The Art of Optimization

Alessandro Astolfi

First draft – August 2018

Preface

This book in an introduction to the *art of optimization.* I have deliberately used the word *art* in the title and placed an artwork in the title page to express that optimization requires imagination, skills and vision. The artwork has been painted by my wife, Elisabetta, and illustrates an optimization problem: one of those problems that we solve every day without even realizing their "technical" nature. I leave it to the reader to guess what the actual problem is. However, if you have attended my Optimization Lectures on a Friday morning in the past 15 years, you have probably solved the problem every week.

This book is the result of the lectures I have given at Imperial College London for Undergraduate, Master and Ph.D. students of all engineering departments (and also of the Mathematics and Physics Departments). I am not an expert in optimization, in the sense that my research activity has only seldom touched upon optimization problems, but I do believe that understanding optimization is essential for all engineers, practitioners and for everyday life. My research work is focused on systems and control: this is why some of the exercises contain a *systems and control* perspective of optimization problems. It is not hard to see that notions such as stationary points and equilibria, convergence and stability, speed of convergence and convergence rate (the former from optimization, the latter from systems and control) are fundamentally identical and one could borrow ideas and tools from systems and control theory to understand optimization problems and design optimization algorithms. Whenever possible, and in particular in the exercises, I have made this connection. Clearly, there are much deeper connections and relations which I have not discussed.

I conclude the preface with two observations. The first one is that the most difficult step in the art of optimization is the formulation of the problem. Only problems which are carefully formulated and in which all physical and engineering insight is captured yield underlying optimization problems for which one can attempt to find a meaningful solution. The second one is that optimization is a much wider *art* than that described in these books: my objective is to stimulate the interest of the reader and open their eyes to a continuously expanding body of knowledge.

Contents

Chapter 1

Introduction

1.1 Introduction

Optimization is the act of achieving the best possible result under given circumstances. In design, construction, maintenance, ..., engineers have to take decisions. The goal of all such decisions is either to minimize effort or to maximize benefit.
The effort or the benefit can be usually expressed as a function of certain design variables. Hence, optimization is the process of finding the conditions that give the maximum or the minimum value of a function.
It is obvious that if a point $x^\star$ corresponds to the minimum value of a function $f(x)$, the same point corresponds to the maximum value of the function $-f(x)$. Thus, optimization can be taken to be minimization.
There is no single method available for solving all optimization problems efficiently. Hence, a number of methods have been developed for solving different types of problems.
Optimum seeking methods are also known as mathematical programming techniques, which are a branch of operations research. Operations research is *coarsely* composed of the following areas.

- Mathematical programming methods. These are useful in finding the minimum of a function of several variables under a prescribed set of constraints.
- Stochastic process techniques. These are used to analyze problems which are described by a set of random variables of known distribution.
- Statistical methods. These are used in the analysis of experimental data and in the construction of empirical models.

These lecture notes deal mainly with the theory and applications of mathematical programming methods. Mathematical programming is a vast area of mathematics and engineering. It includes

- calculus of variations and optimal control;
- linear, quadratic and non-linear programming;
- geometric programming;
- integer programming;
- network methods (PERT);
- game theory.

The foundations of optimization can be traced back to Newton, Lagrange and Cauchy. The development of differential methods for optimization was possible because of the contribution of Newton and Leibnitz. The foundations of the calculus of variations were laid by Bernoulli, Euler, Lagrange and Weierstrasse. Constrained optimization was first studied by Lagrange and the notion of descent was introduced by Cauchy.

Despite these early contributions, very little progress was made till the 20th century, when computer power made the implementation of optimization procedures possible and this in turn stimulated further research methods.
The major developments in the area of numerical methods for unconstrained optimization have been made in the UK. These include the development of the simplex method (Dantzig, 1947), the principle of optimality (Bellman, 1957), necessary and sufficient conditions of optimality (Kuhn and Tucker, 1951).
Optimization in its broadest sense can be applied to solve any engineering problem, *e.g.*

- design of aircraft for minimum weight;
- optimal (minimum time) trajectories for space missions;
- minimum weight design of structures for earthquake;
- optimal design of electric networks;
- optimal production planning, resources allocation, scheduling;
- shortest route;
- design of optimum pipeline networks;
- minimum processing time in production systems;
- optimal control.

1.2 Statement of an optimization problem

An optimization, or a mathematical programming problem can be stated as follows.
Find

$$x = (x^1, x^2, ..., x^n)$$

which minimizes

$$f(x)$$

subject to the constraints

$$g_j(x) \leq 0 \tag{1.1}$$

for $j = 1, \ldots, m$, and

$$l_j(x) = 0 \tag{1.2}$$

for $j = 1, \ldots, p$.
The variable x is called the design vector, $f(x)$ is the objective function, $g_j(x)$ are the inequality constraints and $l_j(x)$ are the equality constraints. The number of variables n and the number of constraints $p + m$ need not be related. If $p + m = 0$ the problem is called an unconstrained optimization problem.

