0 \leq h \leq \bar{h}$. Furthermore, we have

$$f(x+hd) = f(x) + h\nabla^T f(x)d + \frac{1}{2}h^2 d^T \nabla^2 f(x)d + o(h^2\|d\|^2).$$

Since the function f is assumed to be convex on S, it follows from Theorem 3.1 that $f(x+hd) \geq f(x) + h\nabla^T f(x)d$ and therefore,

$$\frac{1}{2}h^2 d^T \nabla^2 f(x)d + o(h^2\|d\|^2) \geq 0.$$

This inequality holds for any h small enough, so dividing the previous inequality by $h^2\|d\|^2$ and letting $h \to \infty$ we obtain that $d^T\nabla^2 f(x)d$ must be nonnegative. The vector d is arbitrary and we conclude that $\nabla^2 f(x) \succeq 0$. Finally, the point x is an arbitrary point from S and we conclude that $\nabla^2 f(x) \succeq 0$ for every $x \in S$.

■

Next, we prove an important property of convex functions.

Theorem 3.4 *Suppose that f is convex on a convex set S. Then, every local minimizer of the function f is also the global minimizer.*

Proof. We prove this statement by contradiction. Let us assume that x^* is a local, but not global minimizer of f. Then, there exists y^* such that $f(y^*) < f(x^*)$. Moreover, the convexity implies that for any $\lambda \in (0,1)$ we have

$$f(x^*+\lambda(y^*-x^*)) = f(\lambda y^*+(1-\lambda)x^*) \leq \lambda f(y^*)+(1-\lambda)f(x^*) < f(x^*).$$

Therefore, one can always find a small enough λ, i.e., the point $z = x^* + \lambda(y^* - x^*)$ in an arbitrary small vicinity of the point x^* such that $f(z) < f(x^*)$ which is in contradiction with the assumption that x^* is a local minimum. ∎

We give two more varieties of convex functions.

Definition 5 *A function f is strongly convex with parameter $m > 0$ on a convex set S if for any $x, y \in S$ and any $\lambda \in [0,1]$ there holds*

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y) - m\frac{1}{2}\lambda(1-\lambda)\|x-y\|^2.$$

If the function f is differentiable, then one can state another characterization of a strongly convex function. If the following holds for every $x, y \in S$

$$f(y) \ge f(x) + \nabla^T f(x)(y-x) + \frac{m}{2}\|x-y\|^2,$$

then the function f is strongly convex with parameter $m > 0$ on a convex set S. Notice that the strong convexity implies the strict convexity. Moreover, if $f \in C^2(S)$, then the function is strongly convex on S with parameter $m > 0$ if $\nabla^2 f(x) \succeq mI$ for every $x \in S$. In other words, the minimal eigenvalue of the Hessian is uniformly bounded away from zero on S.

Definition 6 *A function f is quasi-convex on a convex set S if for any $x, y \in S$ and any $\lambda \in [0,1]$ there holds*

$$f(\lambda x + (1-\lambda)y) \le \max\{f(x), f(y)\}.$$

Every convex function is also a quasi-convex. On the other hand, concave function can be quasi-convex, for example $f(x) = \ln x$. An example of a function which is not quasi-convex is $f(x) = \sin x$.

3.1 Exercises

1. Show that the intersection of convex sets is a convex set.

2. Prove that $S = \{x \in \mathbb{R}^n \mid \|x\| \leq c\}$ is a convex set.

3. Verify that the following functions are convex.

   (a) $f(x) = \max\{g(x), h(x)\}$, where g and h are convex.

   (b) $t(x) = \sum_{i=1}^n x_i^2$.

   (c) $s(x) = \exp(f(x))$, where $f : \mathbb{R}^n \to \mathbb{R}$ is convex and $f \in C^2(\mathbb{R}^n)$.

4. Sketch the level sets of convex functions. What kind of property do they have? Prove that property.

5. Let S be a convex subset of $\mathbb{R}^n$ and

   $$f(y) = \min_{x \in S} \|y - x\|.$$

   This function is convex. Prove this statement for the special case when

   $$S = \{x \in \mathbb{R}^2 \mid ax_1 + bx_2 = c\}.$$

   Provide a geometrical representation.

6. Prove Theorem 3.2.

7. Suppose that f is convex on a convex set S. Prove that the set of minimizers of f is a convex set.

8. Suppose that $f \in C^1(S)$ is convex on a convex set S. If $x^* \in S$ is such that for every $y \in S$

   $$\nabla^T f(x^*)(y - x^*) \geq 0,$$

   then x^* is a global minimizer of f on S. Prove this statement.

9. Find an example of a strictly convex function with the Hessian which is not positive definite, but it is only positive semidefinite.

10. Suppose that $f \in C^2(S)$, S is convex and $\nabla^2 f(x) \succeq mI$ for every $x \in S$. Prove that the function f is strongly convex.

11. A function f is called quasi-concave on a convex set S if for any $x, y \in S$ and any $\lambda \in [0, 1]$ there holds

$$f(\lambda x + (1 - \lambda)y) \geq \min\{f(x), f(y)\}.$$

Show that a monotone function $f : \mathbb{R} \to \mathbb{R}$ is both quasi-convex and quasi-concave.

Chapter 4

Convergence rates

In the subsequent sections we will consider numerical algorithms for solving (1.1). More precisely, we will analyze iterative procedures that generate a sequence $\{x^k\}_{k\in\mathbb{N}}$ of approximate solutions of the problem under consideration. Let us denote a solution of the problem (1.1) by x^*. Sometimes, x^* represents only a local minimizer or just a stationary point. Now, let us assume that the sequence generated by the algorithm converges to that point, i.e.,

$$\lim_{k\to\infty} x^k = x^*.$$

This means that, in general, we need an infinite sequence to reach x^* and for a finite k, the point x^k will be in some neighborhood of the solution for k large enough. Since we can perform only finitely many iterations in practical applications, the velocity of approaching to x^* is highly relevant. Therefore, besides the convergence itself, the convergence rate is one of the key issues for numerical methods. We state the most important definitions bellow.

Definition 7 *Let* $\lim_{k\to\infty} x^k = x^*$. *Then the sequence* $\{x^k\}_{k\in\mathbb{N}}$ *converges to* x^*

a) linearly, if there exists $\mu \in (0,1)$ *such that*

$$\lim_{k\to\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = \mu;$$

b) superlinearly, if

$$\lim_{k\to\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 0;$$

c) with order $p > 1$, *if there exists* $M > 0$ *such that*

$$\lim_{k\to\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|^p} = M.$$

The linear convergence defined above is often referred to as Q-linear convergence. Also, it is often stated as follows: $\{x^k\}_{k\in\mathbb{N}}$ converges to x^* linearly if there exists $\mu \in (0,1)$ such that

$$\|x^{k+1} - x^*\| \leq \mu\|x^k - x^*\|$$

for all k large enough. On the other hand, superlinear (or equivalently, Q-superlinear) convergence is often characterized throughout the sequence $\{\mu_k\}_{k\in\mathbb{N}}$ such that

$$\|x^{k+1} - x^*\| \leq \mu_k\|x^k - x^*\|$$

and $\lim_{k\to\infty} \mu_k = 0$.

One of the most famous special cases of order p convergence is the quadratic (that is, Q-quadratic) convergence - when $p = 2$. So, according to the definitions above, $\{x^k\}_{k\in\mathbb{N}}$ converges to x^* quadratically if there exists an arbitrary large positive constant M such that

$$\|x^{k+1} - x^*\| \leq M\|x^k - x^*\|^2,$$

for k large enough. We say that the convergence rate is cubic if $p = 3$ and so on.

Superlinear convergence implies linear convergence. Therefore, we can say that a sequence that converges superlinearly is faster than a sequence that converges linearly. Moreover, quadratic convergence implies superlinear convergence, cubic convergence implies quadratic and so on.

There are also convergence rates that are inferior to linear. Some of them are stated below.

Definition 8 *Let* $\lim_{k\to\infty} x^k = x^*$. *Then the sequence* $\{x^k\}_{k\in\mathbb{N}}$ *converges to* x^* *sublinearly if*

$$\lim_{k\to\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 1.$$

Definition 9 *Let* $\lim_{k\to\infty} x^k = x^*$. *Then the sequence* $\{x^k\}_{k\in\mathbb{N}}$ *converges to* x^* *R-linearly if*

$$\|x^{k+1} - x^*\| \leq a_k,$$

where $\{a_k\}_{k\in\mathbb{N}}$ *converges to zero Q-linearly.*

4.1 Exercises

1. Define the approximation error as

$$e_k := \|x^k - x^*\|.$$

   Plot the error sequences $\{e_k\}_{k\in\mathbb{N}}$ of the sequences $\{x^k\}_{k\in\mathbb{N}}$ that converge sublinearly, linearly, superlinearly and quadratically.

2. Prove that $\{x^k\}_{k\in\mathbb{N}}$ converges to x^* R-linearly if and only if there exist $B > 0$ and $\rho \in (0, 1)$ such that
$$\|x^k - x^*\| \leq B\rho^k.$$

3. Prove that superlinear convergence implies linear convergence.

4. Prove that quadratic convergence implies superlinear convergence.

5. Prove that superlinear convergence does not depend on the norm.

6. Does linear convergence in one norm imply linear convergence in another norm? Explain.

Chapter 5

Line search methods

Let us consider an unconstrained optimization problem where the objective function f is continuously differentiable. Assume that we have a point x which is not stationary for function f, i.e., $\nabla f(x) \neq 0$. According to the analysis from Chapter 2, we conclude that x is not a minimizer of f and there exists a point y in a vicinity of point x such that $f(y) < f(x)$. This implies that there is a vector $d \in \mathbb{R}^n$ such that $f(x + \alpha d) < f(x)$ for some scalar $\alpha > 0$. The main idea behind the line search method lies in this fact. First, we find a suitable direction d from point x. Then, we search along that line to find a suitable length α. The direction d is usually called a search direction, while the length α is called a step size.[1] Having all this in mind, we define a descent search direction as follows.

Definition 10 *Consider a point x such $\nabla f(x) \neq 0$. A direction d is called descent direction for f at the point x if there exists $\alpha > 0$ such*

[1] Another framework which is widely used is the so called Trust region framework where, within the current iteration, the step size constraint is given first, while the direction is determined latter throughout solving a constrained subproblem. For further introduction, see [16] for instance.

that

$$f(x + \alpha d) < f(x).$$

If f is continuously differentiable, we can state a characterization of descent directions which is frequently used.

Theorem 5.1 *Suppose that $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^1(\mathbb{R}^n)$, and $x \in \mathbb{R}^n$ is such that $\nabla f(x) \neq 0$. Moreover, suppose that the direction d satisfies the following inequality*

$$\nabla^T f(x) d < 0. \tag{5.1}$$

Then, there exists $\bar{\alpha}$ such that $f(x + \alpha d) < f(x)$ for all $\alpha \in (0, \bar{\alpha}]$.

Proof. Let us define $\phi(\alpha) := f(x + \alpha d)$. Then (5.1) implies that $\phi'(0) < 0$. On the other hand,

$$\phi'(0) = \lim_{\alpha \to 0^+} \frac{\phi(\alpha) - \phi(0)}{\alpha}$$

and thus $(\phi(\alpha) - \phi(0))/\alpha < 0$ for all positive α small enough, i.e., there exists $\bar{\alpha}$ such that $f(x + \alpha d) < f(x)$ for all $\alpha \in (0, \bar{\alpha}]$. ■

Notice that the previous theorem states that d is a descent direction for function f at point x if (5.1) holds. Therefore, (5.1) is often used as the main indicator of the descent property.

Now, we state the model algorithm of the line search method.

Algorithm 5.1

Step 0 Input parameters: $x^0 \in \mathbb{R}^n$.

Step 1 Initialization: $k = 0$.

Step 2 Stopping criterion: If $\nabla f(x^k) = 0$ STOP. Otherwise go to Step 3.

Step 3 Search direction: Choose d^k such that $\nabla^T f(x^k)d^k < 0$.

Step 4 Step size: Find $\alpha_k > 0$ such that $f(x^k + \alpha_k d^k) < f(x^k)$.

Step 5 Update: Set $x^{k+1} = x^k + \alpha_k d^k$, $k = k + 1$ and go to Step 2.

Notice that the Algorithm 5.1 stops only if it encounters a stationary point of the function f. Otherwise it generates an infinite sequence of iterates $\{x^k\}_{k\in\mathbb{N}}$ such that $f(x^{k+1}) < f(x^k)$. However, in general, this is not enough to claim that the sequence $\{x^k\}_{k\in\mathbb{N}}$ converges to a minimizer of f. For example, consider $f(x) = x^2$ and define $x^k = 1 + 1/k$. It is easy to see that $x^* = 0$ is the only minimizer and the only stationary point of the considered function. Also, there holds $f(x^{k+1}) < f(x^k)$ for every k, but $\{x^k\}_{k\in\mathbb{N}}$ tends to 1 which is not even a stationary point (see Figure 5.1).

Moreover, the sequence of iterates does not have to be convergent. For instance, consider $f(x) = \sin(x)$. If we define the sequence as $x^{2k+1} = 1/(2k+1)$, $x^{2k} = 2\pi + 1/(2k)$ we obtain again that $f(x^{k+1}) < f(x^k)$ but obviously we have 2 accumulation points of that sequence: 0 and 2π (see Figure 5.2).

Even if the sequence is convergent, we have seen that it does not have to converge to a solution. The reasons for that are different, some of them are: too small steps $(x^{k+1} - x^k)$, too large steps or search directions that are nearly orthogonal to the gradient (see Figure 5.3). In the sequel we provide guidance for overcoming some of these issues.

To avoid steps that are too small, we impose the following condition on the search direction

$$\|d^k\| \geq \sigma \|\nabla f(x^k)\|, \tag{5.2}$$

where $\sigma > 0$. On the other hand, in order to provide sufficient decrease and avoid situations with large but unproductive steps, we impose the following condition on the step size

$$f(x^k + \alpha_k d^k) \leq f(x^k) + \eta \alpha_k \nabla^T f(x^k) d^k, \tag{5.3}$$

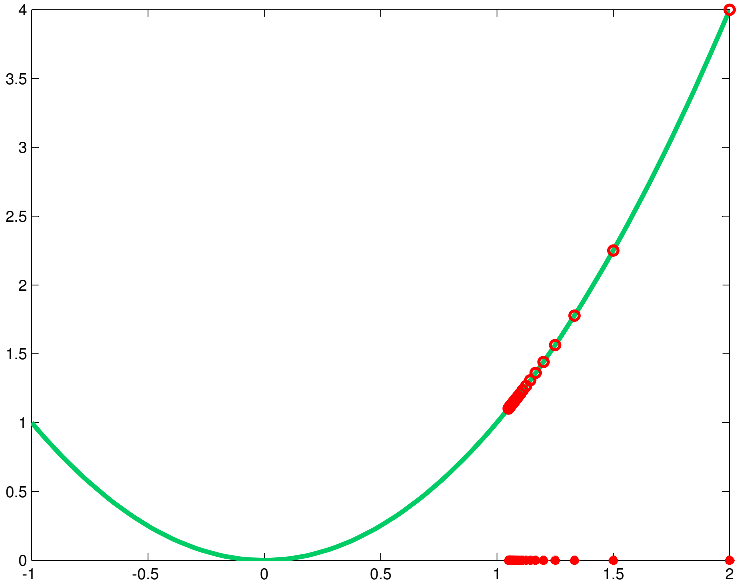

Figure 5.1: Insuficient decrease - small steps.

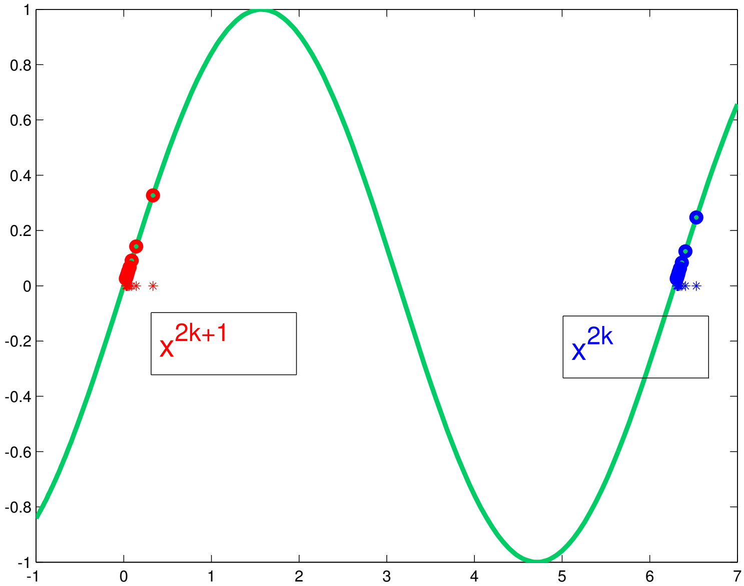

Figure 5.2: Insuficient decrease - large steps.

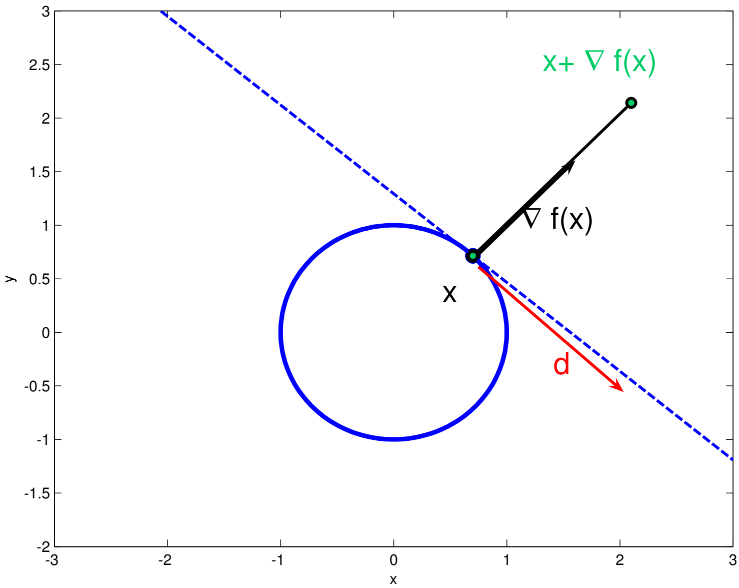

Figure 5.3: Insuficient decrease - insuficiently descent direction.

where $\eta \in (0,1)$ and it is usually set to very small value, for example 10^{-4}. Moreover, we assume that d^k satisfies (5.1), so we obtain the decrease in the function value proportional to scalar product $\nabla^T f(x^k)d^k$. Condition (5.3) is often called the Armijo condition or sufficient decrease condition.

There are many different ways to find a step size that satisfies the Armijo condition. As we will see latter, in some cases we can even find an exact minimizer of $f(x^k + \alpha d^k)$, i.e., the step size α_k which solves the problem

$$\min_{\alpha>0} f(x^k + \alpha d^k). \tag{5.4}$$

In that case we say that the exact line search is performed. However, this does not happen very often since the subproblem (5.4) is not easy to solve in general. Moreover, solving that subproblem exactly is not necessary - an approximate solution which satisfies the Armijo condition does not deteriorate the convergence properties.

One of the common approaches to perform the line search is to use the so called backtracking: starting from $\alpha = 1$, decrease α until the Armijo condition is satisfied. Therefore, the Step 4 of Algorithm 5.1 is often stated as follows.

Step 4.1 Given $\beta \in (0,1)$, find the smallest nonnegative integer j such that $\alpha = \beta^j$ satisfies the Armijo condition (5.3).