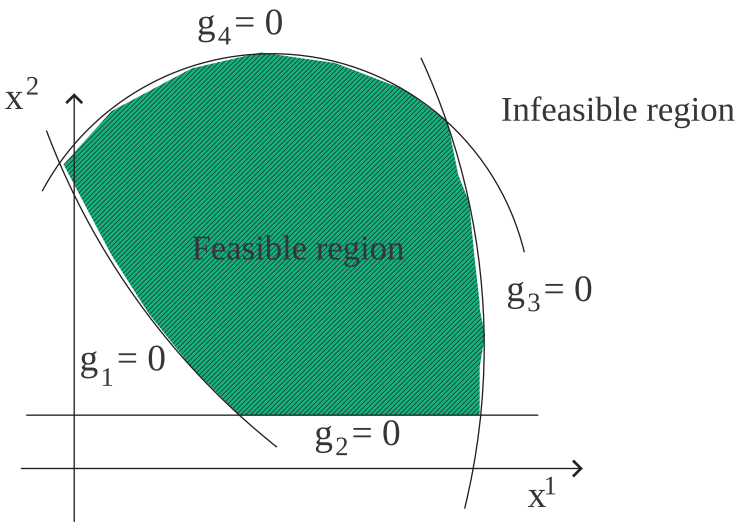


Figure 1.1: Feasible region in a two-dimensional design space. Only inequality constraints are present.

1.2.1 Design vector

Any system is described by a set of quantities, some of which are viewed as variables during the design process, and some of which are preassigned parameters or are imposed by the *environment*. All the quantities that can be treated as variables are called design or decision variables, and are collected in the design vector x.

1.2.2 Design constraints

In practice, the design variables cannot be selected arbitrarily, but have to satisfy certain requirements. These restrictions are called design constraints. Design constraints may represent limitation on the performance or behaviour of the system or physical limitations. Consider, for example, an optimization problem with only inequality constraints, *i.e.* $g_j(x) \leq 0$. The set of values of x that satisfy the equations $g_j(x) = 0$ forms a hypersurface in the design space, which is called constraint surface. In general, if n is the number of design variables, the constraint surface is an $n-1$ dimensional surface. The constraint surface divides the design space into two regions: one in which $g_j(x) < 0$ and one in which $g_j(x) > 0$. The points x on the constraint surface satisfy the constraint critically, whereas the points x such that $g_j(x) > 0$, for some j, are infeasible, *i.e.* are unacceptable, see Figure 1.1.

1.2.3 Objective function

The classical design procedure aims at finding an acceptable design, *i.e.* a design which satisfies the constraints. In general there are several acceptable designs, and the purpose

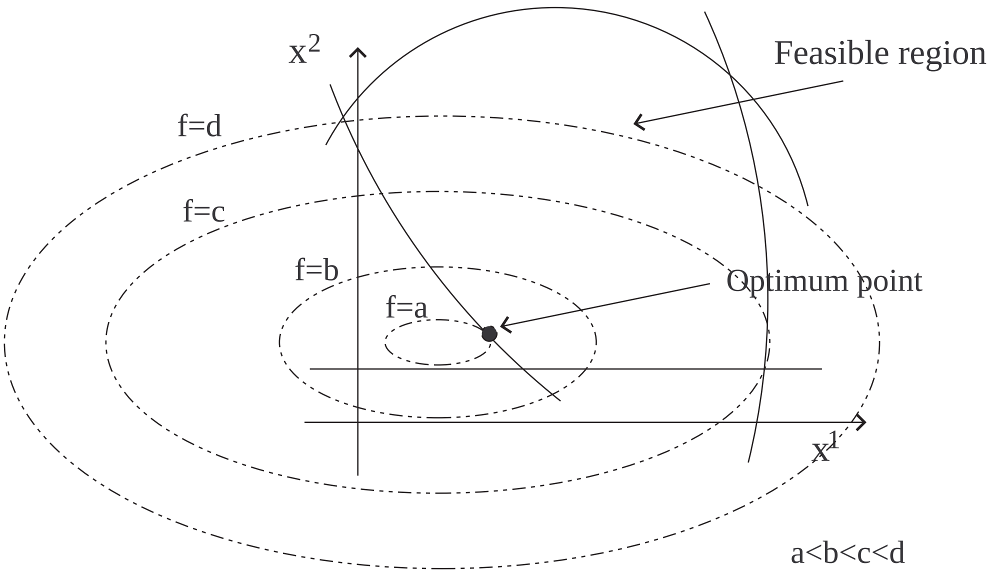


Figure 1.2: Design space, objective functions surfaces, and optimum point.

of the optimization is to single out the best possible design. Thus, a criterion has to be selected for comparing different designs. This criterion, when expressed as a function of the design variables, is known as objective function. The objective function is in general specified by physical or economical considerations. However, the selection of an objective function is not trivial, because what is the optimal design with respect to a certain criterion may be unacceptable with respect to another criterion. Typically there is a trade off performance–cost, or performance–reliability, hence the selection of the objective function is one of the most important decisions in the whole design process. If more than one criterion has to be satisfied we have a multiobjective optimization problem, that may be approximately solved considering a cost function which is a weighted sum of several objective functions.

Given an objective function $f(x)$, the locus of all points x such that $f(x) = c$ forms a hypersurface. For each value of c there is a different hypersurface. The set of all these surfaces are called objective function surfaces.

Once the objective function surfaces are drawn, together with the constraint surfaces, the optimization problem can be easily solved, at least in the case of a two dimensional decision space, as shown in Figure 1.2. If the number of decision variables exceeds two or three, this graphical approach is not viable and the problem has to be solved as a mathematical problem. Note however that more general problems have similar geometrical properties of two or three dimensional problems.

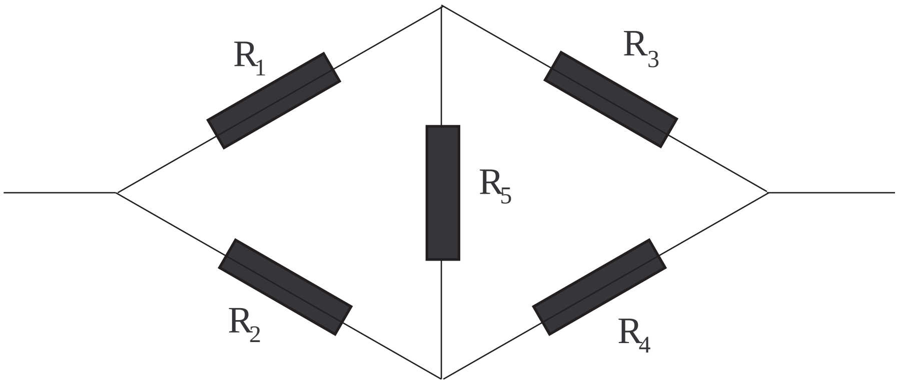


Figure 1.3: Electrical bridge network.

1.3 Classification of optimization problems

Optimization problem can be classified in several ways.

- Existence of constraints. An optimization problem can be classified as a constrained or an unconstrained one, depending upon the presence or not of constraints.
- Nature of the equations. Optimization problems can be classified as linear, quadratic, polynomial, non-linear depending upon the nature of the objective functions and the constraints. This classification is important, because computational methods are usually selected on the basis of such a classification, *i.e.* the nature of the involved functions dictates the type of solution procedure.
- Admissible values of the design variables. Depending upon the values permitted for the design variables, optimization problems can be classified as integer or real valued, and deterministic or stochastic.

1.4 Examples

Example 1 A travelling salesman has to visit n towns. He plans to start from a particular town numbered 1, visit each one of the other $n-1$ towns, and return to the town 1. The distance between town i and j is given by d_{ij}. How should he select the sequence in which the towns are visited to minimize the total distance travelled?

Example 2 The bridge network in Figure 1.3 consists of five resistors R_i, $i = 1, \ldots, 5$. Let I_i be the current through the resistance R_i, find the values of R_i so that the total dissipated power is minimum. The current I_i can vary between the lower limit $\underline{I}_i$ and the upper limit $\bar{I}_i$ and the voltage drop $V_i = R_i I_i$ must be equal to a constant c_i.

Example 3 A manufacturing firm produces two products, A and B, using two limited resources, 1 and 2. The maximum amount of resource 1 available per week is 1000 and the

Article type	w_i	v_i	c_i
1	4	9	5
2	8	7	6
3	2	4	3

Table 1.1: Properties of the articles to load.

maximum amount of resource 2 is 250. The production of one unit of A requires 1 unit of resource 1 and 1/5 unit of resource 2. The production of one unit of B requires 1/2 unit of resource 1 and 1/2 unit of resource 2. The unit cost of resource 1 is $1 - 0.0005u_1$, where u_1 is the number of units of resource 1 used. The unit cost of resource 2 is $3/4 - 0.0001u_2$, where u_2 is the number of units of resource 2 used. The selling price of one unit of A is

$$2 - 0.005x_A - 0.0001x_B$$

and the selling price of one unit of B is

$$4 - 0.002x_A - 0.01x_B,$$

where x_A and x_B are the number of units of A and B sold. Assuming that the firm is able to sell all manufactured units, maximize the weekly profit.

Example 4 A cargo load is to be prepared for three types of articles. The weight, w_i, volume, v_i, and value, c_i, of each article is given in Table 1.1.
Find the number of articles x_i selected from type i so that the total value of the cargo is maximized. The total weight and volume of the cargo cannot exceed 2000 and 2500 units respectively.

Example 5 There are two types of gas molecules in a gaseous mixture at equilibrium. It is known that the Gibbs free energy

$$G(x) = c_1x^1 + c_2x^2 + x^1 log(x^1/x_T) + x^2 log(x^2/x_T),$$

with $x_T = x^1 + x^2$ and c_1, c_2 known parameters depending upon the temperature and pressure of the mixture, has to be minimum in these conditions. The minimization of $G(x)$ is also subject to the mass balance equations:

$$x^1a_{i1} + x^2a_{i2} = b_i,$$

for $i = 1, \ldots, m$, where m is the number of atomic species in the mixture, b_i is the total weight of atoms of type i, and a_{ij} is the number of atoms of type i in the molecule of type j. Show that the problem of determining the equilibrium of the mixture can be posed as an optimization problem.

Chapter 2

Unconstrained optimization

2.1 Introduction

Several engineering, economic and planning problems can be posed as optimization problems, *i.e.* as the problem of determining the points of minimum of a function (possibly in the presence of conditions on the decision variables). Moreover, also numerical problems, such as the problem of solving systems of equations or inequalities, can be posed as an optimization problem.
We start with the study of optimization problems in which the decision variables are defined in $I\!R^n$: unconstrained optimization problems. More precisely we study the problem of determining local minimizers for differentiable functions. Although these methods are seldom used in applications, as in real problems the decision variables are subject to constraints, the techniques of unconstrained optimization are instrumental to solve more general problems: the knowledge of good methods for local unconstrained minimization is a necessary pre-requisite for the solution of constrained and global minimization problems.
The methods that will be studied can be classified from various points of view. The most interesting classification is based on the information available on the function to be optimized, namely

- methods without derivatives (direct search, finite differences);
- methods based on the knowledge of the first derivatives (gradient, conjugate directions, quasi-Newton);
- methods based on the knowledge of the first and second derivatives (Newton).