If the Armijo condition is satisfied immediately, i.e., with $\alpha = 1$, we say that the full step is accepted. The full step is especially important for Newton methods stated in the Chapter 7.

Now, the question is whether Step 4.1 is well defined, i.e., whether it will be finished in a finite number of trials. The following theorem provides the answer.

Theorem 5.2 *Suppose that $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^1(\mathbb{R}^n)$ and $\nabla^T f(x^k)d^k < 0$. Moreover, assume that the function f is bounded from bellow on the line $\{x^k + \alpha d^k \mid \alpha > 0\}$. Then, there exists $\bar{\alpha} > 0$ such that the Armijo condition holds for all $\alpha \in (0, \bar{\alpha}]$.*

Proof. Define $\phi(\alpha) := f(x^k + \alpha d^k)$ and

$$l(\alpha) := f(x^k) + \eta\alpha\nabla^T f(x^k)d^k = \phi(0) + \alpha\eta\phi'(0).$$

Notice that $l(0) = \phi(0)$ and $l(\alpha)$ is a linear function with the negative slope $\eta\phi'(0)$ so it is unbounded from below. Moreover, since $\eta \in (0, 1)$, there holds $l'(0) > \phi'(0)$. On the other hand, $\phi(\alpha)$ is bounded from below under the stated assumptions and there exists at least one point of intersection of these two functions. Let $\bar{\alpha}$ be the first intersection point, i.e., the smallest α such that $l(\alpha) = \phi(\alpha)$. Then, putting all together we conclude that $\phi(\alpha) \le l(\alpha)$ for all $\alpha \in (0, \bar{\alpha}]$ and thus the Armijo condition is satisfied for all sufficiently small $\alpha > 0$. ■

Finally, in order to avoid search directions that are nearly orthogonal to the gradient, we impose the following condition

$$\nabla^T f(x^k)d^k \le -\theta\|\nabla f(x^k)\|\|d^k\|, \tag{5.5}$$

where $\theta \in (0, 1]$. In some sense, this condition implies that the direction d^k is sufficiently descent. Notice that the previous condition implies that the

$$\cos\angle(d^k, \nabla f(x^k)) \le -\theta.$$

In other words, the angle between d^k and the negative gradient $-\nabla f(x^k)$ is sufficiently sharp.

Now, let us specify further the previously stated algorithm and prove the global convergence result.

Algorithm 5.2

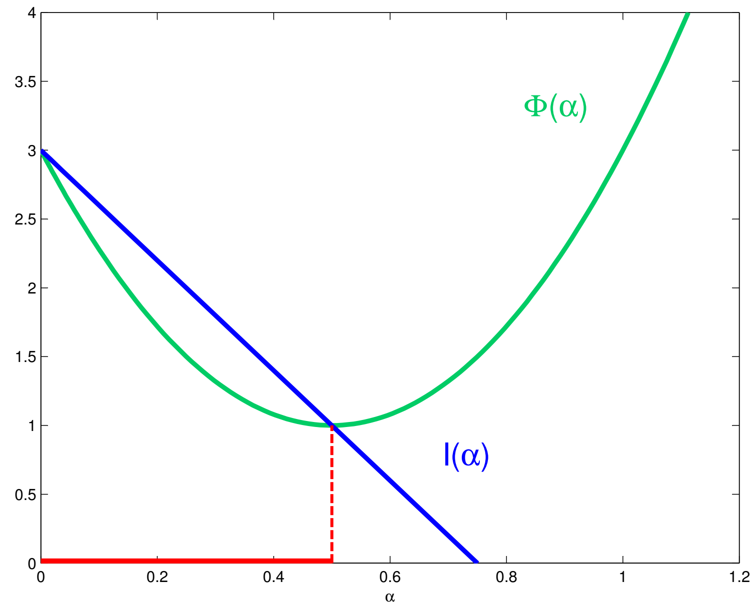

Figure 5.4: Armijo line search - admmisible step sizes.

Step 0 Input parameters: $x^0 \in \mathbb{R}^n$, $\beta, \eta \in (0,1)$, $\theta \in (0,1]$, $\sigma > 0, k = 0$.

Step 1 Stopping criterion: If $\nabla f(x^k) = 0$ STOP. Otherwise go to Step 2.

Step 2 Search direction: Choose d^k such that

$$\|d^k\| \geq \sigma \|\nabla f(x^k)\| \quad \text{and} \quad \nabla^T f(x^k) d^k \leq -\theta \|\nabla f(x^k)\| \|d^k\|.$$

Step 3 Step size: Find the smallest nonnegative integer j such that $\alpha_k = \beta^j$ satisfies the Armijo condition

$$f(x^k + \alpha_k d^k) \leq f(x^k) + \eta \alpha_k \nabla^T f(x^k) d^k.$$

Step 4 Update: Set $x^{k+1} = x^k + \alpha_k d^k$, $k = k + 1$ and go to Step 1.

Theorem 5.3 *Suppose that $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^1(\mathbb{R}^n)$ and f is bounded from bellow. Moreover, assume that the sequence of search directions $\{d^k\}_{k\in\mathbb{N}}$ is bounded. Then, either the Algorithm 5.2 terminates after finite number of iterations $\bar{k}$ at the stationary point $x^{\bar{k}}$ or every accumulation point of the sequence $\{x^k\}_{k\in\mathbb{N}}$ is stationary point of the function f.*

Proof. Algorithm 5.2 terminates only if the stationary point is encountered as stated in Step 1. So, let us consider the case where the number of iterations is infinite. Step 2 implies that for every k we have

$$\nabla^T f(x^k) d^k \leq -\theta \|\nabla f(x^k)\| \|d^k\| \leq -\theta\sigma \|\nabla f(x^k)\|^2. \qquad (5.6)$$

Furthermore, Step 3 yields

$$f(x^{k+1}) \leq f(x^k) - \eta \alpha_k \theta \sigma \|\nabla f(x^k)\|^2.$$

Notice that the Step 3 is well defined since the conditions of Theorem 5.2 are fulfilled. Moreover, applying (5.6) recursively and denoting $c = \eta\theta\sigma$ we obtain

$$f(x^{k+1}) \leq f(x^0) - c\sum_{i=1}^{k} \alpha_i \|\nabla f(x^i)\|^2.$$

Since the function f is assumed to be bounded from bellow, there exists M such that $f(x) \geq M$ for every x and therefore the previous inequality implies

$$c\sum_{i=1}^{k} \alpha_i \|\nabla f(x^i)\|^2 \leq f(x^0) - M.$$

Therefore, $\sum_{k=1}^{\infty} \alpha_k \|\nabla f(x^k)\|^2 < \infty$ and thus

$$\lim_{k\to\infty} \alpha_k \|\nabla f(x^k)\|^2 = 0. \tag{5.7}$$

Now, let x^* be an arbitrary accumulation point of the sequence $\{x^k\}_{k\in\mathbb{N}}$, i.e., there exists $K \subseteq \mathbb{N}$ such that

$$\lim_{k\in K} x^k = x^*.$$

We will show that $\nabla f(x^*) = 0$. If the sequence of step sizes $\{\alpha_k\}_{k\in K}$ is bounded away from zero, then (5.7) implies that $\lim_{k\in K} \|\nabla f(x^k)\|^2 = 0$ and the continuity of ∇f implies that $\|\nabla f(x^*)\| = 0$. So, let us consider the remaining case - assume that there exists $K_1 \subseteq K$ such that $\lim_{k\in K_1} \alpha_k = 0$. Without loss of generality we can assume that $\alpha_k < 1$ for every $k \in K_1$. This means that for every $k \in K_1$ there exists α_k' such that $\alpha_k = \beta\alpha_k'$ and

$$f(x^k + \alpha_k' d^k) > f(x^k) + \eta\alpha_k' \nabla^T f(x^k) d^k. \tag{5.8}$$

Moreover, the Mean value theorem implies that for every $k \in K_1$ there exists $t_k \in (0,1)$ such that

$$f(x^k + \alpha'_k d^k) - f(x^k) = \nabla^T f(x^k + t_k \alpha'_k d^k)\alpha'_k d^k$$

and (5.8) yields

$$\nabla^T f(x^k + t_k \alpha'_k d^k) d^k > \eta \nabla^T f(x^k) d^k$$

for every $k \in K_1$. Recall that the sequence $\{d^k\}_{k\in\mathbb{N}}$ is assumed to be bounded, so there exist $K_2 \subseteq K_1$ and d^* such that $\lim_{k\in K_2} d^k = d^*$. Now, since $\lim_{k\in K_2} \alpha'_k = 0$, taking the limit over K_2 in the previous inequality we get

$$\nabla^T f(x^*) d^* \geq \eta \nabla^T f(x^*) d^*.$$

Since $\eta \in (0,1)$ we conclude that

$$\nabla^T f(x^*) d^* \geq 0. \tag{5.9}$$

On the other hand, for every k there holds $\nabla^T f(x^k) d^k \leq 0$. So, taking the limit over K_2 we obtain

$$\nabla^T f(x^*) d^* \leq 0. \tag{5.10}$$

Combining (5.9) and (5.10) we obtain $\nabla^T f(x^*) d^* = 0$ which together with (5.6) implies

$$0 = \nabla^T f(x^*) d^* = \lim_{k\in K_2} \nabla^T f(x^k) d^k \leq \lim_{k\in K_2} -\theta\sigma\|\nabla f(x^k)\|^2 = -\theta\sigma\|\nabla f(x^*)\|^2$$

and we conclude that $\|\nabla f(x^*)\| = 0$, i.e., x^* must be a stationary point of f. ∎

Notice that the previous theorem can also be proved for the Algorithm 5.2 which takes the following step instead of Step 2.

Step 2' Search direction: Choose d^k such that

$$\nabla^T f(x^k) d^k \leq -m\|\nabla f(x^k)\|^2$$

for some $m > 0$.

5.1 Exercises

1. Let
$$f(x) = \frac{1}{2}x^T Ax + bx + c,$$
where $A \in \mathbb{R}^{n\times n}$, $A^T = A$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$ and let x^* be a local minimizer of f. Prove that x^* is also a global minimizer.

2. If d is a direction such that $\nabla^T f(x)d = 0$, then d can be descent, ascent or neither of these two. Demonstrate this statement.

3. Reformulate the system of nonlinear equations
$$F(x) = 0, \quad F : \mathbb{R}^n \to \mathbb{R}^m$$
into an unconstrained optimization problem.

4. Let
$$f(x) = \frac{1}{2}\|F(x)\|^2, \quad F : \mathbb{R}^n \to \mathbb{R}^n, \quad F \in C^1(\mathbb{R}^n).$$
Suppose that the Jacobian $\nabla F(x)$ is nonsingular for all x and consider an iterative procedure
$$x^{k+1} = x^k - \alpha_k(\nabla F(x^k))^{-1}F(x^k).$$
Prove that the Armijo rule with the parameter $\eta = 0.5$ yields
$$f(x^{k+1}) \leq (1 - \alpha_k)f(x^k).$$

5. Let $f : \mathbb{R} \to \mathbb{R}$, $f \in C^2(\mathbb{R})$, $f'(0) < 0$ and $f''(x) < 0$ for every $x \in \mathbb{R}$. Let $\nu \in (0, 1)$. Prove that the following inequality holds for every $x \geq 0$
$$f(x) \leq f(0) + \nu x f'(0).$$

6. Let $q : \mathbb{R}^n \to \mathbb{R}$ be a convex quadratic function. Assume that we use the exact line search along the descent direction d^k, i.e.,
$$\alpha_k = arg \min_{\alpha>0} q(x^k + \alpha d^k).$$
Show that the step size is given by
$$\alpha_k = \frac{-\nabla^T q(x^k) d^k}{(d^k)^T \nabla^2 q(x^k) d^k}.$$

7. If $f : \mathbb{R}^n \to \mathbb{R}$ is a quadratic function, then $\phi(\alpha) = f(x + \alpha d)$ is parabolic function. Show that the minimizer of $\phi(\alpha)$ is feasible for the Armijo condition if $\eta \in (0, 0.5)$.

8. Let $f : \mathbb{R}^n \to \mathbb{R}$ and suppose $\alpha > 0$ satisfies the Armijo condition. Does an arbitrary $\mu \in (0, \alpha)$ also satisfy the Armijo condition? Prove or provide an counterexample.

9. Suppose that $f : \mathbb{R}^n \to \mathbb{R}, \quad f \in C^2(\mathbb{R}^n), \quad \nabla f(\tilde{x}) = 0$ and $\nabla^2 f(\tilde{x})$ is not positive semidefinite. Prove that then exists a descent direction d at the point $\tilde{x}$.

10. Let x^k be an iterate obtained in optimization process of minimizing function $f : \mathbb{R}^n \to \mathbb{R}, \quad f \in C^1(\mathbb{R}^n)$. Also suppose that x^k is obtained by means of line search with the search direction d^{k-1}. Find a direction d^k such that it is orthogonal to d^{k-1}, descent from x^k and represented as a linear combination of $\nabla f(x^k)$ and d^{k-1}.

11. Let $f : \mathbb{R}^n \to \mathbb{R}, \quad \nabla f(\tilde{x}) \neq 0$ and let $B \in \mathbb{R}^{n\times n}$ be a positive definite matrix. Prove that
$$d = -B \nabla f(\tilde{x})$$
is a descent direction at the point $\tilde{x}$.

Chapter 6

Gradient methods

In this section, we consider a special class of search directions

$$d^k = -\nabla f(x^k). \tag{6.1}$$

Obviously, this direction is a descent direction for function f at point x^k. It is called the steepest descent direction because it minimizes the cosine of the angle between the gradient and the search direction, i.e. it minimizes $\frac{\nabla^T f(x^k) d^k}{\|\nabla^T f(x^k)\| \|d^k\|}$. It is often called simply the negative gradient direction.

The steepest descent direction satisfies inequality (5.2) with $\sigma = 1$. It also satisfies inequality (5.5) with $\theta = 1$. Therefore, in the framework of line search methods, the global convergence result follows directly from Theorem 5.3. However, as we already mentioned, the convergence rate is also an important issue.

Let us start with the strongly convex quadratic objective function

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

where $f : \mathbb{R}^n \to \mathbb{R}$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$ and $A \in \mathbb{R}^{n \times n}$ is symmetric and positive definite. Therefore, there is an unique minimizer x^*.

Moreover, the exact line search along the negative gradient direction produces the step size in the closed form

$$\alpha_k = \frac{\nabla^T f(x^k)\nabla f(x^k)}{\nabla^T f(x^k) A \nabla f(x^k)}. \tag{6.2}$$

Now, denoting the smallest and the largest eigenvalue of A by m and M respectively, we state the following result.

Theorem 6.1 *Let f be a strongly convex quadratic function and $\{x^k\}$ be generated by the steepest descent method with the exact line search. Then*

$$f(x^{k+1}) - f(x^*) \leq \left(\frac{M-m}{M+m}\right)^2 (f(x^k) - f(x^*)).$$

Notice that the previous result states that the convergence is linear with respect functional values $f(x^k) - f(x^*)$ and the following holds

$$\lim_{k\to\infty} f(x^k) = f(x^*).$$

However, this does not imply the standard Q-linear convergence in general. Furthermore, since the solution is unique, the previous result implies that

$$\lim_{k\to\infty} x^k = x^*.$$

In general, that is, if the objective function is not necessarily quadratic, we have the following result.

Theorem 6.2 *Suppose that $f \in C^2(\mathbb{R})$ and that the steepest descent sequence with the exact line search converges to a point x^* such that $\nabla^2 f(x^*)$ is positive definite. Then*

$$f(x^{k+1}) - f(x^*) \leq \left(\frac{M-m}{M+m}\right)^2 (f(x^k) - f(x^*)),$$

where m and M are the smallest and the largest eigenvalue of $\nabla^2 f(x^*)$, respectively.

6.1 Fixed step size

We have already discussed that the exact line search can be very costly in general. This is one of the reasons for introducing fixed step sizes into the negative gradient method, i.e., consider the method which takes fixed α as the step size and the negative gradient as the direction

$$x^{k+1} = x^k - \alpha \nabla f(x^k). \tag{6.3}$$

Assume that the objective function is twice continuously differentiable and its gradient is Lipschitz continuous, i.e., the following holds for all $x, y \in \mathbb{R}^n$

$$\|\nabla f(x) - \nabla f(y)\| \le L\|x - y\|. \tag{6.4}$$

Notice that, under the assumption that $f \in C^2(\mathbb{R}^n)$, the condition (6.4) actually means that the maximal eigenvalues of the Hessian matrix are uniformly bounden from above by L. So, we can prove the convergence result for convex functions. The convergence is ensured only if the step size α is small enough. Moreover, the result stated bellow indicates that the convergence rate is only R-sublinear in terms of $f(x^k)$.

Theorem 6.3 *Suppose that $f \in C^2(\mathbb{R}^n)$ is convex and that (6.4) holds. Then, if $\alpha < 1/L$, the fixed step size negative gradient method defined with (6.3) satisfies*

$$f(x^k) - f(x^*) \le \frac{\|x^0 - x^*\|^2}{2\alpha k}.$$

Proof. Applying the Taylor expansion and using the fact that $\|\nabla^2 f(x)\| \le L$ for every x due to (6.4), we obtain the following inequality for $\alpha < 1/L$

$$f(x^{k+1}) \le f(x^k) - \alpha\|\nabla f(x^k)\|^2 + \frac{L}{2}\alpha^2\|\nabla f(x^k)\|^2 \le f(x^k) - \frac{\alpha}{2}\|\nabla f(x^k)\|^2 \tag{6.5}$$

Now, the convexity of the function f implies

$$f(x^*) \geq f(x^k) + \nabla^T f(x^k)(x^* - x^k),$$

i.e.,

$$f(x^k) \leq f(x^*) + \nabla^T f(x^k)(x^k - x^*). \tag{6.6}$$

Combining (6.6) and (6.5) we obtain

$$f(x^{k+1}) \leq f(x^*) + \nabla^T f(x^k)(x^k - x^*) - \frac{\alpha}{2}\|\nabla f(x^k)\|^2,$$

that is,

$$f(x^{k+1}) - f(x^*) \leq \nabla^T f(x^k)(x^k - x^*) - \frac{\alpha}{2}\|\nabla f(x^k)\|^2. \tag{6.7}$$

Next, we arrange the right hand side of the previous inequality as follows

$$\begin{aligned}
\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2 &= \|x^k - x^*\|^2 - \|x^k - x^* - \alpha \nabla f(x^k)\|^2 \\
&= \|x^k - x^*\|^2 - (\|x^k - x^*\|^2 + \\
&+ \alpha^2\|\nabla f(x^k)\|^2 - 2\alpha(x^k - x^*)^T \nabla f(x^k)) \\
&= 2\alpha\left(\nabla^T f(x^k)(x^k - x^*) - \frac{\alpha}{2}\|\nabla f(x^k)\|^2\right).
\end{aligned}$$

Therefore,

$$\nabla^T f(x^k)(x^k - x^*) - \frac{\alpha}{2}\|\nabla f(x^k)\|^2 = \frac{1}{2\alpha}(\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2). \tag{6.8}$$

Combining the last equality with (6.7) we obtain

$$f(x^{k+1}) - f(x^*) \leq \frac{1}{2\alpha}(\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2). \tag{6.9}$$

Since the previous inequality holds for every $k = 0, 1, ...,$ summing up we get

$$\begin{aligned}
\sum_{k=0}^{s} f(x^{k+1}) - f(x^*) &\leq \frac{1}{2\alpha}\sum_{k=0}^{s}(\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2) \\
&= \frac{1}{2\alpha}(\|x^0 - x^*\|^2 - \|x^{s+1} - x^*\|^2) \\
&\leq \frac{1}{2\alpha}\|x^0 - x^*\|^2.
\end{aligned}$$

Now, letting s tend to infinity and using the fact that $f(x^{k+1}) - f(x^*) \geq 0$ for every k, we conclude that

$$\lim_{k\to\infty} f(x^{k+1}) = f(x^*).$$

Moreover, from (6.5) there follows that $f(x^{s+1}) \leq f(x^s)$ for every s and thus

$$f(x^{s+1}) \leq f(x^k), \quad k = 0, 1, ..., s + 1.$$

Finally,

$$\begin{aligned}
f(x^{s+1}) - f(x^*) &= \frac{1}{s+1}\sum_{k=0}^{s}(f(x^{s+1}) - f(x^*)) \\
&\leq \frac{1}{s+1}\sum_{k=0}^{s}(f(x^{k+1}) - f(x^*)) \\
&\leq \frac{1}{s+1}\left(\frac{1}{2\alpha}\|x^0 - x^*\|^2\right),
\end{aligned}$$

where the last inequality follows from (6.10). ■

6.2 Exercises

1. Suppose that $d^1, ..., d^n$ are linearly independent vectors from $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^1(\mathbb{R}^n)$. Also, suppose that
$$\min_{\lambda\in\mathbb{R}} f(\tilde{x} + \lambda d^j) = f(\tilde{x}), \quad j = 1, ..., n.$$
Prove that $\nabla f(\tilde{x}) = 0$. Does this imply that $\tilde{x}$ is a local minimizer of function f?

2. Let
$$f(x) = \frac{1}{2} x^T Ax + b^T x + c,$$
where $f : \mathbb{R}^n \to \mathbb{R}$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$ and $A \in \mathbb{R}^{n\times n}$ is symmetric and positive definite. Let L_1 and L_2 be two different parallel lines in $\mathbb{R}^n$ with direction d. Let x^1 and x^2 be minimizers of f on L_1 and L_2, respectively. Prove that $(x^2 - x^1)^T Ad = 0$.

3. Let $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^1(\mathbb{R}^n)$. Define an iterative procedure by
$$x^{k+1} = x^k - \lambda_k \nabla f(x^k),$$
where $\lambda_k \geq \lambda > 0$ for every $k \in \mathbb{N}_0$. Suppose that $\lim_{k\to\infty} x^k = \tilde{x}$. Prove that $\nabla f(\tilde{x}) = 0$.

4. Suppose that we have the negative gradient method with the exact line search. Prove that any two consecutive gradients are mutually orthogonal.

5. Suppose that $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^1(\mathbb{R}^n)$. Let y be the point obtained by applying the negative gradient method and the exact line search from a point x. Let z be the point obtained by applying the negative gradient method and the exact line search from the point y. Prove that $z - x$ is a descent direction at the point x.

6. Sketch the level curves of function

$$f(x) = x_1^2 + 4x_2^2 - 4x_1 - 8x_2.$$

Find $\tilde{x}$ that minimizes f. The steepest descent method with the exact line search cannot converge to a point $\tilde{x}$ in finitely many steps if $x^0 = 0$. Is there any x^0 such that the method converges to $\tilde{x}$ in finitely many steps?

7. Consider the negative gradient method for minimizing the strongly convex quadratic function q. Let $\tilde{x}$ be a solution and suppose that

$$x^0 = \tilde{x} + \mu v$$

where v is the eigenvector of $\nabla^2 q(x)$ corresponding to the eigenvalue λ and $\mu \in \mathbb{R}$. Prove that

$$\nabla q(x^0) = \mu\lambda v.$$

If the exact line search is used from x^0, the method converges in one step. Show that the negative gradient method with the exact line search converges in one iteration for every x^0 if $\nabla^2 q(x) = \alpha I$ where $\alpha > 0$ and I is the identity matrix.

8. Suppose that f is quadratic and strongly convex. Prove that if the negative gradient method is applied, starting with a well chosen x^0, such that $\nabla f(x^0) \neq 0$, and finds a minimum in one iteration, then $d = x^1 - x^0$ is an eigenvector of ∇f.

9. For which special class of strongly convex quadratic functions the negative gradient method with the exact line search finds a solution in only one step?

10. For which special class of strongly convex quadratic functions Theorem 6.1 implies the standard Q-linear convergence?

Chapter 7

Newton-like methods

The gradient type methods are widely applicable and relatively easy to implement. However, their rate of convergence is at most linear in general. Therefore, methods that can achieve faster rate of convergence are developed. These methods usually require some information about the Hessian of the objective function. Since the Hessian is the second order derivative, these methods are often called the second order methods.

The most important second order method is the Newton method which assumes that the true Hessian is available at each iteration. Sometimes, the Hessian is unavailable or very hard (costly) to evaluate and the second order derivative is approximated yielding the so called quasi-Newton methods. Moreover, even if the true Hessian is available, obtaining the Newton direction assumes solving a system of linear equations. If the dimension of the problem is large, then this system is large too and it is hard (or costly) to solve it exactly at each iteration. When the considered system is solved only approximately, we are talking about Inexact Newton methods.

There are a lot of modifications of the Newton method. We focus here only on the basic concepts.

7.1 The Newton method

Suppose that we are at an iteration x^k such that $\nabla f(x^k) \neq 0$. Ideally, we would like to find a solution in the next step so we would like to obtain x^{k+1} such that $\nabla f(x^{k+1}) = 0$. Denote $d^k = x^{k+1} - x^k$. Then, using the Taylor expansion we get the following approximation

$$\nabla f(x^{k+1}) \approx \nabla f(x^k) + \nabla^2 f(x^k) d^k.$$

So, instead of searching for x^{k+1} such that the left hand side is equal to zero, we search for d^k such that the right hand side is equal to zero. This way we obtain the famous Newton's equation

$$\nabla f(x^k) + \nabla^2 f(x^k) d^k = 0. \tag{7.1}$$

The vector d^k that satisfies (7.1)is called the Newton step or the Newton direction. Notice that this step is a solution of the system of linear equations

$$\nabla^2 f(x^k) d^k = -\nabla f(x^k).$$

This step is not unique in general, but if the Hessian matrix $\nabla^2 f(x^k)$ is non-singular, then we can express the Newton step by

$$d^k = -(\nabla^2 f(x^k))^{-1} \nabla f(x^k). \tag{7.2}$$

To see the power of the Newton method, let us start with the quadratic objective function.

Theorem 7.1 *Suppose that the function f is quadratic and strongly convex. Then, the Newton method applied from an arbitrary x^0 provides a global minimizer of function f in one iteration.*

Proof. Let us consider an arbitrary starting point x^0 and

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

The objective function is assumed to be strongly convex and thus the Hessian matrix A is positive definite and the Newton direction can be stated as

$$d^0 = -A^{-1}(Ax^0 + b) = -x^0 - A^{-1}b.$$

Now, the next iterate is $x^1 = x^0 + d^0 = -A^{-1}b$ and

$$\nabla f(x^1) = Ax^1 + b = A(-A^{-1}b) + b = 0.$$

Therefore, the point x^1 is stationary for function f and since the objective function is strongly convex we conclude that x^1 is a global minimizer. ∎

Now, let us consider a general case. Notice that the Newton direction does not have to be a descent direction even if the Hessian matrix is non-singular. However, it is a descent direction if the Hessian is positive definite. In that case the inverse Hessian is also positive definite and there holds

$$\nabla^T f(x^k) d^k = -\nabla^T f(x^k)(\nabla^2 f(x^k))^{-1} \nabla f(x^k),$$

which is negative unless x^k is a stationary point. Assume further that the Hessian is uniformly bounded. Then there exists $M > 0$ such that the maximal eigenvalue of each Hessian is bounded form above, i.e., $\lambda_{max}(\nabla^2 f(x)) \leq M$ for every x. This furthermore implies that the eigenvalues of the inverse Hessian matrices are uniformly bounded away from zero, i.e.,

$$\lambda_{min}(\nabla^2 f(x))^{-1}) \geq \frac{1}{M} := m$$

for every x. Therefore,

$$\nabla^T f(x^k) d^k = -\nabla^T f(x^k)(\nabla^2 f(x^k))^{-1} \nabla f(x^k) \leq -m \|\nabla f(x^k)\|^2. \tag{7.3}$$

Using this inequality we can prove the global convergence for the line search Newton method. Indeed, let us consider the following algorithm which differers from Algorithm 5.2 only in Step 2.

Algorithm 7.1

Step 0 Input parameters: $x^0 \in \mathbb{R}^n$, $\beta, \eta \in (0,1), k = 0$.

Step 1 Stopping criterion: If $\nabla f(x^k) = 0$ STOP.

Step 2 Search direction: Compute the Newton direction d^k such that

$$\nabla^2 f(x^k) d^k = -\nabla f(x^k).$$

Step 3 Step size: Find the smallest nonnegative integer j such that $\alpha_k = \beta^j$ satisfies the Armijo condition

$$f(x^k + \alpha_k d^k) \leq f(x^k) + \eta \alpha_k \nabla^T f(x^k) d^k.$$

Step 4 Update: Set $x^{k+1} = x^k + \alpha_k d^k$, $k = k + 1$ and go to Step 1.

Following the same steps as in the proof of Theorem 5.3 we can prove the global convergence stated in the following theorem.

Theorem 7.2 *Suppose that $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^2(\mathbb{R}^n)$ and f is bounded from bellow. Moreover, assume that the Hessian matrices are uniformly bounded and positive definite and that the sequence of search directions $\{d^k\}_{k\in\mathbb{N}}$ is bounded. Then, either the Algorithm 7.1 terminates after a finite number of iterations $\bar{k}$ at the stationary point $x^{\bar{k}}$ or every accumulation point of the sequence $\{x^k\}_{k\in\mathbb{N}}$ is a stationary point of the function f.*

Although the global convergence can be achieved under the stated conditions, Newton method is famous for the local convergence result that we state below. The key property of the Newton method is its quadratic local convergence.

Theorem 7.3 *Suppose that the function $f \in C^2(\mathbb{R}^n)$ and there exists $\delta > 0$ such that $\nabla^2 f(x) \succ 0$ and $\nabla^2 f(x)$ is Lipschitz continuous with the constant L for all $x \in B(x^*, \delta)$. Then there exists $\epsilon > 0$ such that the Newton method converges quadratically to the solution x^* for all $x^0 \in B(x^*, \epsilon)$. Moreover, the sequence of the gradient norms converges quadratically to zero.*

Proof. Since $\nabla^2 f(x) \succ 0$ for all $x \in B(x^*, \delta)$ and there holds[1]

$$\|(\nabla^2 f(x))^{-1}\| \leq 2\|(\nabla^2 f(x^*))^{-1}\| := 2\gamma.$$

Now, let us take an arbitrary $\tau \in (0,1)$ and define

$$\epsilon = \min\{\frac{\tau}{\gamma L}, \delta\},$$

Assume that $x^0 \in B(x^*, \epsilon)$. Then, using the Mean value theorem and the fact that $\nabla f(x^*) = 0$ we obtain

$$\begin{aligned}
& \|x^1 - x^*\| = \|x^0 - (\nabla^2 f(x^0))^{-1} \nabla f(x^0) - x^*\| \\
= \;& \|(\nabla^2 f(x^0))^{-1}(\nabla^2 f(x^0)(x^0 - x^*) - (\nabla f(x^0) - \nabla f(x^*)))\| \\
\leq \;& \|(\nabla^2 f(x^0))^{-1}\| \left\| \int_0^1 \left(\nabla^2 f(x^0)(x^0 - x^*) - \nabla^2 f(x^0 + t(x^0 - x^*))(x^0 - x^*)\right) dt \right\| \\
\leq \;& 2\gamma \int_0^1 \|\nabla^2 f(x^0) - \nabla^2 f(x^0 + t(x^0 - x^*))\| \|x^0 - x^*\| dt \\
\leq \;& 2\gamma \int_0^1 Lt \|x^0 - x^*\|^2 dt \\
= \;& \gamma L \|x^0 - x^*\|^2
\end{aligned}$$

Furthermore,

$$\|x^1 - x^*\| \leq \gamma L \|x^0 - x^*\| \|x^0 - x^*\| \leq \gamma L \epsilon \|x^0 - x^*\| \leq \tau \|x^0 - x^*\| < \|x^0 - x^*\|$$

[1] See Theorem 7.6.

and using the induction arguments we obtain that

$$\|x^{k+1} - x^*\| \le \tau \|x^k - x^*\|, \quad k = 0, 1, ... \tag{7.4}$$

and

$$\|x^{k+1} - x^*\| \le \gamma L \|x^k - x^*\|^2, \quad k = 0, 1, ... \tag{7.5}$$

Since $\tau \in (0,1)$, the inequality (7.4) implies that $\lim_{k\to\infty} x^k = x^*$ and the inequality (7.5) implies that the convergence is quadratic.

Now, using the similar arguments we prove the quadratic convergence of the gradient norms. Indeed,

$$\begin{aligned}
\|\nabla f(x^{k+1})\| &= \|\nabla f(x^{k+1}) - (\nabla f(x^k) + \nabla^2 f(x^k) d^k)\| \\
&= \| \int_0^1 \nabla^2 f(x^k + t d^k) d^k dt - \int_0^1 \nabla^2 f(x^k) d^k dt \| \\
&\le \int_0^1 L t \|d^k\|^2 dt = \frac{L}{2} \|d^k\|^2 \\
&= \frac{L}{2} \| - (\nabla^2 f(x^k))^{-1} \nabla f(x^k)\|^2 \\
&\le \frac{L}{2} \|(\nabla^2 f(x^k))^{-1}\|^2 \|\nabla f(x^k)\|^2 \\
&\le \frac{L}{2} (2\gamma)^2 \|\nabla f(x^k)\|^2 = 2\gamma^2 L \|\nabla f(x^k)\|^2.
\end{aligned}$$

This completes the proof. ∎

The next statement claims that the line search method combines well with the local Newton method, i.e., that the step size eventually becomes equal to 1 and thus the full step is taken if the objective function is strongly convex. Therefore the local quadratic convergence is preserved under the line search.

Theorem 7.4 *Suppose that the conditions of Theorem 7.3 hold and that the function f is strongly convex. Let $\{x^k\}$ be a sequence generated* *by Algorithm 7.1. Then there exists $\eta \in (0,1)$ small enough and $\bar{k} \in \mathbb{N}$ such that $\alpha_k = 1$ for all $k \geq \bar{k}$.*

Proof. Let x^k be generated by Algorithm 7.1. Theorem 7.3 implies that $\lim_{k\to\infty} x^k = x^*$. Given that f is strongly convex, there is $\mu > 0$ such that $\nabla^2 f(x^k) \succeq \mu I$ and thus $(\nabla^2 f(x^k))^{-1} \preceq (\mu)^{-1} I$ for k large enough. The Taylor expansion yields

$$f(x^k + d^k) = f(x^k) + \nabla f(x^k)^T d^k + \frac{1}{2}(d^k)^T \nabla^2 f(\theta^k) d^k,$$

for some $\theta^k \in B(x^k, \|d^k\|)$. Therefore,

$$\begin{aligned} f(x^k + d^k) &= f(x^k) + \nabla^T f(x^k) d^k + \frac{1}{2}(d^k)^T \nabla^2 f(x^k) d^k \\ &+ \frac{1}{2}(d^k)^T \left(\nabla^2 f(\theta^k) - \nabla^2 f(x^k)\right) d^k \end{aligned} \tag{7.6}$$

and

$$\|d^k\| = \left\|(\nabla^2 f(x^k))^{-1} \nabla f(x^k)\right\| \leq \frac{1}{\mu}\|\nabla f(x^k)\|. \tag{7.7}$$

The definition of d^k yields

$$-(d^k)^T \nabla f(x^k) = (d^k)^T \nabla^2 f(x^k) d^k \geq \mu \|d^k\|^2 \tag{7.8}$$

and hence

$$\|d^k\|^2 \leq -\frac{1}{\mu}(d^k)^T \nabla f(x^k). \tag{7.9}$$

Take $\varepsilon > 0$ such that $1 - \varepsilon/\mu \in (0,1)$. Given that $x^k \to x^*$ and $\nabla f(x^*) = 0$, we have $\|d^k\| \to 0$ and $\|\nabla^2 f(\theta^k) - \nabla^2 f(x^k)\| \leq \varepsilon$ for all k large enough. So, (7.6) - (7.9) imply

$$f(x^k + d^k) \leq f(x^k) + \frac{1}{2}(1 - \varepsilon/\mu)(d^k)^T \nabla f(x^k)$$

and the full step is accepted for $\eta \leq \frac{1}{2}(1 - \varepsilon/\mu)$. ■

7.2 Quasi-Newton methods

The Newton method requires evaluation of the second order derivatives at every iteration which can be very costly. In order to avoid these costs, quasi-Newton methods are developed. The main idea is to construct a matrix that approximates the Hessian or the inverse Hessian by updating the previous approximation and using the first order (gradient) information. In general, quasi-Newton methods can achieve only the superlinear rate of convergence, but the cost of obtaining the search direction is significantly smaller then in the Newton method.

Let us denote by B_k the approximation of the Hessian matrix $\nabla^2 f(x^k)$. Then, quasi-Newton direction d^k satisfies

$$B_k d^k = -\nabla f(x^k). \tag{7.10}$$

Assume that we have an approximation B_k and that we performed the iteration to obtain x^{k+1}. Then we need to update the Hessian approximation B_{k+1}. One requirement is that B_{k+1} satisfies the secant equation

$$B_{k+1} s^k = y^k, \tag{7.11}$$

where $s^k = x^{k+1} - x^k$ and

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

The motivation for the secant equation comes from the mean value property

$$y^k = \int_0^1 \nabla^2 f(x^k + ts^k) s^k dt.$$

Thus B_{k+1} aims to approximate the right-hand side of the previous equality, i.e.,

$$B_{k+1} s^k \approx \int_0^1 \nabla^2 f(x^k + ts^k) s^k dt.$$

The secant equation does not determine an unique B_{k+1}. Therefore, besides the symmetry which is a natural requirement for Hessian approximation, other conditions are imposed. The least-change update condition is the most successful one and it states that the next approximation matrix should be as close as possible to the current approximation. Therefore B_{k+1} is a solution of the following optimization problem

$$\min \|B - B_k\| \quad \text{subject to} \quad B^T = B, \quad Bs^k = y^k. \tag{7.12}$$

Clearly, the solution of the above problem depends on the norm we use in the objective function. Two update formulas are the most famous, BFGS and DFP. The BFGS (Broyden-Fletcher-Goldfarb-Shanno) formula is given as

$$B_{k+1} = B_k + \frac{y^k (y^k)^T}{(y^k)^T s^k} - \frac{B_k s^k (s^k)^T B_k}{(s^k)^T B_k s^k} \tag{7.13}$$

The update (7.13) is well defined if the curvature condition $(y^k)^T s^k > 0$ holds. It can be shown that this condition is satisfied if the objective function is strongly convex. Moreover, if B_k is positive definite, curvature condition also implies that B_{k+1} is positive definite as well. Otherwise, if the curvature condition does not hold, a safeguarding is necessary and the common strategy is to skip the update, i.e., to take $B_{k+1} = B_k$.