2.2 Definitions and existence conditions

Consider the following general optimization problem.

Problem 1 *Minimize*

$$f(x) \qquad \text{subject to } x \in \mathcal{F}$$

in which $f : I\!R^n \to I\!R$ *and*[1] $\mathcal{F} \subset I\!R^n$.

With respect to this problem we introduce the following definitions.

Definition 1 *A point $x \in \mathcal{F}$ is a global minimizer for the Problem 1 if*

$$f(x) \leq f(y)$$

for all $y \in \mathcal{F}$.
A point $x \in \mathcal{F}$ is a strict (or isolated) global minimizer for the Problem 1 if

$$f(x) < f(y)$$

[1]The set $\mathcal{F}$ may be specified by equations of the form (1.1) and/or (1.2).

for all $y \in \mathcal{F}$ and $y \neq x$.
A point $x \in \mathcal{F}$ is a local minimiser for the Problem 1 if there exists $\rho > 0$ such that

$$f(x) \le f(y)$$

for all $y \in \mathcal{F}$ such that $\|y - x\| < \rho$.
A point $x \in \mathcal{F}$ is a strict (or isolated) local minimizer for the Problem 1 if there exists $\rho > 0$ such that

$$f(x) < f(y)$$

for all $y \in \mathcal{F}$ such that $\|y - x\| < \rho$ and $y \neq x$.

Definition 2 *If $x \in \mathcal{F}$ is a local minimizer for the Problem 1 and if x is in the interior of $\mathcal{F}$ then x is an unconstrained local minimizer of f in $\mathcal{F}$.*

The following result provides a sufficient, but not necessary, condition for the existence of a global minimum for Problem 1.

Proposition 1 *Let $f : I\!R^n \to I\!R$ be a continuous function and let $\mathcal{F} \subset I\!R^n$ be a compact set[2]. Then there exists a global minimum of f in $\mathcal{F}$.*

In unconstrained optimization problems the set $\mathcal{F}$ coincides with $I\!R^n$, hence the above statement cannot be used to establish the existence of global minima. To address the existence problem it is necessary to consider the structure of the level sets of the function f. See also Section 1.2.3.

Definition 3 *Let $f : I\!R^n \to I\!R$. A level set of f is any non-empty set described by*

$$\mathcal{L}(\alpha) = \{x \in I\!R^n \ : \ f(x) \le \alpha\},$$

with $\alpha \in I\!R$.

For convenience, if $x_0 \in I\!R^n$ we denote with $\mathcal{L}_0$ the level set $\mathcal{L}(f(x_0))$. Using the concept of level sets it is possible to establish a simple sufficient condition for the existence of global solutions for an unconstrained optimization problem.

Proposition 2 *Let $f : I\!R^n \to I\!R$ be a continuous function. Assume there exists $x_0 \in I\!R^n$ such that the level set $\mathcal{L}_0$ is compact. Then there exists a point of global minimum of f in $I\!R^n$.*

Proof. By Proposition 1 there exists a global minimizer $x_\star$ of f in $\mathcal{L}_0$, *i.e.* $f(x_\star) \le f(x)$ for all $x \in \mathcal{L}_0$. However, if $x \notin \mathcal{L}_0$ then $f(x) > f(x_0) \ge f(x_\star)$, hence $x_\star$ is a global minimizer of f in $I\!R^n$. $\triangleleft$

It is obvious that the structure of the level sets of the function f plays a fundamental role in the solution of Problem 1. The following result provides a necessary and sufficient condition for the compactness of all level sets of f.

[2] A compact set is a bounded and closed set.

Proposition 3 *Let $f : I\!R^n \to I\!R$ be a continuous function. All level sets of f are compact if and only if for any sequence $\{x_k\}$ one has*

$$\lim_{k\to\infty} \|x_k\| = \infty \quad \Rightarrow \quad \lim_{k\to\infty} f(x_k) = \infty.$$

Remark. In general $x_k \in I\!R^n$, namely

$$x_k = \begin{bmatrix} x_k^1 \\ x_k^2 \\ \vdots \\ x_k^n \end{bmatrix},$$

i.e. we use superscripts to denote components of a vector. $\diamond$

A function that satisfies the condition of the above proposition is said to be radially unbounded.

Proof. We only prove the necessity. Suppose all level sets of f are compact. Then, proceeding by contradiction, suppose there exist a sequence $\{x_k\}$ such that $\lim_{k\to\infty} \|x_k\| = \infty$ and a number $\gamma > 0$ such that $f(x_k) \leq \gamma < \infty$ for all k. As a result

$$\{x_k\} \subset \mathcal{L}(\gamma).$$

However, by compactness of $\mathcal{L}(\gamma)$ it is not possible that $\lim_{k\to\infty} \|x_k\| = \infty$. $\triangleleft$

Definition 4 *Let $f : I\!R^n \to I\!R$. A vector $d \in I\!R^n$ is said to be a descent direction for f in $x_\star$ if there exists $\delta > 0$ such that*

$$f(x_\star + \lambda d) < f(x_\star),$$

for all $\lambda \in (0, \delta)$.