The DFP (Davidon-Fletcher-Powell) formula is given by

$$B_{k+1} = \left(I - \frac{1}{(y^k)^T s^k} y^k (s^k)^T \right) B_k \left(I - \frac{y^k (s^k)^T}{(y^k)^T s^k} \right) + \frac{y^k (y^k)^T}{(y^k)^T s^k} \tag{7.14}$$

and it has similar properties as BFGS although BFGS is more popular and efficient in general.

Since the search direction is obtained from the quasi Newton linear system, $B_k d^k = -\nabla f(x^k)$, one might be interested in an inverse

Hessian approximation, $H_k \approx (\nabla^2 f(x^k))^{-1}$. The Sherman-Morrison-Woodbury formula[2] provides the updating formula that correspond to the DFP update

$$H_{k+1} = H_k - \frac{H_k y^k (y^k)^T H_k}{(y^k)^T H_k y^k} + \frac{s^k (s^k)^T}{(y^k)^T s^k}. \tag{7.15}$$

Since this is an approximation of the inverse Hessian, i.e. $H_k = B_k^{-1}$, the secant equation becomes

$$H_{k+1} y^k = s^k. \tag{7.16}$$

Analogously, the BFGS (Broyden-Fletcher-Goldfarb-Shanno) formula yields

$$H_{k+1} = \left(I - \frac{s^k (y^k)^T}{(y^k)^T s^k}\right) H_k \left(I - \frac{y^k (s^k)^T}{(y^k)^T s^k}\right) + \frac{s^k (s^k)^T}{(y^k)^T s^k}. \tag{7.17}$$

The most simple approximation of the inverse Hessian is clearly a scalar matrix,

$$H_k = \gamma_k I, \quad \gamma_k \in \mathbb{R} \tag{7.18}$$

Let us consider the secant equation (7.16) and solve it for γ_k,

$$\gamma_k = \arg\min_{\gamma > 0} \|\gamma y^{k-1} - s^{k-1}\|.$$

This problem can be solved analytically and the solution is given by

$$\gamma_k = \frac{(s^{k-1})^T y^{k-1}}{\|y^{k-1}\|^2}. \tag{7.19}$$

Clearly, the coefficient γ_k contains the minimal second order information and the corresponding quasi Newton equation yields the so called spectral gradient direction

$$d_k = -\gamma_k \nabla f(x^k). \tag{7.20}$$

[2] see Exercise 12

Notice that if the curvature condition $(s^{k-1})^T y^{k-1} > 0$ does not hold, then γ_k can be negative and the search direction is not necessary descent. Thus the safeguard

$$\bar{\gamma}_k = \min\{\gamma_{max}, \max\{\gamma_k, \gamma_{min}\}\}$$

is needed with arbitrary values $\gamma_{\min} < \gamma_{\max}$ and in practice it is common to take $0 < \gamma_{min} << 1 << \gamma_{max} < \infty$. Perhaps surprisingly this simple modification of the gradient methods is very efficient in practice.

We conclude this section by providing the general superlinear convergence result for Quasi-Newton methods without specifying the method.

Theorem 7.5 *Suppose that $f \in C^2(\mathbb{R}^n)$. Let $\{x^k\}$ be a sequence generated by a quasi Newton method (7.10) and assume that $\{x^k\}_{k\in\mathbb{N}}$ converges to a point x^* such that $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*) \succ 0$. Then $\{x^k\}_{k\in\mathbb{N}}$ converges superlinearly if*

$$\lim_{k\to\infty} \frac{\|(B_k - \nabla^2 f(x^*))d^k\|}{\|d^k\|} = 0. \tag{7.21}$$

Proof. Let us start with difference between the Newton direction d_N^k and quasi-Newton direction d^k. Since $f \in C^2(\mathbb{R}^n)$, $\nabla^2 f(x^*) \succ 0$ and $\lim_{k\to\infty} x^k = x^*$ we know that $\nabla^2 f(x^k) \succ 0$ for all k large enough. Therefore,

$$\begin{aligned} d^k - d_N^k &= (\nabla^2 f(x^k))^{-1}(\nabla^2 f(x^k)d^k + \nabla f(x^k)) \\ &= (\nabla^2 f(x^k))^{-1}(\nabla^2 f(x^k)d^k - B_k d^k) \\ &= (\nabla^2 f(x^k))^{-1}((\nabla^2 f(x^k) - \nabla^2 f(x^*))d^k + (\nabla^2 f(x^*) - B_k)d^k) \end{aligned}$$

Now, since $\|(\nabla^2 f(x^k))^{-1}\| \leq C_1$ for some $C_1 > 0$ and all k large enough, there holds

$$\frac{\|d^k - d_N^k\|}{\|d^k\|} \leq C_1 \left(\frac{\|\nabla^2 f(x^k) - \nabla^2 f(x^*)\|\|d^k\|}{\|d^k\|} + \frac{\|(\nabla^2 f(x^*) - B_k)d^k\|}{\|d^k\|} \right). \tag{7.22}$$

Taking the limit and using the continuity of the Hessian and (7.21) we obtain that

$$\|d^k - d_N^k\| = o(\|d^k\|). \tag{7.23}$$

Notice that under the stated conditions, Theorem 7.3 implies that the Newton method converges quadratically. Therefore, there exists a constant C_2 such that for k large enough we obtain the following

$$\begin{aligned}
\|x^{k+1} - x^*\| &= \|x^k + d^k - x^*\| \\
&= \|x^k + d^k - x^* + d_N^k - d_N^k\| \\
&\leq \|x^k + d_N^k - x^*\| + \|d^k - d_N^k\| \\
&\leq C_2\|x^k - x^*\|^2 + \|d^k - d_N^k\|,
\end{aligned}$$

so

$$\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} \leq C_2\|x^k - x^*\| + \frac{\|d^k - d_N^k\|}{\|x^k - x^*\|}. \tag{7.24}$$

Notice that

$$\|d^k\| = \|x^{k+1} - x^k\| \leq \|x^{k+1} - x^*\| + \|x^* - x^k\|.$$

Since x^k tends to x^* there must exist $C_3 > 0$ such that $\|x^{k+1} - x^k\| \leq C_3\|x^k - x^*\|$ and thus

$$\|d^k\| \leq (1 + C_3)\|x^k - x^*\|.$$

Furthermore, we obtain

$$\frac{\|d^k - d_N^k\|}{\|x^k - x^*\|} \leq (1 + C_3)\frac{\|d^k - d_N^k\|}{\|d^k\|} \tag{7.25}$$

and thus (7.23) implies that

$$\lim_{k\to\infty} \frac{\|d^k - d_N^k\|}{\|x^k - x^*\|} = 0.$$

Finally, taking the limit in (7.24) we obtain the result. ∎

7.3 Inexact Newton methods

In this section we consider the search direction which is an approximation of the Newton direction in a sense that it solves the Newton equation (7.1) inexactly. The error is usually gradient-related so the inexact Newton direction d^k satisfies

$$\|\nabla^2 f(x^k)d^k + \nabla f(x^k)\| \leq \eta_k \|\nabla f(x^k)\|, \tag{7.26}$$

where $\eta_k \geq 0$. Notice that the choice $\eta_k = 0$ provides the Newton direction. Moreover, as we will see in the sequel, the choice of η_k controls the rate of convergence. In order to prove the main result, we need the following result.

Theorem 7.6 *Suppose that $f \in C^2(\mathbb{R}^n)$ and that the Hessian is Lipschitz continuous on the neighborhood of x^* such that $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*) \succ 0$. Then there exists $\epsilon > 0$ such that for all $x \in B(x^*, \epsilon)$ the Hessian is positive definite and the following holds:*

a)

$$\|\nabla^2 f(x)\| \leq 2\|\nabla^2 f(x^*)\|.$$

b)

$$\|(\nabla^2 f(x))^{-1}\| \leq 2\|(\nabla^2 f(x^*))^{-1}\|.$$

c)

$$\frac{\|x - x^*\|}{2\|(\nabla^2 f(x^*))^{-1}\|} \leq \|\nabla f(x)\| \leq 2\|\nabla^2 f(x^*)\|\|x - x^*\|.$$

Theorem 7.7 *Suppose that $f \in C^2(\mathbb{R}^n)$ and that the Hessian is Lipschitz continuous on the neighborhood of x^* such that $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*) > 0$. Then there exist $C, \delta > 0$ such that for all $x^k \in B(x^*, \delta)$ the inexact Newton iteration satisfies*

$$\|x^{k+1} - x^*\| \leq C(\|x^k - x^*\| + \eta_k)\|x^k - x^*\|.$$

Proof. Notice that under the stated conditions Theorem 7.3 implies the existence of $C_1 > 0$ and $\varepsilon > 0$ such that if $x^k \in B(x^*, \varepsilon)$ there holds

$$\|x^k + d_N^k - x^*\| = \|x^k - (\nabla^2 f(x^k))^{-1}\nabla f(x^k) - x^*\| \leq C_1\|x^k - x^*\|^2. \tag{7.27}$$

Now, let us define $\delta = \min\{\epsilon, \varepsilon\}$ where ϵ is such that the results of Theorem 7.6 hold on $B(x^*, \epsilon)$. Suppose that $x^k \in B(x^*, \delta)$. Now, let us denote by r_k the residual of the Newton's equation, i.e.,

$$r_k := \nabla^2 f(x^k) d^k + \nabla f(x^k),$$

so there holds $\|r_k\| \leq \eta_k \|\nabla f(x^k)\|$. Obviously,

$$d^k = (\nabla^2 f(x^k))^{-1} r_k - (\nabla^2 f(x^k))^{-1}\nabla f(x^k)$$

and we obtain

$$\begin{aligned}
\|x^{k+1} - x^*\| &= \|x^k + (\nabla^2 f(x^k))^{-1} r_k - (\nabla^2 f(x^k))^{-1}\nabla f(x^k) - x^*\| \\
&\leq C_1\|x^k - x^*\|^2 + \|(\nabla^2 f(x^k))^{-1}\|\|r_k\| \\
&\leq C_1\|x^k - x^*\|^2 + 2\|(\nabla^2 f(x^*))^{-1}\|\eta_k\|\nabla f(x^k)\| \\
&\leq C_1\|x^k - x^*\|^2 + 2\|(\nabla^2 f(x^*))^{-1}\|\eta_k 2\|\nabla^2 f(x^*)\|\|x^k - x^*\| \\
&\leq C(\|x^k - x^*\| + \eta_k)\|x^k - x^*\|,
\end{aligned}$$

where $C = \max\{C_1, 4\|(\nabla^2 f(x^*))^{-1}\|\|\nabla^2 f(x^*)\|\}$. ∎

Finally, we state the convergence rate result.

Theorem 7.8 *Suppose that $f \in C^2(\mathbb{R}^n)$ and that the Hessian is Lipschitz continuous on the neighborhood of x^* such that $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*) \succ 0$. Then there exist $\delta > 0$ and $\bar{\eta} > 0$ such that the full Inexact Newton method converges linearly to x^* if $x^0 \in B(x^*, \delta)$ and $\{\eta_k\}_{k\in\mathbb{N}} \subseteq [0, \bar{\eta}]$. Moreover, if $\lim_{k\to\infty} \eta_k = 0$, the convergence rate is superlinear and if $\eta_k \leq C\|\nabla f(x^k)\|^p$ for some $C > 0$ and $p \in (0, 1]$, then the convergence rate is of order $1 + p$.*

7.4 Exercises

1. Consider a function $f(x,y) = 0.5(x^2 - y^2)$. Show that the Newton direction is an ascent direction from point $z = (0,1)^T$. Is the negative Newton direction a descent direction from that point?

2. Consider a function $f(x,y) = x^4 + xy + (1+y)^2$. Show that neither Newton nor negative Newton direction are descent in a sense that $\nabla^T f(z)d < 0$ if we consider the point $z = (0,0)^T$.

3. Prove Theorem 7.2.

4. Let $f(x) = 0.5(x_1^2 - x_2)^2 + 0.5(1-x_1)^2$. Solve $\min_{x\in\mathbb{R}^2} f(x)$ analytically. Perform one Newton step from $x^0 = (-1, 1.2)^T$. Calculate $f(x^0)$ and $f(x^1)$ to find out if this is a good step.

5. Consider Newton method applied to minimize the function $f(x) = \sin(x)$, $f : \mathbb{R} \to \mathbb{R}$ with $x^0 \in [-\pi, \pi]$. A local minimizer is $\tilde{x} = -\pi/2$. Suppose that $\varepsilon > 0$ is sufficiently small. If $x^0 = -\varepsilon$, show that $x_1 \approx -1/\varepsilon$. What happens if $x^0 = \varepsilon$ but $f''(x^0)$ is replaced with a small positive number?

6. Prove Theorem 7.6.

7. Newton method can converge to a local maximizer. To verify this, apply Newton method to minimize function $f : \mathbb{R} \to \mathbb{R}$,
$$f(x) = -x^4/4 + x^3/3 + x^2,$$
with $x^0 = 1$. What happens to Newton method if we minimize $f(x) = x^3/3 + x$?

8. Let $f(x) = \sum_{i=1}^{n}(a_i x_i^2 + b_i x_i)$ where $a_i, b_i \in \mathbb{R}$ for $i = 1, ..., n$. Find conditions under which the Newton direction is well defined and descent.

9. Show that γ_k given by (7.19) is the solution of the problem

$$\min_{\gamma\in\mathbb{R}} \|\gamma y^{k-1} - s^{k-1}\|^2.$$

10. Find the spectral gradient direction if the other form of the secant equation (7.11) is considered.

11. Prove Theorem 7.8.

12. Assume that A, C, U, V are matrices of the corresponding dimensions and that A is nonsingular matrix. The Sherman- Morrison-Woodbury formula states that

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1},$$

assuming that $C^{-1} + VA^{-1}U$ is nonsingular. Using this formula prove (7.15) and derive the corresponding expression for the BFGS update.

Chapter 8

Least squares problems

One of the important special classes of unconstrained optimization problems are the least squares problems which are written in the form

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}\sum_{j=1}^{m}(r_j(x))^2, \tag{8.1}$$

where $r_j : \mathbb{R}^n \to \mathbb{R}, j = 1, ..., m$ are usually assumed to be smooth functions. Functions r_j are often referred to as residuals. One typical example of the least squares problem comes from data fitting and in that case the residual is actually the error produced by employing the model function. For example, if linear regression is considered then the model function $\Phi(x; t) = x^T t$ approximates the real values of the outcome y according to the measurements t and x represents the unknown vector of the relevant coefficients. More precisely, assuming that there are m measurements, one can define the error as $r_j(x) = \Phi(x; t_j) - y_j$ and, minimization of the cumulative squared error yields (8.1).

Let us denote by r the following function from $\mathbb{R}^n$ to $\mathbb{R}^m$

$$r(x) = (r_1(x), ..., r_m(x))^T.$$

Furthermore, denote the Jacobian of this function by J, i.e.,

$$J(x) := \nabla r(x).$$

Now, the problem (8.1) can be stated equivalently as

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2}\|r(x)\|^2$$

and the derivatives of the objective function are

$$\nabla f(x) = J^T(x) r(x) = \sum_{j=1}^{m} r_j(x) \nabla r_j(x), \tag{8.2}$$

$$\nabla^2 f(x) = \sum_{j=1}^{m} \nabla r_j(x) \nabla^T r_j(x) + \sum_{j=1}^{m} r_j(x) \nabla^2 r_j(x) := G(x) + H(x). \tag{8.3}$$

Notice that $G(x) = J^T(x)J(x)$.

First, we consider the linear least squares problem where each residual is a linear function. One of the examples of the linear least squares problem is the linear regression problem mentioned above. In that case, r can be represented by

$$r(x) = Ax + b, \tag{8.4}$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$. Moreover, notice that $J(x) = A$ and $\nabla f(x) = A^T(Ax + b)$. Also, we have that $\nabla^2 r_j(x) = 0$, $j = 1, ..., m$ which implies $H(x) = 0$ and

$$\nabla^2 f(x) = A^T A.$$

This furthermore implies that the problem is convex and thus a stationary point of function f cannot be a maximizer. Therefore a candidate solution x^* is a point that satisfies $\nabla f(x^*) = 0$. In other words, linear least squares problem reduces to the problem of solving system of linear equations

$$A^T A x^* = -A^T b. \tag{8.5}$$

The system (8.5) is often called the system of normal equations and it is usually solved numerically (by means of factorization for instance). Of course, convex optimization tools are also applicable.

Now, we consider the general case of (8.1), i.e., nonlinear least squares problem and the Gauss-Newton method. The Gauss-Newton direction d^k is defined by

$$J^T(x^k)J(x^k)d^k = -J^T(x^k)r(x^k). \tag{8.6}$$

Notice that this is in fact a Quasi-Newton method since the Hessian is approximated only by the first sum in (8.3), i.e.,

$$\nabla^2 f(x^k) \approx G(x^k).$$

This way the calculation of the second order derivatives $\nabla^2 r_j(x^k)$ is avoided which can bring significant savings in the optimization process. Moreover, $G(x^k)$ often dominates $H(x^k)$ and thus a good approximation of the Hessian is provided. This furthermore implies that the Gauss-Newton direction is close to the Newton direction and thus fast convergence is possible.

Another advantage of the Gauss-Newton direction is its descent property. Indeed,

$$\begin{aligned}
(d^k)^T \nabla f(x^k) &= (d^k)^T J^T(x^k) r(x^k) \\
&= (d^k)^T(-J^T(x^k)J(x^k)d^k) \\
&= -\|J(x^k)d^k\|^2 \leq 0.
\end{aligned}$$

Since, according to (8.6), $J(x^k)d^k = 0$ implies $J^T(x^k)r(x^k) = 0$, i.e., $\nabla f(x^k) = 0$, we conclude that the Gauss-Newton direction is a descent direction provided that x^k is not a stationary point. Finally, notice that d^k is a solution of the linear least squares problem

$$\min_{d\in\mathbb{R}^n} \frac{1}{2}\|J(x^k)d + r(x^k)\|^2, \tag{8.7}$$

so the linear least squares problem can be viewed as a subproblem in the nonlinear least squares setup.

Let us now consider the convergence of the Gauss-Newton method incorporated in the Armijo line search backtracking framework. The convergence result presented after the Algorithm 8.1 a consequence of Theorem 5.3.

Algorithm 8.1

Step 0 Input parameters: $x^0 \in \mathbb{R}^n$, $\beta, \eta \in (0,1)$.

Step 1 Initialization: $k = 0$, $x^k = x^0$.

Step 2 Stopping criterion: If $\nabla f(x^k) = 0$ STOP. Otherwise go to Step 3.

Step 3 Search direction: Compute the Gauss-Newton direction d^k by solving (8.6).

Step 4 Step size: Find the smallest nonnegative integer j such that $\alpha_k = \beta^j$ satisfies the Armijo condition

$$f(x^k + \alpha_k d^k) \le f(x^k) + \eta\alpha_k \nabla^T f(x^k) d^k.$$

Step 5 Update: Set $x^{k+1} = x^k + \alpha_k d^k$, $k = k+1$ and go to Step 2.

Theorem 8.1 *Suppose that $r \in C^1(\mathbb{R}^n)$ and that the level set $\mathcal{L}(x^0) = \{x \in \mathbb{R}^n \mid f(x) \le f(x^0)\}$ is bounded. Assume that $G(x)$ is uniformly positive definite on an open set that contains $\mathcal{L}(x^0)$. Then either the Algorithm 8.1 terminates after a finite number of iterations $\bar{k}$*

at the stationary point $x^{\bar{k}}$ or every accumulation point of the sequence $\{x^k\}_{k\in\mathbb{N}}$ is a stationary point of the function f.

Proof. First, notice that the objective function is nonnegative and thus bounded from below. Also, $f \in C^1(\mathbb{R}^n)$ as the same was assumed for r. Therefore, the line search step is well defined since the Gauss-Newton direction is a descent direction unless the current point is stationary. Moreover, Armijo condition implies that all iterates belong to the level set $\mathcal{L}(x^0)$. Since the level set is assumed to be bounded, the same is true for the sequence $\{x^k\}_{k\in\mathbb{N}}$ and the continuity of the gradient ∇f implies that the sequence $\{\nabla f(x^k)\}_{k\in\mathbb{N}}$ is bounded as well, i.e., there exists a constant $M > 0$ such that $\|\nabla f(x^k)\| \leq M$ for every k. On the other hand, since $G(x)$ is assumed to be uniformly positive definite, there exists $m > 0$ such that $\lambda_{min}(G(x^k)) \geq m$ and

$$\|d^k\| = \| - (G(x^k))^{-1}\nabla f(x^k)\| \leq \|(G(x^k))^{-1}\|\|\nabla f(x^k)\| \leq \frac{1}{m}M.$$

Therefore, the sequence of search directions is also bounded. Moreover, the definition of $G(x^k)$ also implies the existence of a constant $S > 0$ such that $\|G(x^k)\| \leq S$, i.e., $\lambda_{min}((G(x^k))^{-1}) \geq 1/S$ and thus

$$\nabla^T f(x^k)d^k = -\nabla^T f(x^k)(G(x^k))^{-1}\nabla f(x^k) \leq -\frac{1}{S}\|\nabla f(x^k)\|^2.$$

Having all this in mind, we conclude that the conditions of the Theorem 5.3 are fulfilled and that (5.6) holds, so the rest of the proof is the same as for the Theorem 5.3. ∎

One of the key assumptions that makes the Gauss-Newton method convergent is that the matrix of the system (8.6) (i.e., the Quasi-Newton matrix) is uniformly positive definite. In order to ensure that the Quasi-Newton matrix is positive definite, one can use the so called regularization. More precisely, having in mind that the matrix

$G(x) = J^T(x^k)J(x^k) \succeq 0$, adding a constant positive definite matrix will provide that the eigenvalues of the new matrix are bounded away from zero. Let us fix $\rho > 0$ as a regularization parameter. Then, using the feature of symmetric quadratic matrices, we obtain that the smallest eigenvalue λ_{min} of the regularized Hessian satisfies

$$\lambda_{min}(J^T(x^k)J(x^k) + \rho I) \geq \lambda_{min}(J^T(x^k)J(x^k)) + \lambda_{min}(\rho I) \geq \rho.$$

The method described above is called the Levenberg-Marquardt method and the relevant search direction is obtained as the solution of the following system

$$(J^T(x^k)J(x^k) + \rho I)d^k = -J^T(x^k)r(x^k). \tag{8.8}$$

Notice that the Levenberg-Marquardt direction is uniquely determined while the Gauss-Newton direction is not since $G(x^k)$ may be singular. The convergence analysis of the Levenberg-Marquardt method can be conducted as in the case of Gauss-Newton method.

Now, let us comment on the rate of convergence of the Gauss-Newton method. Recall that the Hessian (8.3) is approximated by $G(x)$. Therefore, we expect that the error of that approximation ($H(x)$) will play a key role in determining the convergence rate. Of course, if $H(x) = 0$, the convergence rate is quadratic since the Newton direction is recovered. On the other hand, if $H(x)$ is close to zero, then the method can exhibit near-quadratic behavior. Small $H(x)$ occurs in practice since the residuals are often very small or nearly linear in the vicinity of the solution. We conclude this section by proving the linear convergence, while the conditions for superlinear rate can be analyzed throughout the Quasi-Newton approach.

Theorem 8.2 *Suppose that $r \in C^2(\mathbb{R}^n)$ and consider the Gauss-Newton algorithm, i.e., $x^{k+1} = x^k + d^k$ where d^k is defined by (8.6). Let us suppose that H is Lipschitz continuous. Assume further that*

$\{x^k\}_{k\in\mathbb{N}}$ converges to a point x^* such that $\nabla f(x^*) = 0$ and $G(x^*) \succ 0$. Then $\{x^k\}_{k\in\mathbb{N}}$ converges linearly if

$$\|(G(x^*))^{-1}H(x^*)\| < 1. \tag{8.9}$$

Proof. Since $r \in C^2(\mathbb{R}^n)$, $G(x^*) \succ 0$ and $\{x^k\}_{k\in\mathbb{N}}$ converges to x^*, then $(G(x^k))^{-1}$ exists and there exists $m > 0$ such that$\|(G(x^k))^{-1}\| \leq 1/m := M$ for all k large enough. Moreover, G is also Lipschitz continuous and without loss of generality we can denote by L the Lipschitz constant for G and H. From now on, assume that k is large enough. Then, using the Mean value theorem we obtain the following.

$$\begin{aligned}
x^{k+1} - x^* &= x^k - (G(x^k))^{-1}\nabla f(x^k) - x^* \\
&= (G(x^k))^{-1}\left(G(x^k)(x^k - x^*) - \nabla f(x^k) + \nabla f(x^*)\right) \\
&= (G(x^k))^{-1}(\int_0^1 G(x^k)(x^k - x^*)dt \\
&- \int_0^1 \nabla^2 f(x^k + t(x^k - x^*))(x^k - x^*)dt) \\
&= (G(x^k))^{-1}\left(\int_0^1 (G(x^k) - G(x^k + t(x^k - x^*)))(x^k - x^*)dt\right) \\
&+ \int_0^1 G(x^k))^{-1}H(x^k + t(x^k - x^*))(x^k - x^*)dt.
\end{aligned}$$

Furthermore, using the Lipschitz continuity we obtain

$$\begin{aligned}
\|x^{k+1} - x^*\| &\leq \|(G(x^k))^{-1}\| \int_0^1 \|G(x^k) - G(x^k + t(x^k - x^*))\|\|x^k - x^*\|dt \\
&+ \int_0^1 \|(G(x^k))^{-1}H(x^k + t(x^k - x^*))\|\|(x^k - x^*)\|dt \\
&\leq ML\|x^k - x^*\|^2 \\
&+ \int_0^1 \|(G(x^k))^{-1}H(x^k + t(x^k - x^*))\|\|(x^k - x^*)\|dt.
\end{aligned}$$

Now, notice that we can bound $\|(G(x^k))^{-1}H(x^k + t(x^k - x^*))\|$ as follows

$$\begin{array}{rcl} \|(G(x^k))^{-1}H(x^k+t(x^k-x^*))\| & = & \|(G(x^k))^{-1}(H(x^k+t(x^k-x^*)) \pm H(x^k))\| \\ & \leq & \|(G(x^k))^{-1}H(x^k)\| \\ & + & \|(G(x^k))^{-1}(H(x^k+t(x^k-x^*)) - H(x^k))\| \\ . & \leq & \|(G(x^k))^{-1}H(x^k)\| + ML\|x^k - x^*\|. \end{array}$$

Putting all together we obtain

$$\|x^{k+1} - x^*\| \leq 2ML\|x^k - x^*\|^2 + \|(G(x^k))^{-1}H(x^k)\|\|x^k - x^*\|, \quad (8.10)$$

or equivalently,

$$\|x^{k+1} - x^*\| \leq \left(\|(G(x^k))^{-1}H(x^k)\| + 2ML\|x^k - x^*\|\right)\|x^k - x^*\|. \quad (8.11)$$

Since we assumed (8.9), the continuity implies that $\|(G(x^k))^{-1}H(x^k)\| < 1$ for all k large enough. Finally, we conclude that there exists $\gamma \in (0,1)$ such that for all k large enough

$$\|(G(x^k))^{-1}H(x^k)\| + 2ML\|x^k - x^*\| \leq \gamma$$

and the result follows immediately. ■

8.1 Exercises

1. Consider the linear least squares problem

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}\|Ax + b\|^2,$$

where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$. Under which conditions on A it has an unique solution?

2. Consider a linear regression model function $\Phi(x; a^i) = x^T a^i + c_i$, where $a^i \in \mathbb{R}^n, c_i \in \mathbb{R}$, $i = 1, ..., m$ are attribute measurements, $x \in \mathbb{R}^n$ is a decision variable and $y_1, ..., y_m$ are the measured values of the target function. State the least squares problem. State the optimality conditions.

3. Prove that the Gauss-Newton method is equivalent to Newton method if the linear least squares problem is considered.

4. Consider nonlinear regression model function $\Phi(x; t) = x_1 e^{x_2 t}$ where t represents time. If we know that the real values of the target variable are $y_1, ..., y_m$, state the least squares problem. Find the Gauss-Newton step if $t_1 = 1, t_2 = 2, y_1 = 3, y_2 = 4$ and $x^k = (1, 0)^T$.

5. Analyze conditions for superlinear convergence of the Gauss-Newton and the Levenberg-Marquardt method.

Chapter 9

Constrained optimization

From now on, we consider constrained optimization problems in the standard form which was stated in Chapter 1

$$\min_{x \in S} f(x), \ S = \{x \in \mathbb{R}^n \mid h(x) = 0, \ g(x) \leq 0\}. \tag{9.1}$$

Recall that the objective function $f : D \to \mathbb{R}$, so D actually represents the set of implicit constraints while S states the explicit constraints. We say that $h : \mathbb{R}^n \to \mathbb{R}^m$ represents the equality constraints and $g : \mathbb{R}^n \to \mathbb{R}^p$ determines the inequality constraints.[1]

Now, denote by f^* the optimal value of the problem (9.1), i.e.,

$$f^* = \inf_{x \in S} f(x). \tag{9.2}$$

We say that the problem is infeasible if S is empty and in that case we define

$$f^* = \infty.$$

For example, take $S = \{x \in \mathbb{R}^n \mid x_1 - x_2 = 0, \ x_1 - x_2 = 2\}$ which is an obviously empty. On the other hand, if the objective function is

[1] Other relevant definitions are stated in Chapter 1.

unbounded from below on the feasible set S, we have

$$f^* = -\infty.$$

An example would be the problem $\min_{x \le 0} x$.

Constrained optimization deals with two important concepts: optimality and feasibility. We say that a point x is feasible if it satisfies both implicit and explicit constraints, i.e., if

$$x \in D \cap S.$$

A feasible point x^* is optimal if

$$f(x^*) = f^*.$$

Recall that a point $\tilde{x}$ is locally optimal if there exists $\varepsilon > 0$ such that $\tilde{x}$ is a solution of the problem

$$\min_{x \in S,\ \|x - \tilde{x}\| \le \varepsilon} f(x). \tag{9.3}$$

Let us consider some simple examples without explicit constraints.

Example 1 $f(x) = x^{-2}$. In this case, $D = \mathbb{R} \backslash \{0\}$ and $f^* = 0$, but there is no optimal point.

Example 2 $f(x) = \ln(x)$. In this case, $D = \mathbb{R}_+ \backslash \{0\}$ and $f^* = -\infty$.

Example 3 $f(x) = x \ln(x)$. In this case, $D = \mathbb{R}_+ \backslash \{0\}$, $f^* = -e^{-1}$ and the optimal point is $x^* = e^{-1}$.

Example 4 $f(x) = x^3 - 3x$. In this case, there are no implicit constraints, the optimal value is $f^* = -\infty$, but there is one local minimum at $\tilde{x} = 1$.

Now, let us consider the problem

$$\min_{x\in\mathbb{R}^n} -\sum_{i=1}^{k} \ln(b_i - x^T a_i). \tag{9.4}$$

This is considered as unconstrained problem since there are no explicit constraints, but implicit constraints are present. The equivalent problem can be stated as

$$\min_{x\in S} -\sum_{i=1}^{k} \ln(b_i - x^T a_i),\ S = \{x \in \mathbb{R}^n \mid x^T a_i < b_i, i = 1, ..., k.\}. \tag{9.5}$$

Moreover, constraints functions h and g can produce implicit constraints as well. Denote by $D(f)$ the domain of considered function f.Then, the implicit constraints of the optimization problem (9.1) are in fact

$$D = D(f) \bigcap (\cap_{i=1}^{m} D(h_i)) \bigcap (\cap_{j=1}^{p} D(g_j)). \tag{9.6}$$

The implicit constraints related to h and g are often written as explicit. For example, the problem

$$\min_{x\in S} f(x),\ S = \{x \in \mathbb{R} \mid \ \ln(x) \leq 2\}.$$

can be represented as

$$\min_{x\in S} f(x),\ S = \{x \in \mathbb{R} \mid \ \ln(x) \leq 2,\ x > 0\}.$$

Therefore, without loss of generality, we can consider the case where $D = D(f)$.

Finding a feasible point itself can be a challenging problem. This particular problem can be stated as

$$\min_{x\in S} c, \tag{9.7}$$

where c is an arbitrary constant. Thus, $f^* = c$ and every feasible point is optimal if S is nonempty.

9.1 Convex problems

Within this section, we focus on an important class of constrained optimization problems - convex problems. The problem (9.1) is convex if the objective function f and the inequality constraints functions $g_1, ..., g_m$ are convex, while the equality constraints functions $h_1, ..., h_p$ are affine, i.e., $h_i(x) = x^T a_i - b_i$, for $i = 1, ..., p$.

Convex problems have some very nice properties. One of them is that the feasible set of a convex problem is convex. Lot of methods developed for this kind of problems use this fact and make the optimization processes more efficient. Therefore, it is important to notice if the problem is convex or if it can be transformed into the convex one. For example, consider

$$\min_{x \in S} \|x\|^2, \ S = \{x \in \mathbb{R}^2 \mid x_1/(1 + x_2^2) \leq 0, \ (x_1 + x_2)^2 = 0\}. \quad (9.8)$$

The objective function is convex, but the equality constraint function is not affine and thus the problem is not convex. However, notice that the constraints are equivalent to $x_1 \leq 0$, $x_1 + x_2 = 0$ and the problem can be reformulated as the convex one

$$\min_{x \in S} \|x\|^2, \ S = \{x \in \mathbb{R}^2 \mid x_1 \leq 0, \ x_1 + x_2 = 0\}. \quad (9.9)$$

Similar to the unconstrained case, one can show that any local solution is also a global solution. We prove this statement as follows.

Theorem 9.1 *Every local solution of a convex constrained problem is a global solution of the same problem.*

Proof. Let us assume that x^* is a local, but not a global solution of the convex problem $\min_{x \in S} f(x)$. In other words, there exists $y^* \in S$ such that $f(y^*) < f(x^*)$. Moreover, since x^* is a local optimum, there

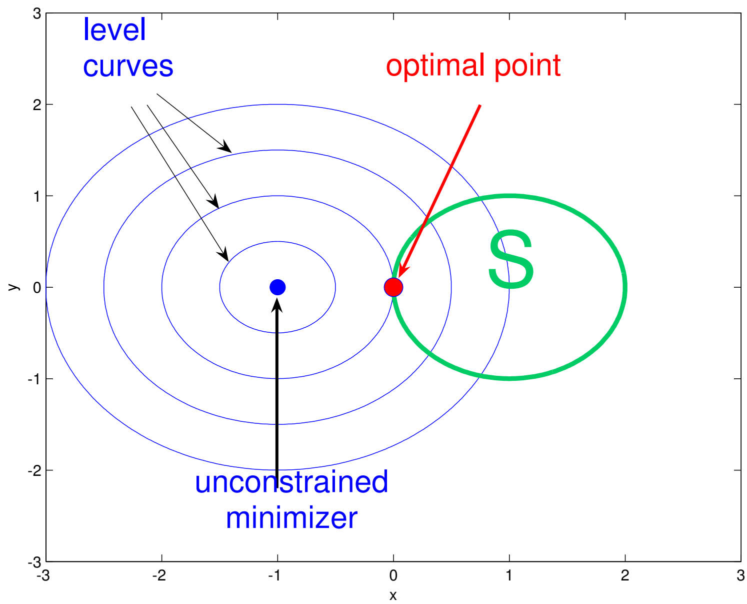

Figure 9.1: Convex constrained problem.

exists $\varepsilon > 0$ such that x^* solves the problem (9.3). Notice that in that case there holds $\|y^* - x^*\| > \varepsilon$.

Now, let us consider a point $z = \lambda y^* + (1-\lambda)x^*$, where $\lambda = \varepsilon/(2\|y^* - x^*\|)$. Notice that $\lambda \in (0, 1/2)$. Moreover, since the feasible set S is convex, we conclude that $z \in S$. On the other hand,

$$\|z - x^*\| = \|\lambda(y^* - x^*)\| = \frac{\varepsilon}{2\|y^* - x^*\|}\|y^* - x^*\| = \frac{\varepsilon}{2} < \varepsilon.$$

Therefore, z is a feasible point of the problem (9.3) and there holds $f(x^*) \leq f(z)$. However, using the convexity of the objective function we obtain the following contradiction

$$f(z) \leq \lambda f(y^*) + (1-\lambda)f(x^*) < \lambda f(x^*) + (1-\lambda)f(x^*) = f(x^*),$$

which completes the proof. ■

Next, we state the optimality criterion provided that the objective function is continuously differentiable.

Theorem 9.2 *Suppose that $f \in C^1(\mathbb{R}^n)$ and that the problem is convex. Then, x^* is optimal if and only if $x^* \in S$ and for every $y \in S$ there holds*

$$\nabla^T f(x^*)(y - x^*) \geq 0. \tag{9.10}$$

Proof. Suppose that $x^* \in S$ satisfies (9.10). Furthermore, since f is convex and differentiable, we obtain

$$f(y) \geq f(x^*) + \nabla^T f(x^*)(y - x^*) \geq f(x^*),$$

for every $y \in S$. Therefore, x^* is an optimal point.

Now, assume that x^* is an optimal point and that (9.10) does not hold, i.e., there exists $z \in S$ such that

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

Consider points of the form $u(\lambda) = \lambda z + (1-\lambda)x^*$, where $\lambda \in [0,1]$. Since the feasible set is convex, these points are feasible. Consider the function $g(\lambda) = f(u(\lambda))$. Notice that $g'(\lambda) = \nabla^T f(u(\lambda))(z - x^*)$ and

$$g'(0) = \nabla^T f(x^*)(z - x^*) < 0.$$

So, we conclude that $g'(\lambda) < 0$ for $\lambda > 0$ small enough, i.e., g is a decreasing function of λ for $\lambda > 0$ small enough. Therefore, there exists $\lambda > 0$ such that $g(\lambda) < g(0)$, or equivalently,

$$f(u(\lambda)) < f(u(0)) = f(x^*).$$

Since such $u(\lambda)$ is feasible, x^* cannot be an optimal point. This is a contradiction, so we conclude that the solution must satisfy (9.10) for every $y \in S$. ∎

9.2 Exercises

1. Prove that the feasible set of a convex problem is convex.

2. Consider the convex problem

$$\min_{Ax=b} f(x),$$

where $A \in \mathbb{R}^{p\times n}, b \in \mathbb{R}^p$. Prove that a point x^* is a solution of that problem if and only if $Ax^* = b$ and there exists $\mu \in \mathbb{R}^p$ such that $\nabla f(x^*) = A^T \mu$.

3. Consider the convex problem

$$\min_{x\geq 0} f(x).$$

Prove that a point x^* is a solution of that problem if and only if $x^* \geq 0$, $\nabla f(x^*) \geq 0$ and $x_i^* [\nabla f(x^*)]_i = 0$ for $i = 1, ..., n$.