If the function f is differentiable it is possible to give a simple condition guaranteeing that a certain direction is a descent direction.

Proposition 4 *Let $f : I\!R^n \to I\!R$ and assume[3] ∇f exists and is continuous. Let $x_\star$ and d be given. Then, if $\nabla f(x_\star)' d < 0$ the direction d is a descent direction for f at $x_\star$.*

Proof. Note that $\nabla f(x_\star)' d$ is the directional derivative of f (which is differentiable by hypothesis) at $x_\star$ along d, *i.e.*

$$\nabla f(x_\star)' d = \lim_{\lambda\to 0^+} \frac{f(x_\star + \lambda d) - f(x_\star)}{\lambda},$$

[3] We denote with ∇f the gradient of the function f, *i.e.* $\nabla f = [\frac{\partial f}{\partial x^1}, \cdots, \frac{\partial f}{\partial x^n}]'$. Note that ∇f is a column vector.

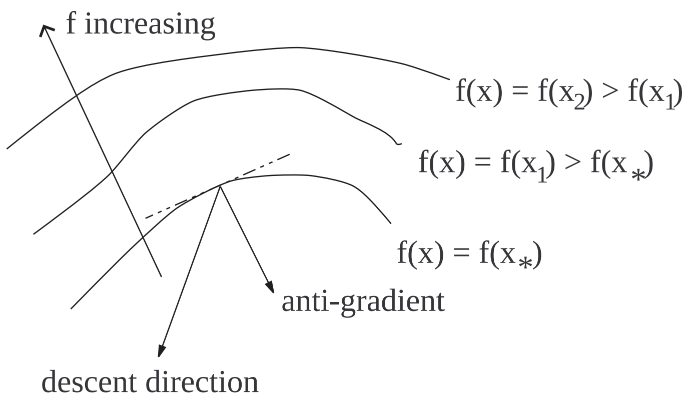


Figure 2.1: Geometrical interpretation of the anti-gradient.

and this is negative by hypothesis. As a result, for $\lambda > 0$ and sufficiently small

$$f(x_\star + \lambda d) - f(x_\star) < 0,$$

hence the claim. $\triangleleft$

The proposition establishes that if $\nabla f(x_\star)'d < 0$ then for sufficiently small positive displacements along d and starting at $x_\star$ the function f is decreasing. It is also obvious that if $\nabla f(x_\star)'d > 0$, d is a direction of *ascent*, *i.e.* the function f is increasing for sufficiently small positive displacements from $x_\star$ along d. If $\nabla f(x_\star)'d = 0$, d is orthogonal to $\nabla f(x_\star)$ and it is not possible to establish, without further knowledge on the function f, what is the nature of the direction d.

From a geometrical point of view (see also Figure 2.1), the sign of the *directional derivative* $\nabla f(x_\star)'d$ gives information on the angle between d and the direction of the gradient at $x_\star$, provided $\nabla f(x_\star) \neq 0$. If $\nabla f(x_\star)'d > 0$ the angle between $\nabla f(x_\star)$ and d is acute. If $\nabla f(x_\star)'d < 0$ the angle between $\nabla f(x_\star)$ and d is obtuse. Finally, if $\nabla f(x_\star)'d = 0$, and $\nabla f(x_\star) \neq 0$, $\nabla f(x_\star)$ and d are orthogonal. Note that the gradient $\nabla f(x_\star)$, if it is not identically zero, is a direction orthogonal to the level surface $\{x : f(x) = f(x_\star)\}$ and it is a direction of ascent, hence the anti-gradient $-\nabla f(x_\star)$ is a descent direction.

Remark. The scalar product $x'y$ between the two vectors x and y can be used to define the angle between x and y. For, define the angle between x and y as the number $\theta \in [0, \pi]$ such that[4]

$$\cos\theta = \frac{x'y}{\|x\|_E \|y\|_E}.$$

If $x'y = 0$ one has $\cos\theta = 0$ and the vectors are orthogonal, whereas if x and y have the same direction, *i.e.* $x = \lambda y$ with $\lambda > 0$, $\cos\theta = 1$. $\diamond$

[4] $\|x\|_E$ denotes the Euclidean norm of the the vector x, *i.e.* $\|x\|_E = \sqrt{x'x}$.

We are now ready to state and prove some necessary conditions and some sufficient conditions for a local minimizer.

Theorem 1 *[First order necessary condition] Let $f : I\!R^n \to I\!R$ and assume ∇f exists and is continuous. The point $x_\star$ is a local minimizer of f only if*

$$\nabla f(x_\star) = 0.$$

Remark. A point $x_\star$ such that $\nabla f(x_\star) = 0$ is called a stationary point of f. $\diamond$

Proof. If $\nabla f(x_\star) \neq 0$ the direction $d = -\nabla f(x_\star)$ is a descent direction. Therefore, in a neighborhood of $x_\star$ there is a point $x_\star + \lambda d = x_\star - \lambda \nabla f(x_\star)$ such that

$$f(x_\star - \lambda \nabla f(x_\star)) < f(x_\star),$$

and this contradicts the hypothesis that $x_\star$ is a local minimizer. $\triangleleft$

Theorem 2 *[Second order necessary condition] Let $f : I\!R^n \to I\!R$ and assume[5] $\nabla^2 f$ exists and is continuous. The point $x_\star$ is a local minimizer of f only if*

$$\nabla f(x_\star) = 0$$

and

$$x' \nabla^2 f(x_\star) x \geq 0$$

for all $x \in I\!R^n$.

Proof. The first condition is a consequence of Theorem 1. Note now that, as f is two times differentiable, for any $x \neq x_\star$ one has

$$f(x_\star + \lambda x) = f(x_\star) + \lambda \nabla f(x_\star)' x + \frac{1}{2} \lambda^2 x' \nabla^2 f(x_\star) x + \beta(x_\star, \lambda x),$$

where

$$\lim_{\lambda \to 0} \frac{\beta(x_\star, \lambda x)}{\lambda^2 \|x\|^2} = 0,$$

or what is the same (note that x is fixed)

$$\lim_{\lambda \to 0} \frac{\beta(x_\star, \lambda x)}{\lambda^2} = 0.$$

[5] We denote with $\nabla^2 f$ the Hessian matrix of the function f, *i.e.*

$$\begin{bmatrix} \frac{\partial^2 f}{\partial x^1 \partial x^1} & \cdots & \frac{\partial^2 f}{\partial x^1 \partial x^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x^n \partial x^1} & \cdots & \frac{\partial^2 f}{\partial x^n \partial x^n} \end{bmatrix}.$$

Note that $\nabla^2 f$ is a square matrix and that, under suitable regularity conditions, the Hessian matrix is symmetric.

Moreover, the condition $\nabla f(x_\star) = 0$ yields

$$\frac{f(x_\star + \lambda x) - f(x_\star)}{\lambda^2} = \frac{1}{2}x'\nabla^2 f(x_\star)x + \frac{\beta(x_\star, \lambda x)}{\lambda^2}. \tag{2.1}$$

However, as $x_\star$ is a local minimizer, the left hand side of equation (2.1) must be non-negative for all λ sufficiently small, hence

$$\frac{1}{2}x'\nabla^2 f(x_\star)x + \frac{\beta(x_\star, \lambda x)}{\lambda^2} \geq 0,$$

and

$$\lim_{\lambda \to 0} \left(\frac{1}{2}x'\nabla^2 f(x_\star)x + \frac{\beta(x_\star, \lambda x)}{\lambda^2} \right) = \frac{1}{2}x'\nabla^2 f(x_\star)x \geq 0,$$

which proves the second condition. $\triangleleft$

Theorem 3 (Second order sufficient condition) *Let $f : I\!R^n \to I\!R$ and assume $\nabla^2 f$ exists and is continuous. The point $x_\star$ is a strict local minimizer of f if*

$$\nabla f(x_\star) = 0$$

and

$$x'\nabla^2 f(x_\star)x > 0$$

for all non-zero $x \in I\!R^n$.

Proof. To begin with, note that as $\nabla^2 f(x_\star) > 0$ and $\nabla^2 f$ is continuous, then there is a neighborhood Ω of $x_\star$ such that for all $y \in \Omega$

$$\nabla^2 f(y) > 0.$$

Consider now the Taylor series expansion of f around the point $x_\star$, *i.e.*

$$f(y) = f(x_\star) + \nabla f(x_\star)'(y - x_\star) + \frac{1}{2}(y - x_\star)'\nabla^2 f(\xi)(y - x_\star),$$

where $\xi = x_\star + \theta(y - x_\star)$, for some $\theta \in [0, 1]$. By the first condition one has

$$f(y) = f(x_\star) + \frac{1}{2}(y - x_\star)'\nabla^2 f(\xi)(y - x_\star),$$

and, for any $y \in \Omega$ such that $y \neq x_\star$,

$$f(y) > f(x_\star),$$

which proves the claim. $\triangleleft$

The above results can be easily modified to derive necessary conditions and sufficient conditions for a local maximizer. Moreover, if $x_\star$ is a stationary point and the Hessian matrix

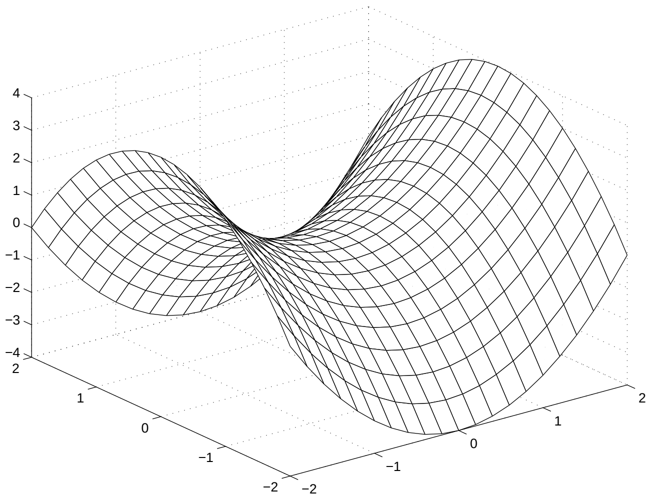


Figure 2.2: A saddle point in $I\!R^2$.