4. Solve the problem
$$\min_{x \geq 0} f(x)$$
with the following objective functions:

   (a) $f(x) = x_1 + x_2$.
   (b) $f(x) = x_1 + x_2^2$.
   (c) $f(x) = (x_1 + 1)^2 + x_2^2$.
   (d) $f(x) = (x_1 + 1)^2 + (x_2 - 1)^2$.
   (e) $f(x) = (x_1 - 1)^2 + (x_2 - 1)^2$.
   (f) $f(x) = \|x\|^2$.

   Sketch the level curves of each objective function.

5. Solve the following convex problems:

   (a) $\min_{x \in S} f(x),\ S = \{x \in \mathbb{R}^2 \mid 2x_1 + x_2 = 1,\ x_1 - x_2 = 0\}$.
   (b) $\min_{x \in S}(x_1 - 1)^2 + x_2^2,\ S = \{x \in \mathbb{R}^2 \mid x_1 = x_2\}$.
   (c) $\min_{x \in S} \|x\|^2,\ S = \{x \in \mathbb{R}^3 \mid 2x_1 + x_2 = 1,\ x_1 - x_2 = 0\}$.
   (d) $\min_{x \in S} 0.5\|x - 1\|^2,\ S = \{x \in \mathbb{R}^3 \mid x_1 = x_2,\ x_2 + x_3 = 2\}$.

6. Consider the problem $\min_{x \in S} f(x)$ where
$$S = \{x \in \mathbb{R}^2 \mid 2x_1 + x_2 \geq 1,\ x_1 + 3x_2 \geq 1,\ x_1 \geq 0,\ x_2 \geq 0\}.$$
Sketch the feasible set. For each of the following objective functions, find the set of optimal solutions and determine the optimal value.

   (a) $f(x) = x_1 + x_2$.
   (b) $f(x) = -x_1 - x_2$.

(c) $f(x) = x_1$.

(d) $f(x) = \max\{x_1, x_2\}$.

(e) $f(x) = x_1^2 + 9x_2^2$.

Chapter 10

Optimality conditions - constrained optimization

Within this chapter we consider characterizations of optimal points for constrained optimization problems in the standard form (9.1). Let us assume that the domain D is nonempty and all the relevant functions are continuously differentiable. We form a function which combines optimality and feasibility throughout the weighted sum of the objective function and the constraints functions. This function is called the Lagrangian and it is defined as follows:

$$L(x, \lambda, \mu) := f(x) + \lambda^T g(x) + \mu^T h(x) = f(x) + \sum_{i=1}^{p} \lambda_i g_i(x) + \sum_{j=1}^{m} \mu_j h_j(x), \tag{10.1}$$

where $\lambda = (\lambda_1, ..., \lambda_p)^T \in \mathbb{R}^p$ are the Lagrange multipliers associated to inequality constraints and $\mu = (\mu_1, ..., \mu_m)^T \in \mathbb{R}^m$ are the Lagrange multipliers associated to equality constraints. Vectors λ and μ are also called the dual variables.

10.1 Duality

In order to derive optimality conditions, we form the Lagrange dual function

$$l(\lambda, \mu) := \inf_{x \in D} L(x, \lambda, \mu). \tag{10.2}$$

Notice that the Lagrangian L is an affine function with respect to λ and μ. It can be shown that the infimum of a family of affine functions is always concave, so we conclude that l is a concave function regardless of the convexity of the problem (9.1). We use this fact in the sequel to obtain the lower bound on the optimal value of problem (9.1).

Suppose that $x \in D$ is a feasible point and that $\lambda \geq 0$. Then we have $g(x) \leq 0$, $h(x) = 0$ and

$$\sum_{i=1}^{p} \lambda_i g_i(x) + \sum_{j=1}^{m} \mu_j h_j(x) \leq 0.$$

The above inequality and (10.1) yield $L(x, \lambda, \mu) \leq f(x)$ and thus

$$l(\lambda, \mu) = \inf_{y \in D} L(y, \lambda, \mu) \leq L(x, \lambda, \mu) \leq f(x). \tag{10.3}$$

Since the inequality holds for an arbitrary feasible x, we conclude that

$$l(\lambda, \mu) \leq \min_{x \in S} f(x) = f^*. \tag{10.4}$$

Therefore, for any $\lambda \geq 0$ and $\mu \in \mathbb{R}^m$, the Lagrange dual function $l(\lambda, \mu)$ represents a lower bound of the optimal value of problem (9.1).

The inequality (10.4) trivially holds if $l(\lambda, \mu) = -\infty$, but does not provide any useful information. Therefore, we are interested in points that belong to the domain of the Lagrange dual function. So, the points of our interest are $(\lambda, \mu) \in D(l)$ such that $\lambda \geq 0$. These pairs of dual variables are called dual feasible.

In order to obtain the best possible lower bound, we pose the Lagrange dual problem

$$\max_{\lambda \geq 0} l(\lambda, \mu). \tag{10.5}$$

Notice that this problem is convex. Therefore, there exist an unique solution (λ^*, μ^*) which is called the dual optimal or we refer to (λ^*, μ^*) as the optimal Lagrange multipliers.

10.2 KKT conditions and the strong duality

In this section we derive the set of optimality conditions called the KKT (Karush-Kuhn-Tucker) conditions. We saw that $l(\lambda, \mu) \leq f^*$, i.e.,

$$-f^* \leq -l(\lambda, \mu) \tag{10.6}$$

if (λ, μ) is dual feasible. Furthermore, assuming that x is primal feasible, i.e., feasible for problem (9.1), we obtain

$$0 \leq f(x) - f^* \leq f(x) - l(\lambda, \mu) := \epsilon(x, \lambda, \mu) \tag{10.7}$$

and we say that x is ϵ-suboptimal with $\epsilon = \epsilon(x, \lambda, \mu)$. We call this ϵ the duality gap between primal and dual variables.

If there is no duality gap, i.e., if $\epsilon(x^*, \lambda^*, \mu^*) = 0$, then x^* is primal optimal and (λ^*, μ^*) is dual optimal. Indeed, if $\epsilon(x^*, \lambda^*, \mu^*) = 0$ then $f(x^*) = f^*$ and, since x^* is assumed to be primal feasible, we conclude that x^* is an optimal point of problem (9.1). On the other hand, since f^* is actually an upper bound for l in the case of feasible dual variables, having $l(\lambda^*, \mu^*) = f(x^*)$ means that (λ^*, μ^*) is a solution of the dual problem (10.5). This fact is often used as a termination criterion for algorithms - they are stopped if the duality gap is smaller than some tolerance $\delta > 0$, i.e., if

$$\epsilon(x^k, \lambda^k, \mu^k) \le \delta. \tag{10.8}$$

Let us define the strong duality needed for further derivation.

Definition 11 *Strong duality holds if the primal and dual optimal values are attained and equal.*

Assume that the strong duality holds. Then $\epsilon(x^*, \lambda^*, \mu^*) = 0$ if and only if x^* is primal optimal and (λ^*, μ^*) is dual optimal. Furthermore,

$$\begin{aligned} f(x^*) &= l(\lambda^*, \mu^*) \\ &= \inf_{x \in D} L(x, \lambda^*, \mu^*) \\ &\le L(x^*, \lambda^*, \mu^*) \\ &= f(x^*) + (\lambda^*)^T g(x^*) + (\mu^*)^T h(x^*) \\ &\le f(x^*). \end{aligned}$$

Thus, we conclude that the previous derivation holds with the equalities. Two important conclusions can be derived from this fact. First,

$$\inf_{x \in D} L(x, \lambda^*, \mu^*) = L(x^*, \lambda^*, \mu^*), \tag{10.9}$$

i.e., x^* is a minimizer of $L(x, \lambda^*, \mu^*)$ and there holds

$$\nabla_x L(x^*, \lambda^*, \mu^*) = 0. \tag{10.10}$$

However, it does not have to be unique. Second conclusion is that

$$f(x^*) + (\lambda^*)^T g(x^*) + (\mu^*)^T h(x^*) = f(x^*),$$

that is,

$$(\lambda^*)^T g(x^*) + (\mu^*)^T h(x^*) = 0. \tag{10.11}$$

Since x^* is feasible, we have that $h(x^*) = 0$ and thus $(\lambda^*)^T g(x^*) = 0$. Furthermore, since $g(x^*) \leq 0$ and $\lambda^* \geq 0$, we conclude that

$$\lambda_i^* g_i(x^*) = 0, \quad i = 1, ..., p. \tag{10.12}$$

The condition (10.12) is known as the complementarity condition. It implies that if the Lagrange multiplier $\lambda_i^* > 0$, then $g_i(x^*) = 0$, i.e., x^* is on the boundary determined by the constraint g_i. In general, if $g_i(x) = 0$, we say that the ith constraint is active at x. For instance, if we have $g_i(x) = x_1^2 + x_2^2 - 1$, then the feasible points are within the unit circle and if this constraint is active at x^* we know that this point is on the boundary of the circle. On the other hand, if $g_i(x^*) < 0$, i.e., if the point x^* is strictly inside of the area determined by g_i, then $\lambda_i^* = 0$. Recall that these conclusions are derived only under the assumption of strong duality. We sum them up into the famous KKT conditions. Notice that these are necessary conditions provided that the strong duality holds.

Definition 12 *KKT conditions are:*

a) $g(x^*) \leq 0$ *(feasibility - inequality constraints).*

b) $h(x^*) = 0$ *(feasibility - equality constraints).*

c) $\lambda^* \geq 0$ *(dual feasibility).*

d) $\lambda_i^* g_i(x^*) = 0, \quad i = 1, ..., p$ *(complementarity).*

e) $\nabla f(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla g_i(x^*) + \sum_{i=j}^{m} \mu_j^* \nabla h_j(x^*) = 0$ *(optimality).*

In general, it is very hard to check if the strong duality holds in advance. Nevertheless, KKT conditions are often used as a guidance for finding candidate solutions. The conditions are often used in construction of many algorithms. The KKT system has $n + p + m$ unknowns and this is the number of equality equations in the KKT system as well. So, the candidate solutions may be obtained by solving the system of equations defined by the KKT conditions b) and d)-e), while the true solutions must also satisfy the remaining conditions a) and c). However, the system of equations is nonlinear in general and therefore not easy to solve.

10.2.1 Convex problems

An important special case is a convex problem (9.1). In that case, KKT conditions are sufficient. Indeed, assume that KKT conditions hold. Then, a)-b) imply that the point x^* is feasible. Moreover, c) implies that the Lagrangian is convex and thus e) implies that x^* is a global minimizer of $L(x, \lambda^*, \mu^*)$, so

$$l(\lambda^*, \mu^*) = \inf_{x \in D} L(x, \lambda^*, \mu^*) = L(x^*, \lambda^*, \mu^*) = f(x^*)$$

where the last equality comes from b) and d). Therefore, the optimality gap vanishes $\epsilon(x^*, \lambda^*, \mu^*) = 0$ which implies that x^* is primal optimal and (λ^*, μ^*) is dual optimal as we already discussed. We formalize this as follows.

Theorem 10.1 *Suppose that x^* and (λ^*, μ^*) are such that the KKT conditions are satisfied and the problem (9.1) is convex. Then x^* is a solution of the problem (9.1).*

We end this subsection by stating the Slater's condition which implies the strong duality of a convex problem.

Definition 13 *Suppose that the problem (9.1) is convex. Slater's condition holds if there exist at least one feasible point $\tilde{x}$ such that $g_i(\tilde{x}) < 0$ for all $i = 1, ..., p$.*

Finally, having all the previous discussions in mind, we state the following result for convex problems.

Theorem 10.2 *Suppose that the Slater's condition holds. Then x^* is a solution of a convex problem (9.1) if and only if the KKT conditions hold.*

10.3 KKT conditions and LICQ

First and second order optimality conditions are often analyzed under the assumption of linear independence constraint qualification (LICQ).

Definition 14 *LICQ holds at point x^* if the gradients of active constraints at the point x^* are linearly independent.*

Recall that constraint g_i is active at x^* if $g_i(x^*) = 0$. Also notice that if the point x^* is feasible, all the equality constraints h_i are active. We state the first order optimality conditions as follows.

Theorem 10.3 *Suppose that x^* is a local solution of the problem (9.1) and that LICQ holds at the point x^*. Then there are Lagrange multipliers (λ^*, μ^*) such that the KKT conditions are satisfied.*

Next, we state second order conditions. They imply that the Hessian of the Lagrangian $\nabla_x^2 L(x^*, \lambda^*, \mu^*)$ has to be positive semidefinite on the special subset of $\mathbb{R}^n$. Roughly speaking, this subset contains directions which are nearly feasible. In order to define such subset, let us start with the following subsets.

Let x^* and (λ^*, μ^*) be primal and dual variables that satisfy KKT conditions. Then

$$A_1 = \{d \in \mathbb{R}^n \mid \nabla^T h_i(x^*)d = 0, i = 1, ..., m\}, \tag{10.13}$$

$$A_2 = \{d \in \mathbb{R}^n \mid \nabla^T g_i(x^*)d = 0 \text{ for all active constraints with } \lambda_i^* > 0\},$$

$$A_3 = \{d \in \mathbb{R}^n \mid \nabla^T g_i(x^*)d \leq 0 \text{ for all active constraints with } \lambda_i^* = 0\},$$

$$A = A_1 \cap A_2 \cap A_3. \tag{10.14}$$

The directions from the above defined sets retain feasibility with respect to linear approximations of the constraints. For example, assume that $d \in A_1$. Then, for arbitrary i we have

$$h_i(x^* + d) \approx h_i(x^*) + \nabla^T h_i(x^*)d = h_i(x^*) = 0.$$

Furthermore, if the constraint h_i is linear, then $h_i(x^* + d) = 0$ and the point $x^* + d$ remains feasible with respect to the constraint h_i. Considering A_2, if the constraint g_i is linear, it remains active at $x^* + d$ for $d \in A_2$. Then the corresponding Lagrange multiplier remains positive since the complementarity condition holds. On the other hand, if $\lambda_i^* = 0$ and we consider $d \in A_3$, then we can allow g_i to became inactive, i.e., we aim for $g_i(x^* + d) \leq 0$. Notice that if a constraint is not active, then we can always find small enough step size for an arbitrary direction d such that the feasibility with respect to the constraint is not violated. The second order necessary conditions are as follows.

Theorem 10.4 *Suppose that x^* is a local solution of the problem (9.1) and that LICQ holds at the point x^*. Suppose that the Lagrange multipliers (λ^*, μ^*) are such that the KKT conditions hold. Then,*

$$d^T \nabla_x^2 L(x^*, \lambda^*, \mu^*) d \geq 0 \textit{ for all } d \in A.$$

Finally, we state the second order sufficient conditions.

Theorem 10.5 *Suppose that x^* and (λ^*, μ^*) are such that the KKT conditions are satisfied and*

$$d^T \nabla_x^2 L(x^*, \lambda^*, \mu^*) d > 0 \textit{ for all } d \in A \backslash \{0\}.$$

Then x^* is a strict local solution of the problem (9.1).

10.4 Exercises

1. Prove that the pointwise infimum of a family of affine functions is concave.
2. Show that the problem (10.5) is convex.
3. Show that the Lagrangian related to the convex problem (9.1) is also convex with respect to x, provided that λ is dual feasible.
4. Assume that x^* and (λ^*, μ^*) satisfy KKT conditions. Show that $\nabla^T f(x^*)d = 0$ for every $d \in A$ defined by (10.14).
5. State the KKT conditions and solve the following problem

$$\min_{x \in S} f(x) = (x_1 - 2)^2 + x_2^2$$

   with

   (a) $S = \{x \in \mathbb{R}^2 \mid x_1 + x_2 \geq 1\}$.
   (b) $S = \{x \in \mathbb{R}^2 \mid x_1 + x_2 \leq 1\}$.
   (c) $S = \{x \in \mathbb{R}^2 \mid x_1 + x_2 \leq 1, x_1^2 + x_2 \leq 1, x_2 - x_1 = -2\}$.

6. Analyze the optimality conditions for the following problem

$$\min_{x \in S} f(x) = c^T x,$$

   where $c \in \mathbb{R}^n$ and

   (a) $S = \{x \in \mathbb{R}^n \mid Ax = b\}$, $b \in \mathbb{R}^m, A \in \mathbb{R}^{m \times n}$.
   (b) $S = \{x \in \mathbb{R}^n \mid a^T x \leq b\}$, $b \in \mathbb{R}, a \in \mathbb{R}^n, a \neq 0$.
   (c) $S = \{x \in \mathbb{R}^n \mid d \leq x \leq u\}$, $d, u \in \mathbb{R}^n$.
   (d) $S = \{x \in \mathbb{R}^n \mid \mathbf{1}^T x = 1, x \geq 0\}$.

7. Form the dual problem of

$$\min_{Ax=b} \|x\|^2,$$

where $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m\times n}$.

8. Form the dual problem of

$$\min_{Ax\leq b} \|x\|^2,$$

where $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m\times n}$.

9. Consider the problem

$$\min_{x_1+x_2=1,\ x\in\mathbb{R}^2} \|x\|^2.$$

Solve this problem and its dual problem. Compare the optimal values. Does strong duality hold?

10. Consider the problem

$$\min_{x_1+x_2\leq 1,\ x\in\mathbb{R}^2} \|x\|^2.$$

Solve this problem and its dual problem. Compare the optimal values. Does strong duality hold?

Chapter 11

Some special classes of constrained problems

In general, solving constrained optimization problems is not an easy task. There is no algorithm which works well for all classes of problems. However, methods specialized for a certain class of problems are developed, so classifying a problem properly is an important part of problem solving followed by choice of an appropriate method. The second part of the solving procedure is to choose an appropriate method. Within this chapter, we will consider two important special classes of problems - problems with linear equality constraints and problems with the so called box constraints. We will also provide representative algorithms. The first class of problems serves as a good example of the adapted Newton method for the constrained optimization framework. The second class, box-constrained problems is of importance by itself but the box constrained problems are also encountered as subproblems in general constrained problems. More details are available in the next chapter.

11.1 Problems with linear equality constraints

Within this section we observe problems of the form

$$\min_{Ax=b} f(x), \tag{11.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^2(\mathbb{R}^n)$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ and $rank(A) = m < n$. Moreover, we will focus on the case when the considered problem is convex leaving the general (nonconvex) case for exercises.

Recall that x^* is a KKT point of the problem (11.1) if there exist $\mu^* \in \mathbb{R}^m$ such that

$$\nabla f(x^*) + A^T\mu^* = 0 \quad \text{and} \quad Ax^* = b. \tag{11.2}$$

Also recall the KKT conditions are both necessary and sufficient if the problem (11.1) is convex. This fact motivated the method described below.

11.1.1 The Newton method for constrained problems

Let us first consider the problem with quadratic objective function

$$f(x) = \frac{1}{2}x^TGx + x^Tq + c. \tag{11.3}$$

Since the problem is assumed to be convex and $\nabla^2 f(x) = G$, the matrix G is symmetric and positive semidefinite. Furthermore, the KKT conditions are reduced to

$$Gx^* + q + A^T\mu^* = 0, \quad Ax^* = b, \tag{11.4}$$

and the KKT system in the matrix form can be stated as follows

$$\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \mu^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}. \tag{11.5}$$

Therefore, solving the problem (11.1) reduces to solving the (linear) KKT system. Let us denote the matrix of that system by B. If B is nonsingular, there exists an unique pair of primal-dual optimal points (x^*, μ^*). This is true if the matrix G is positive definite on the null space of A. Of course, if the objective function is strongly convex, then G is positive definite on the whole $\mathbb{R}^n$ and the matrix B is nonsingular.

If B is singular, but the KKT system is solvable, then any solution is optimal for the original problem. Finally, if the KKT system is not solvable, then the problem (11.1) is unbounded or infeasible.

Having the quadratic case in mind, we can develop Newton method for the general case as follows.

Assume that x is a feasible point. If we want to retain the feasibility, we need a step d which belongs to the null space of A, i.e.,

$$Ad = 0. \tag{11.6}$$

Then, applying any step size α we have

$$A(x + \alpha d) = Ax + \alpha Ad = b.$$

Now, Taylor's expansion of the second order gives

$$f(x + d) \approx f(x) + \nabla^T f(x) d + \frac{1}{2} d^T \nabla^2 f(x) d := \tilde{f}(d). \tag{11.7}$$

Identifying $f(x), \nabla f(x)$ and $\nabla^2 f(x)$ with c, q and G from (11.3), respectively, and observing the subproblem

$$\min_{Ad=0} \tilde{f}(d), \tag{11.8}$$

using the same reasoning as above, we obtain the following (approximate) KKT system

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} d \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}. \tag{11.9}$$

Solving this problem yields the Newton step d and the approximate dual variable w. Notice that, besides (11.6), the Newton step satisfies

$$\nabla^2 f(x)d + A^T w = -\nabla f(x). \tag{11.10}$$

Another interpretation of the Newton step comes from linearization of the KKT conditions. Namely, the ideal step would be the one that reaches a KKT point, i.e., that satisfies $\nabla_x L(x+d, w) = 0$ for some w. Approximating the gradient of the Lagrangian by

$$\nabla_x L(x+d, w) \approx \nabla_x L(x, w) + \nabla_x^2 L(x, w)d$$

and making the right hand side equal to zero one gets (11.10). Furthermore, using the same equality we conclude that the Newton step is nonascent. Indeed, using (11.6) and the convexity we obtain

$$d^T \nabla f(x) = -d^T \nabla^2 f(x)d - d^T A^T w = -d^T \nabla^2 f(x)d \leq 0. \tag{11.11}$$

Furthermore, if we assume that f is strongly convex with parameter m then

$$d^T \nabla f(x) = -d^T \nabla^2 f(x)d \leq -m\|d\|^2. \tag{11.12}$$

Therefore, in that case, d is a descent direction whenever $\|d\| \neq 0$. However, notice that if $d = 0$ then x is a KKT point since it is feasible and (11.10) implies $A^T w + \nabla f(x) = 0$. Notice that (11.12) implies

$$\|d\|^2 \leq \frac{1}{m} d^T \nabla^2 f(x)d,$$

so $d^T \nabla^2 f(x) d$ can be viewed as a measure of optimality. The quantity

$$z(x) := \sqrt{d^T \nabla^2 f(x) d}$$

is called the Newton decrement and it is common to have the stopping criterion of the form $z(x^k) \leq \varepsilon$. Let us state the algorithm.

Algorithm 11.1

Step 0 Input parameters: Find $x^0 \in \mathbb{R}^n$ such that $Ax^0 = b$, $\beta, \eta \in (0,1), k = 0$.

Step 1 Search direction: Compute the Newton direction d^k that satisfies (11.9). If $d^k = 0$ STOP. Otherwise go to Step 2.

Step 2 Step size: Find the smallest nonnegative integer j such that $\alpha_k = \beta^j$ satisfies the Armijo condition

$$f(x^k + \alpha_k d^k) \leq f(x^k) + \eta \alpha_k \nabla^T f(x^k) d^k.$$

Step 3 Update: Set $x^{k+1} = x^k + \alpha_k d^k$, $k = k+1$ and go to Step 1.

Notice that the feasibility is retained at every iteration. Furthermore, using the descent property (11.12) and the fact that $\|d\| = 0$ implies that the corresponding x is stationary point. Following the same ideas as in the proof of Theorem 5.3 we can prove the convergence result.

Theorem 11.1 *Suppose that $f : \mathbb{R}^n \to \mathbb{R}$, f is strongly convex and bounded from bellow on the feasible set $S = \{x \in \mathbb{R}^n \mid Ax = b\}$ and $f \in C^2(S)$. Moreover, assume that the sequence of search directions $\{d^k\}_{k\in\mathbb{N}}$ is bounded. Then, either the Algorithm 11.1 terminates after finite number of iterations $\bar{k}$ at a KKT point $x^{\bar{k}}$ of the problem (11.1) or every accumulation point of the sequence $\{x^k\}_{k\in\mathbb{N}}$ is a KKT point of the problem (11.1).*

Now, let us comment the case when the Newton method is applied with possibly infeasible starting point x, i.e., when we have $Ax \neq b$ in general. Still, we would like to obtain a direction which takes us to a KKT point, so we are searching for d such that $x + d \approx x^*$ where x^* satisfies

$$Ax^* = b, \nabla f(x^*) + A^T \mu^* = 0.$$

Let us approximate $w \approx \mu^*$ and

$$\nabla f(x + d) \approx \nabla f(x) + \nabla^2 f(x) d.$$

Putting this in KKT conditions we obtain

$$\nabla f(x) + \nabla^2 f(x) d + A^T w = 0. \tag{11.13}$$

We also impose the feasibility by asking for

$$A(x + d) = b. \tag{11.14}$$

The previous two inequalities can be stated in a matrix form

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} d \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ b - Ax \end{bmatrix}. \tag{11.15}$$

So, if we have infeasible starting point x^0, one possibility is to perform one step as above, i.e., $x^1 = x^0 + d^0$ where d^0 satisfies (11.15) and then use the Algorithm 11.1 to ensure the global convergence. On the other hand, imposing the feasibility may reduce the optimality progress of the algorithm. So, let us see what happens if we apply the line search from the start to obtain the sufficient reduction and not necessary a feasible point. Let us denote the residual, i.e., the measure of infeasibility, by

$$r(x) = Ax - b.$$

Then

$$
\begin{aligned}
r(x^1) &= A(x^0 + \alpha_0 d^0) - b \\
&= Ax^0 - b + \alpha_0 A d^0 \\
&= r(x^0) + \alpha_0 (b - Ax^0) \\
&= (1 - \alpha_0) r(x^0).
\end{aligned}
$$

Applying this recursively we obtain

$$
r(x^k) = r(x^0) \prod_{i=0}^{k-1} (1 - \alpha_i). \tag{11.16}
$$

This clearly reveals that the feasibility is reached if the full step is accepted at any point. On the other hand, the sequence of iterates tends to the feasible point if, for instance, step size sequence is uniformly bounded away from zero.

11.2 Box constrained optimization

Within this subsection we consider problems of the form

$$
\min_{l \le x \le u} f(x), \tag{11.17}
$$

where $l, u \in \mathbb{R}^n_\infty$ and f is continuously differentiable on a feasible set $S = \{x \in \mathbb{R}^n : l \le x \le u\}$. Although the approach presented in this section may easily be extended to the case of a generic convex set S, we focus our attention to the box constrained case. Let us denote an orthogonal distance of point x from a set S by $dist_S(x)$, i.e.,

$$
dist_S(x) = \inf_{y \in S} \|y - x\|. \tag{11.18}
$$

If S is a box, this distance is easy to calculate. Moreover, one may use minimum instead of infimum and we define an orthogonal projection of point x on a set S by

$$P_S(x) = \arg\min_{y \in S} \|y - x\|. \tag{11.19}$$

We define the projected gradient direction by

$$d = d(x) = P_S(x - \nabla f(x)) - x. \tag{11.20}$$

This direction enjoys properties stated in Theorem 11.2. Recall that Theorem 9.2 implies that x^* is a solution of a convex problem (11.17) if and only if $\nabla^T f(x^*)(x - x^*) \geq 0$ for every $x \in S$. However, if the objective function f is not necessary convex, then we say that x^* is a (constrained) stationary point of problem (11.17) if

$$\nabla^T f(x^*)(x - x^*) \geq 0 \tag{11.21}$$

holds for every $x \in S$.

Theorem 11.2 *[9] Suppose that $f \in C^1(S)$ and $x \in S$. Then the projected gradient direction d defined by (11.20) satisfies the following:*

a) $d^T \nabla f(x) \leq -\|d\|^2$.

b) $d = 0$ *if and only if* x *is a stationary point for the problem (11.17).*

Proof. First, notice that the projection $z^* := P_S(z)$ is a solution of the following convex problem

$$\min_{y \in S} g(y) := \frac{1}{2}\|y - z\|^2. \tag{11.22}$$

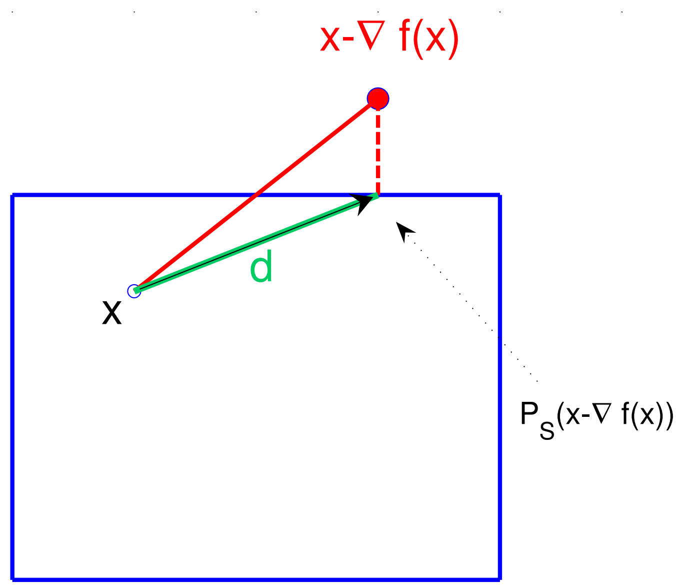

Figure 11.1: Projected gradient direction.

Therefore, using the optimality conditions stated in Theorem 9.2, we conclude that $\nabla^T g(z^*)(y - z^*) \geq 0$ for every $y \in S$. Since this inequality holds for an arbitrary z, having in mind the definition of the function g and the projection z^*, we obtain that for every $z \in \mathbb{R}^n$ and every $y \in S$

$$(P_S(z) - z)^T (y - P_S(z)) \geq 0. \tag{11.23}$$

Now, specifying $z = x - \nabla f(x)$ and $y = x$, we get

$$(P_S(x - \nabla f(x)) - x + \nabla f(x))^T (x - P_S(x - \nabla f(x))) \geq 0, \tag{11.24}$$

i.e., $-d^T d - \nabla^T f(x) d \geq 0$ which obviously implies

$$d^T \nabla f(x) \leq -\|d\|^2. \tag{11.25}$$

Let us prove the second part of the statement.

First, assume that x^* is a stationary point for the problem (11.17), i.e., (11.21) holds for every $x \in S$. Furthermore, (11.20) implies that $x^* + d(x^*) = P_S(x^* - \nabla f(x^*)) \in S$, so replacing x with $x^* + d(x^*)$ in (11.21) we obtain

$$\nabla^T f(x^*) d(x^*) \geq 0. \tag{11.26}$$

On the other hand, (11.25) implies that $\nabla^T f(x^*) d(x^*) \leq -\|d(x^*)\|^2 \leq 0$, so combining this inequality with (11.26) yields $d(x^*) = 0$.

Now we prove the reverse implication. Assume that $d(x^*) = 0$. Then, replacing z with $x^* - \nabla f(x^*)$ in (11.23) we obtain that $\nabla^T f(x^*)(y - x^*) \geq 0$ for every $y \in S$ and the statement follows. ■

Theorem 11.2 implies that d is a descent direction provided that x is not a stationary point of (11.17). Therefore, the line search algorithm can be applied as follows.

Algorithm 11.2

Step 0 Input parameters: $x^0 \in S$, $\beta, \eta \in (0, 1), k = 0$.

Step 1 Search direction: Compute the projected gradient direction d defined by (11.20). If $d^k = 0$ STOP. Otherwise go to Step 2.

Step 2 Step size: Find the smallest nonnegative integer j such that $\alpha_k = \beta^j$ satisfies the Armijo condition

$$f(x^k + \alpha_k d^k) \leq f(x^k) + \eta \alpha_k \nabla^T f(x^k) d^k.$$

Step 3 Update: Set $x^{k+1} = x^k + \alpha_k d^k$, $k = k + 1$ and go to Step 1.

Notice that the previous algorithm retains feasibility and fits into the common line search framework. Therefore, following the ideas from the proof of Theorem 5.3 and using Theorem 11.17, one can show the following result.

Theorem 11.3 *Suppose that $f : \mathbb{R}^n \to \mathbb{R}$, f is bounded from bellow on the feasible set $S = \{x \in \mathbb{R}^n \mid l \leq x \leq u\}$ and $f \in C^1(S)$. Moreover, assume that the sequence of search directions $\{d^k\}_{k \in \mathbb{N}}$ is bounded. Then, either the Algorithm 11.2 terminates after a finite number of iterations $\bar{k}$ at a stationary point $x^{\bar{k}}$ of the problem (11.17) or every accumulation point of the sequence $\{x^k\}_{k \in \mathbb{N}}$ is a stationary point of the problem (11.17).*

11.3 Exercises

1. Consider the problem (11.1) and discuss the case of condition $rank(A) = m < n$ being violated.

2. Prove Theorem 11.1.

3. Consider the problem (11.1). Let $\tilde{x}$ be a feasible point.

   (a) Show that all feasible points can be represented by $x = \tilde{x} + u$ where $u \in Null(A)$.

(b) Show that $x = \tilde{x} + \alpha u$ is feasible for all $u \in Null(A)$ and $\alpha \in \mathbb{R}$.

(c) Define the function $\varphi : \mathbb{R}^{n-m} \to \mathbb{R}$ as

$$\varphi(\gamma) = f(\tilde{x} + Z\gamma),$$

where $Z \in \mathbb{R}^{n\times(n-m)}$ represents a basis of $Null(A)$. Consider the unconstrained problem

$$\min \varphi(\gamma). \tag{11.27}$$

Show that γ^* is a local (global) solution of the problem (11.27) if and only if $x^* = \tilde{x}+Z\gamma^*$ is a local (global) solution of problem (11.1).

(d) Show that the first order optimality conditions for the problem (11.1) can be expressed as

$$Ax^* = b, Z^T\nabla f(x^*) = 0. \tag{11.28}$$

(e) Show that (11.28) is equivalent to

$$Ax^* = b, \quad \nabla f(x^*) = A^T\mu^* \quad \text{for some} \quad \mu^* \in \mathbb{R}^m. \tag{11.29}$$

(f) Show that the sufficient second order optimality conditions for the problem (11.1) can be expressed as

$$Ax^* = b, \quad Z^T\nabla f(x^*) = 0, \quad Z^T\nabla^2 f(x^*)Z \succ 0. \tag{11.30}$$

(g) Analyze the optimality conditions if the objective function is strictly convex.

4. Define $\varphi_k(\gamma) = f(x^k+Z\gamma)$ where Z is as in the previous exercise. Suppose that x^k is feasible, but it is not a solution of the problem (11.1). Define

$$d^k = -ZZ^T\nabla f(x^k).$$

(a) Prove that d^k is a descent direction for the function f at point x^k.

(b) Prove that the line search along the direction d^k maintains feasibility.

(c) Prove that (a) and (b) are also true for

$$d^k = -ZH_k\nabla\varphi_k(0)$$

for any positive definite matrix $H_k \in \mathbb{R}^{(n-m)\times(n-m)}$.

(d) Suppose that f is twice continuously differentiable and consider Newton step d^k defined through (11.9). Show that d^k also satisfies

$$d^k = Zv^k, \quad \text{where} \quad \nabla^2\varphi_k(0)v^k = -\nabla\varphi_k(0). \tag{11.31}$$

5. Consider a problem

$$\min_{x_1+2x_2+3x_3=6} x_1^2 + 3x_2^2 + 2x_3^2.$$

Let $x^0 = (1,1,1)^T$ be a starting point. Compute d^0 satisfying (11.31) on the relevant problem of the form (11.27) to obtain x^1. Check if x^1 is an optimal point.

6. Solve the previous task by applying the Newton step defined in (11.9). Do we obtain the same point x^1?

7. Consider the problem

$$\min_{Ax=b} \frac{1}{2}x^TQx - c^Tx,$$

where $Q \in \mathbb{R}^{n\times n}$ and $Q^T = Q$. Prove that x^* is a local solution if and only if it is a global solution of the considered problem.

8. Consider a problem
$$\min_{x_1+x_2=1} x_1^2 + x_2^2.$$
   (a) Find an optimal point x^*.
   (b) Consider a penalty problem
   $$\min x_1^2 + x_2^2 + \rho(x_1 + x_2 - 1)^2,$$
   where $\rho > 0$. Find an optimal point of this problem $x^*(\rho)$.
   (c) Show that $\lim_{\rho\to\infty} x^*(\rho) = x^*$.

9. Find the analytical expression for the projection function P_S if:
   (a) S is a box.
   (b) S is a unit ball.

10. Sketch the direction d defined by (11.20).

11. Show that all the iterates of the Algorithm 11.2 belong to the feasible set S provided that it is convex.

12. Prove the Theorem 11.3.

Chapter 12

Penalty methods

Within this section, we consider one of the approaches for solving constrained optimization problems of the generic form (9.1), i.e.,

$$\min_{x \in S} f(x),\ S = \{x \in \mathbb{R}^n \mid h(x) = 0,\ g(x) \leq 0\}. \tag{12.1}$$

The methods that we consider are referred to as penalty methods and they aim to solve considered problems by employing unconstrained optimization technics. For other approaches such as Interior point methods and Active set methods, one can see [16] for instance.

The penalty method transforms the problem (12.1) into the unconstrained optimization problem

$$\min_{x \in \mathbb{R}^n} \Phi(x), \tag{12.2}$$

where $\Phi(x)$ is called the penalty function, or into a problem with some simple constraints (for example, box) which can be solved by using methods of projected gradient type for instance.

The form of the penalty function is usually as follows

$$\Phi(x, \tau) = f(x) + \tau\rho(x), \tag{12.3}$$

where $\rho(x)$ represents constraint violation measure and $\tau > 0$ is the penalty parameter. So, the penalty function combines optimality and feasibility into one objective function. The function $\rho : \mathbb{R}^n \to \mathbb{R}_+$ satisfies the following condition

$$\rho(x) = 0 \iff x \in S. \tag{12.4}$$

For example, it may be defined as the distance of the point x from the feasible set S.

Most penalty methods assume that a sequence of penalty problems of the form

$$\min_{x \in \mathbb{R}^n} \Phi(x, \tau_k), \tag{12.5}$$

are solved, where the sequence of penalty parameters tends to infinity, i.e.,

$$\lim_{k \to \infty} \tau_k = \infty. \tag{12.6}$$

That way, the infeasibility is more and more penalized to force the iterates to converge to the feasible set.

If there exists penalty parameter large enough such that the problems (12.1) and (12.5) are equivalent, then we say that the penalty function is exact. More precisely, we have the following definition.

Definition 15 *The penalty function Φ is exact if there exists $\bar{\tau} > 0$ such that for all $\tau \geq \bar{\tau}$ any local solution of the problem (12.1) is a local minimizer of the penalty function $\Phi(x, \tau)$.*

An example of exact penalty function is

$$Q_1(x, \tau) = f(x) + \tau\left(\sum_{i=1}^{m} |h_i(x)| + \sum_{i=1}^{p} \max\{0, g_i(x)\}\right).$$

The definition implies that we can obtain a local solution of the constrained problem by solving a single unconstrained problem provided that the penalty parameter is large enough. However, finding a suitable penalty parameter is not that easy. So, usually one still has to solve the (finite) sequence of unconstrained optimization problems increasing the penalty parameter gradually. Another problem that arises with exact penalty function is that such functions are usually not differentiable which makes the straightforward application of the unconstrained optimization methods described in some of the previous chapters impossible.