$\nabla^2 f(x_\star)$ is indefinite, the point $x_\star$ is neither a local minimizer neither a local maximizer. Such a point is called a saddle point (see Figure 2.2 for a geometrical illustration).
If $x_\star$ is a stationary point and $\nabla^2 f(x_\star)$ is semi-definite it is not possible to draw any conclusion on the point $x_\star$ without further knowledge on the function f. Nevertheless, if $n = 1$ and the function f is infinitely times differentiable it is possible to establish the following necessary and sufficient condition.

Proposition 5 *Let $f : I\!R \to I\!R$ and assume f is infinitely times differentiable. The point $x_\star$ is a local minimizer if and only if there exists an even integer $r > 1$ such that*

$$\frac{d^k f(x_\star)}{dx^k} = 0$$

for $k = 1, 2, \ldots, r-1$ and

$$\frac{d^r f(x_\star)}{dx^r} > 0.$$

Necessary and sufficient conditions for $n > 1$ can be only derived if further hypotheses on the function f are added, as shown for example in the following fact.

Proposition 6 (Necessary and sufficient condition for convex functions) *Let $f : I\!R^n \to I\!R$ and assume ∇f exists and it is continuous. Suppose f is convex,* i.e.

$$f(y) - f(x) \geq \nabla f(x)'(y - x) \tag{2.2}$$

for all $x \in I\!R^n$ and $y \in I\!R^n$. The point $x_\star$ is a global minimizer if and only if $\nabla f(x_\star) = 0$.

Remark. The convexity condition (2.2) can be rewritten as

$$f(y) \geq f(x) + \nabla f(x)'(y - x).$$

This reveals that the tangent plane for f at x is always below the graph of the function, *i.e.* the function is supported from below by the tangent planes. $\diamond$

Proof. The necessity is a consequence of Theorem 1. For the sufficiency note that, by equation (2.2), if $\nabla f(x_\star) = 0$ then

$$f(y) \geq f(x_\star),$$

for all $y \in I\!R^n$. $\triangleleft$

From the above discussion it is clear that to establish the property that $x_\star$, satisfying $\nabla f(x_\star) = 0$, is a global minimizer it is enough to assume that the function f has the following property: for all x and y such that

$$\nabla f(x)'(y - x) \geq 0$$

one has

$$f(y) \geq f(x).$$

A function f satisfying the above property is said pseudo-convex. Note that a differentiable convex function is also pseudo-convex, but the opposite is not true. For example, the function $x + x^3$ is pseudo-convex but it is not convex. Finally, if f is strictly convex or strictly pseudo-convex the global minimizer (if it exists) is also unique.

2.3 General properties of minimization algorithms

Consider the problem of minimizing the function $f : I\!R^n \to I\!R$ and suppose that ∇f and $\nabla^2 f$ exist and are continuous. Suppose that such a problem has a solution, and moreover that there exists x_0 such that the level set

$$\mathcal{L}(f(x_0)) = \{x \in I\!R^n \ : \ f(x) \leq f(x_0)\}$$

is compact.

General unconstrained minimization algorithms allow only to determine stationary points of f, *i.e.* to determine points in the set

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

Moreover, for almost all algorithms, it is possible to exclude that the points of Ω yielded by the algorithm are local maximizer. Finally, some algorithms yield points of Ω that satisfy also the second order necessary conditions.

2.3.1 General unconstrained minimization algorithm

An *algorithm* for the solution of the considered minimization problem is a sequence $\{x_k\}$, obtained starting from an initial point x_0, having some convergence properties in relation with the set Ω. Most of the algorithms that will be studied in this notes can be described in the following general way.

a) Fix a point $x_0 \in I\!\!R^n$ and set $k = 0$.

b) If $x_k \in \Omega$ STOP.

c) Compute a direction of research $d_k \in I\!\!R^n$.

d) Compute a step $\alpha_k \in I\!\!R$ along d_k.

e) Let $x_{k+1} = x_k + \alpha_k d_k$. Set $k = k + 1$ and go back to 2.

The existing algorithms differ in the way the direction of research d_k is computed and on the criteria used to compute the step α_k. However, independently from the particular selection, it is important to study the following issues:

- the existence of accumulation points for the sequence $\{x_k\}$;
- the behavior of such accumulation points in relation with the set Ω;
- the speed of convergence of the sequence $\{x_k\}$ to the points of Ω.

2.3.2 Existence of accumulation points

To make sure that any subsequence of $\{x_k\}$ has an accumulation point it is necessary to assume that the sequence $\{x_k\}$ remains bounded, *i.e.* that there exists $M > 0$ such that $\|x_k\| < M$ for any k. If the level set $\mathcal{L}(f(x_0))$ is compact, the above condition holds if $\{x_k\} \in \mathcal{L}(f(x_0))$. This property, in turn, is guaranteed if

$$f(x_{k+1}) < f(x_k),$$

for any k such that $x_k \notin \Omega$. The algorithms that satisfy this property are denominated descent methods. For such methods , if $\mathcal{L}(f(x_0))$ is compact and if ∇f is continuous one has

- $\{x_k\} \in \mathcal{L}(f(x_0))$ and any subsequence of $\{x_k\}$ admits a subsequence converging to a point of $\mathcal{L}(f(x_0))$;
- the sequence $\{f(x_k)\}$ has a limit, *i.e.* there exists $\bar{f} \in I\!\!R$ such that
$$\lim_{k\to\infty} f(x_k) = \bar{f};$$
- there always exists an element of Ω in $\mathcal{L}(f(x_0))$. In fact, as f has a minimizer in $\mathcal{L}(f(x_0))$, this minimizer is also a minimizer of f in $I\!\!R^n$. Hence, by the assumptions of ∇f, such a minimizer must be a point of Ω.