In the following section, we consider quadratic penalty function. This penalty function is not exact, but it is convenient since it is differentiable - at least in the case of equality constraints.

12.1 Quadratic penalty function

Quadratic penalty function takes the following form

$$Q(x,\tau) = f(x) + \frac{\tau}{2}\left(\sum_{i=1}^{m}(h_i(x))^2 + \sum_{i=1}^{p}(\max\{0, g_i(x)\})^2\right). \tag{12.7}$$

Notice that if the inequality constraints are present, the quadratic penalty function may be less smooth than the objective and the constraints functions. However, there are some techniques that can transform inequality into equality constraints as we will see latter. Therefore, we restrict our attention to the equality constrained case

$$\min_{h(x)=0} f(x). \tag{12.8}$$

The framework algorithm is given below.

Algorithm 12.1

Step 0 Input parameters: Take $x^0 \in \mathbb{R}^n, \varepsilon_0 \geq 0, \tau_0 > 0, k = 0.$

Step 1 Initialization: $x^0_{start} = x^0$.

Step 2 Solve the subproblem $\min Q(x, \tau_k)$ approximately: Start with x^k_{start}, terminate when

$$\|\nabla_x Q(x^k, \tau_k)\| \le \varepsilon_k. \tag{12.9}$$

Step 3 Update the penalty parameter: Choose $\tau_{k+1} > \tau_k$.

Step 4 Update the tolerance: Choose $\varepsilon_{k+1} \in [0, \varepsilon_k)$.

Step 5 Update the starting point: Set $x^{k+1}_{start} = x^k$.

Step 6 Set $k = k + 1$ and go to Step 2.

Large penalty parameter may yield difficult subproblems to solve if the derivatives become ill-conditioned. Thus, the starting penalty parameter is usually rather modest. The rule presented in Step 5 is often called the warm start as the solution of the previous subproblem serves as an initial approximation for the subsequent problem. This strategy may reduce the number of inner (subproblem) iterations significantly.

Notice that the subproblems do not have to be solved exactly. However, if we assume to have exact subproblem solutions, we obtain the following result.

Theorem 12.1 *Suppose that $f, h \in C^1(\mathbb{R}^n)$ and that each x^k is the exact global minimizer of function $Q(x, \tau_k)$. Suppose that (12.6) holds. Then every accumulation point of the sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by Algorithm 12.1 is a solution of the problem (12.8).*

Proof. Let $\bar{x}$ be a global solution of the problem (12.8). Then there holds that $h(\bar{x}) = 0$ and $f(\bar{x}) \le f(x)$ for every $x \in S$. Now, since x^k is a global minimizer of $Q(x, \tau_k)$, there holds

$$Q(x^k, \tau_k) \le Q(\bar{x}, \tau_k), \tag{12.10}$$

so we conclude

$$f(x^k) + \frac{1}{2}\tau_k\|h(x^k)\|^2 \leq f(\bar{x}) + \frac{1}{2}\tau_k\|h(\bar{x})\|^2 = f(\bar{x}). \tag{12.11}$$

The last inequality further implies

$$\|h(x^*)\|^2 \leq \frac{2}{\tau_k}(f(\bar{x}) - f(x^k)). \tag{12.12}$$

Now, let x^* be an arbitrary accumulation point of the sequence $\{x^k\}_{k\in\mathbb{N}}$, i.e., $\lim_{k\in K} x^k = x^*$ for some $K \subseteq \mathbb{N}$. Taking the limit over K in (12.12) we obtain

$$\|h(x^*)\|^2 \leq (\lim_{k\in K}\frac{2}{\tau_k})(f(\bar{x}) - f(x^*)) = 0. \tag{12.13}$$

Thus, $h(x^*) = 0$ and we conclude that x^* is a feasible point.

On the other hand, (12.11) implies

$$f(x^k) \leq f(\bar{x}) - 0.5\tau_k\|h(x^k)\|^2 \leq f(\bar{x}). \tag{12.14}$$

Again, taking the limit over K and using the continuity argument we obtain $f(x^*) \leq f(\bar{x})$ which together with the feasibility implies that x^* is a global solution of the problem (12.8). ■

Now, let us assume that the subproblems are solved approximately with the increasing accuracy. Then the following can be proved.

Theorem 12.2 *Suppose that $f, h \in C^1(\mathbb{R}^n)$ and that $\lim_{k\to\infty}\varepsilon_k = 0$. Suppose that (12.6) holds. Then every accumulation point x^* of the sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by Algorithm 12.1 at which LICQ holds is a KKT point of the problem (12.8). Moreover, Lagrange multipliers associated with $x^* = \lim_{k\in K} x^k$ are given by*

$$\lim_{k\in K}\tau_k h(x^k) = \mu^*. \tag{12.15}$$

Proof. Consider (12.9). This inequality implies

$$\tau_k \| \sum_{i=1}^{m} h_i(x^k) \nabla h_i(x^k) \| - \|\nabla f(x^k)\| \leq \varepsilon_k \tag{12.16}$$

and

$$\| \sum_{i=1}^{m} h_i(x^k) \nabla h_i(x^k) \| \leq \frac{1}{\tau_k} (\|\nabla f(x^k)\| + \varepsilon_k). \tag{12.17}$$

Taking the limit over K we obtain

$$\sum_{i=1}^{m} h_i(x^*) \nabla h_i(x^*) = 0$$

and LICQ implies that $h(x^*) = 0$, i.e., x^* is feasible.

Now, denote the Jacobian of the equality constraints by A, i.e., $A(x) = \nabla h(x)$ and define

$$\mu^k := \tau_k h(x^k). \tag{12.18}$$

Then, using the definition of quadratic penalty function we obtain

$$\nabla f(x^k) - \nabla Q_x(x^k, \tau_k) = - \sum_{i=1}^{m} \tau_k h_i(x^k) \nabla h_i(x^k) = -A^T(x^k) \mu^k. \tag{12.19}$$

Multiplying the previous inequality by $A(x^k)$ from the left we obtain

$$-A(x^k) A^T(x^k) \mu^k = A(x^k)(\nabla f(x^k) - \nabla_x Q(x^k, \tau_k)).$$

Furthermore, LICQ implies that $rank(A(x^*)) = m$, i.e., $A(x^*)A^T(x^*)$ is nonsingular. Using the continuity of the Jacobian we conclude that $A(x^k)A^T(x^k)$ is also nonsingular for $k \in K$ large enough and

$$\mu^k = -(A(x^k)A^T(x^k))^{-1} A(x^k)(\nabla f(x^k) - \nabla_x Q(x^k, \tau_k)). \tag{12.20}$$

Now, (12.9) implies that $\lim_{k \in K} \nabla_x Q(x^k, \tau_k) = 0$ and taking the limit over K in (12.20) we conclude that

$$\mu^* = \lim_{k \in K} \mu^k = -(A(x^*)A^T(x^*))^{-1} A(x^*) \nabla f(x^*).$$

Finally, returning to (12.19) and taking the limit over K we obtain

$$0 = \nabla f(x^*) + A^T(x^*)\mu^* = \nabla f(x^*) + \nabla^T h(x^*)\mu^* = \nabla_x L(x^*, \mu^*)$$

which together with the feasibility of x^* implies that x^* is the KKT point of the problem (12.8) and μ^* is the corresponding vector of Lagrange multipliers. ∎

Now, let us consider the general case (12.1), with both equality and inequality constraints. One way to cope with this problem is to introduce the so called slack variables $s \in \mathbb{R}^p$ and form the equivalent problem

$$\min_{y \in \tilde{S}} f(x), \;\; \tilde{S} = \{y \in \mathbb{R}^{n+p} \mid y = (x; s), \;\; h(x) = 0, \;\; g(x) + s = 0, \;\; s \geq 0\}. \tag{12.21}$$

Then, quadratic penalty function takes the form

$$Q(y, \tau) = f(x) + \frac{\tau}{2}(\|h(x)\|^2 + \|g(x) + s\|^2)$$

and every subproblem has nonnegativity constraints $s \geq 0$, i.e., we have

$$\min_{s \geq 0} Q(y, \tau_k). \tag{12.22}$$

Such subproblem can be solved by means of algorithms presented in Section 11.2 for instance. Then, the framework algorithm (Algorithm 12.1) has to be altered since the stopping criterion in Step 2 is no longer valid. Following the results from Theorem 11.2, one possible choice would be to stop when the norm of the projected gradient direction is smaller than the tolerance ε_k.

12.2 Augmented Lagrangian method

The Augmented Lagrangian method can be viewed as a combination of the quadratic penalty function and the Lagrangian function. One of its main advantages is that it reduces the possibility of ill-conditioned subproblems by a specific update of Lagrange multipliers.

Let us consider the equality constrained problem (12.8). Then the Augmented Lagrangian function takes the form

$$L_A(x, \mu, \tau) = f(x) - \sum_{i=1}^{m} \mu_i h_i(x) + \frac{1}{2}\tau \sum_{i=1}^{m} (h_i(x))^2. \tag{12.23}$$

Let us consider the gradient of the Augmented Lagrangian function

$$\nabla_x L_A(x, \mu, \tau) = \nabla f(x) - \sum_{i=1}^{m} (\mu_i - \tau h_i(x)) \nabla h_i(x). \tag{12.24}$$

Looking at the proof of Theorem 12.2 (see (12.19)) and keeping in mind that the values associated with $\nabla h_i(x)$ represent estimates of the Lagrange multipliers, following the same ideas we conclude that

$$\mu_i^* \approx \mu_i - \tau h_i(x).$$

Therefore, the Lagrange multipliers update is as follows

$$\mu^{k+1} = \mu^k - \tau h(x^k). \tag{12.25}$$

The framework algorithm is similar to Algorithm 12.1.

Algorithm 12.2

Step 0 Input parameters: Find $x^0 \in \mathbb{R}^n, \varepsilon_0 \geq 0, \tau_0 > 0, \mu^0 \in \mathbb{R}^m$.

Step 1 Initialization: $k = 0$, $x^0_{start} = x^0$, $\varepsilon_k = \varepsilon_0$, $\tau_k = \tau_0$, $\mu^k = \mu^0$.

Step 2 Solve the subproblem $\min L_A(x, \mu^k, \tau_k)$ approximately: Start with x^k_{start}, terminate when

$$\|\nabla_x L_A(x^k, \mu^k, \tau_k)\| \leq \varepsilon_k. \tag{12.26}$$

Step 3 Update the penalty parameter: Choose $\tau_{k+1} > \tau_k$.

Step 4 Update the tolerance: Choose $\varepsilon_{k+1} \in [0, \varepsilon_k)$.

Step 5 Update the starting point: Set $x^{k+1}_{start} = x^k$.

Step 6 Update the Lagrange multiplier: Set

$$\mu^{k+1} = \mu^k - \tau h(x^k).$$

Step 7 Set $k = k + 1$ and go to Step 2.

Next, we state that the Augmented Lagrangian is an exact-type penalty function under the appropriate conditions.

Theorem 12.3 *Let x^* be a local solution of the problem (12.8) at which LICQ holds and the second order conditions given in Theorem 10.5 are satisfied with the Lagrange multiplier μ^*. Then there exists $\tau^* > 0$ such that for all $\tau \geq \tau^*$ x^* is a strict local minimizer of $L_A(x, \mu^*, \tau)$.*

Now, let us consider the case with inequality constraints. For simplicity, let us restrict our attention to the following problem

$$\min_{g(x) \leq 0} f(x). \tag{12.27}$$

The general case (9.1) may be approached following the same ideas. Similar to the quadratic penalty case, one way to approach this problem is to add the slack variables and obtain an equivalent problem of the form

$$\min_{y \in \tilde{S}} f(x),\ \tilde{S} = \{y \in \mathbb{R}^{n+p} \mid y = (x; s),\ g(x) + s = 0,\ s \geq 0\}. \tag{12.28}$$

The Augmented Lagrangian associated with this problem is

$$L_A(y,\mu,\tau) = f(x) - \sum_{i=1}^{p} \mu_i(g_i(x)+s_i) + 0.5\tau \sum_{i=1}^{p} (g_i(x)+s_i)^2 \quad (12.29)$$

and the subproblem is of the form

$$\min_{s \geq 0} L_A(y, \mu^k, \tau_k). \quad (12.30)$$

Now, there are two usual ways to proceed. One is to modify the Augmented Lagrangian algorithm in the same manner as it was done for the quadratic penalty method taking into account box constraints that yield different stopping criteria with respect to (12.26). On the other hand, notice that the corresponding Augmented Lagrangian function is quadratic and convex with respect to s and solving the problem (12.30) with respect to s (with fixed x) one obtains that the optimal solution (see the Exercise 9.2, task 3.) is given by

$$s(x) = \max\{\frac{1}{\tau_k}\mu^k - g(x), 0\}. \quad (12.31)$$

That way we obtain the subproblem

$$\begin{aligned} \min_x \tilde{L}_A(x,\mu^k,\tau_k) \quad &= \quad f(x) - \sum_{i=1}^{p} \mu_i^k (g_i(x) + \max\{\frac{1}{\tau_k}\mu_i^k - g_i(x), 0\}) \\ &+ \quad 0.5\tau_k \sum_{i=1}^{p} (g_i(x) + \max\{\frac{1}{\tau_k}\mu_i^k - g_i(x), 0\})^2. \end{aligned}$$

The obtained problem is unconstrained, but nonsmooth in general so one may have to use some generalized gradient method to cope with this problem.

Finally, let us see how the Lagrange multipliers associated with the inequality constraints are updated. The Augmented Lagrangian related directly to the problem (12.27) is

$$L_A(x, \lambda^k, \tau_k) = f(x) + \sum_{i=1}^{p} \lambda_i^k g_i(x) + 0.5\tau_k \sum_{i=1}^{p} (\max\{0, g_i(x)\})^2. \tag{12.32}$$

It is nondifferentiable, but using the fact that

$$\max\{0, g_i(x)\} = \begin{cases} 0 & g_i(x) \leq 0 \\ g_i(x) & g_i(x) > 0 \end{cases}$$

the derivative of the Augmented Lagrangian can be represented by

$$\nabla_x L_A(x, \lambda^k, \tau_k) = \nabla f(x) + \sum_{i=1}^{p} \lambda_i^k \nabla g_i(x) + \tau_k \sum_{i=1}^{p} \max\{0, g_i(x)\} \nabla g_i(x).$$

Arranging this further we obtain

$$\nabla_x L_A(x, \lambda^k, \tau_k) = \nabla f(x) + \sum_{i=1}^{p} (\lambda_i^k + \tau_k \max\{0, g_i(x)\}) \nabla g_i(x).$$

Following the same ideas as before, we consider the terms multiplying $\nabla g_i(x)$ as the approximates of the optimal Lagrange multipliers and set

$$\lambda^{k+1} = \lambda^k + \tau_k \max\{0, g(x^k)\}.$$

This way, we retain the nonnegativity of the relevant multipliers λ^k if the starting guess is nonnegative.

12.3 Exercises

1. Consider the problem (12.8) and the penalty function of the form

$$\Phi(x, l, u) = f(x) + \sum_{i=1}^{m} u_i e^{l_i h_i(x)/u_i},$$

where $l_i \in \mathbb{R}, u_i > 0$ for $i = 1, ..., m$. Assume that $f, h \in C^1$. Let x^* be a solution of the problem (12.8) with multipliers l^* such that LICQ holds. Prove that x^* is a stationary point of $\Phi(x, l^*, u)$.

2. Consider a problem
$$\min_{x \in S} f(x). \tag{12.33}$$
Let ρ be a nonnegative function defined as in (12.4). Suppose that the penalty function $\Phi(x, \tau) = f(x) + \tau\rho(x)$ has a global minimizer x^* for $\tau = \tau^*$ and assume that $x^* \in S$. Prove that x^* is a global solution of (12.33).

3. Let x^* be a global minimizer of f over S. Assume that $\bar{x}$ is not feasible. Prove that there is a penalty parameter $\bar{\tau} > 0$ such that $\Phi(x^*, \tau) \le \Phi(\bar{x}, \tau)$ for every $\tau \ge \bar{\tau}$ where Φ and ρ are as in the previous exercise.

4. Consider the problem (12.8) where $f, h \in C^1$. Let x^* be a regular solution of that problem. Suppose that at least one of the Lagrange multipliers associated with x^* is not equal to zero. Prove that there is no finite τ such that x^* is a local minimizer of the relevant quadratic penalty function.

5. State the Augmented Lagrangian algorithm for the general case with both equality and inequality constraints - use both approaches described in Section 12.2.

6. Prove that the optimal solution of the problem
$$\min_{s \ge 0} q(s) := L_A((x; s), \mu^k, \tau_k)$$
is given by (12.31).

Bibliography

[1] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.

[2] E. G. Birgin, J.M. Martínez, Practical Augmented Lagrangian Methods for Constrained Optimization, SIAM 2014.

[3] J.R. Dennis, R.B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equation, SIAM 1996.

[4] C.D. Meyer, Matrix Analysis and Applied Linear Algebra, SIAM ???

[5] L.N. Trefeten, D. Bau, Numerical Linear Algebra, SIAM ???

[6] C.T. Kelley, Iterative Methods for Linear and Nonlinear Equations, SIAM, 1997?

[7] A. Beck, Introduction to Nonlinear Optimization, Theory, ALgorithms and Applications with Matlab, SIAM, 2014

[8] E. G. Birgin, N. Krejić, J. M. Martínez, Globaly convergent inexact quasi-Newton methods for solving nonlinear systems, *Numerical Algorithms, 32 (2003), pp. 249-260.*

[9] E. G. Birgin, J. M. Martínez, M. Raydan, Nonmonotone Spectral Projected Gradient Methods on Convex Sets, *SIAM Journal on Optimization, Vol. 10, Issue 4 (2006), pp.1196-1211.*

[10] A. R. Conn, N. I. M. Gould, Ph. L. Toint, Trust-region Methods, *SIAM, 2000.*

[11] A. Friedlander, Elementos de programacao nao-linear, Unicamp, SP, Brasil, `https://www.ime.unicamp.br/~friedlan/livro.htm`

[12] A. R. Conn, K. Scheinberg, L. N. Vicente, Introduction to Derivative-Free Optimization, *MPS-SIAM Book Series on Optimization, SIAM, Philadelphia, 2009.*

[13] E. D. Dolan, J. J. Moré, Benchmarking optimization software with performance profiles, *Mathematical Programming, Ser. A 91 (2002), pp. 201-213*

[14] M. P. Friedlander, M. Schmidt, Hybrid deterministic-stochastic methods for data fitting, *SIAM Journal on Scientific Computing, 34 (3) (2012), pp. 1380-1405.*

[15] N. Krejić, N. Krklec, Line search methods with variable sample size for unconstrained optimization, *Journal of Computational and Applied Mathematics, Vol. 245 (2013), pp. 213-231.*

[16] J. Nocedal, S. J. Wright, Numerical Optimization, *Springer, 1999.*

[17] M. Raydan, The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem, *SIAM Journal on Optimization, 7 (1997), pp. 26-33.*

[18] J. C. Spall, Introduction to Stochastic Search and Optimization, *Wiley-Interscience Series in Discrete Mathematics, New Jersey, 2003.*

[19] http://www.uni-graz.at/imawww/kuntsevich/solvopt/results /moreset.html.