Remark. To guarantee the descent property it is necessary that the research directions d_k be directions of descent. This is true if

$$\nabla f(x_k)' d_k < 0,$$

for all k. Under this condition there exists an interval $(0, \alpha_\star]$ such that

$$f(x_k + \alpha d_k) < f(x_k),$$

for any $\alpha \in (0, \alpha_\star]$. $\diamond$

Remark. The existence of accumulation points for the sequence $\{x_k\}$ and the convergence of the sequence $\{f(x_k)\}$ do not guarantee that the accumulation points of $\{x_k\}$ are local minimizers of f or stationary points. To obtain this property it is necessary to impose further restrictions on the research directions d_k and on the steps α_k. $\diamond$

2.3.3 Condition of angle

The condition which is in general imposed on the research directions d_k is the so-called condition of angle, that can be stated as follows.

Condition 1 *There exists $\epsilon > 0$, independent from k, such that*

$$\nabla f(x_k)' d_k \leq -\epsilon \|\nabla f(x_k)\| \|d_k\|,$$

for any k.

From a geometric point of view the above condition implies that the cosine of the angle between d_k and $-\nabla f(x_k)$ is larger than a certain quantity. This condition is imposed to avoid that, for some k, the research direction is orthogonal to the direction of the gradient. Note moreover that, if the angle condition holds, and if $\nabla f(x_k) \neq 0$ then d_k is a descent direction. Finally, if $\nabla f(x_k) \neq 0$, it is always possible to find a direction d_k such that the angle condition holds. For example, the direction $d_k = -\nabla f(x_k)$ is such that the angle condition is satisfied with $\epsilon = 1$.

Remark. Let $\{B_k\}$ be a sequence of matrices such that

$$mI \leq B_k \leq MI,$$

for some $0 < m < M$, and for any k, and consider the directions

$$d_k = -B_k \nabla f(x_k).$$

Then a simple computation shows that the angle condition holds with $\epsilon = m/M$. $\diamond$

The angle condition imposes a constraint only on the research directions d_k. To make sure that the sequence $\{x_k\}$ converges to a point in Ω it is necessary to impose further conditions on the step α_k, as expressed in the following statements.

Theorem 4 *Let $\{x_k\}$ be the sequence obtained by the algorithm*

$$x_{k+1} = x_k + \alpha_k d_k,$$

for $k \geq 0$. Assume that

(H1) ∇f is continuous and the level set $\mathcal{L}(f(x_0))$ is compact.

(H2) There exists $\epsilon > 0$ such that

$$\nabla f(x_k)' d_k \leq -\epsilon \|\nabla f(x_k)\| \|d_k\|,$$

for any $k \geq 0$.

(H3) $f(x_{k+1}) < f(x_k)$ for any $k \geq 0$.

(H4) The property

$$\lim_{k\to\infty} \frac{\nabla f(x_k)' d_k}{\|d_k\|} = 0$$

holds.

Then

(C1) $\{x_k\} \in \mathcal{L}(f(x_0))$ and any subsequence of $\{x_k\}$ has an accumulation point.

(C2) $\{f(x_k)\}$ is monotonically decreasing and there exists $\bar{f}$ such that

$$\lim_{k\to\infty} f(x_k) = \bar{f}.$$

(C3) $\{\nabla f(x_k)\}$ is such that

$$\lim_{k\to\infty} \|\nabla f(x_k)\| = 0.$$

(C4) Any accumulation point $\bar{x}$ of $\{x_k\}$ is such that $\nabla f(\bar{x}) = 0$.

Proof. Conditions (C1) and (C2) are a simple consequence of (H1) and (H3). Note now that (H2) implies

$$\epsilon \|\nabla f(x_k)\| \leq \frac{|\nabla f(x_k)' d_k|}{\|d_k\|},$$

for all k. As a result, and by (H4),

$$\lim_{k\to\infty} \epsilon \|\nabla f(x_k)\| \leq \lim_{k\to\infty} \frac{|\nabla f(x_k)' d_k|}{\|d_k\|} = 0$$

hence (C3) holds. Finally, let $\bar{x}$ be an accumulation point of the sequence $\{x_k\}$, *i.e.* there is a subsequence that converges to $\bar{x}$. For such a subsequence, and by continuity of f, one has

$$\lim_{k\to\infty} \nabla f(x_k) = \nabla f(\bar{x}),$$

and, by (C3),

$$\nabla f(\bar{x}) = 0,$$

which proves (C4). ◁

Remark. Theorem 4 does not guarantee the convergence of the sequence $\{x_k\}$ to a unique accumulation point. Obviously $\{x_k\}$ has a unique accumulation point if either $\Omega \cap \mathcal{L}(f(x_0))$ contains only one point or $x, y \in \Omega \cap \mathcal{L}(f(x_0))$, with $x \neq y$ implies $f(x) \neq f(y)$. Finally, if the set $\Omega \cap \mathcal{L}(f(x_0))$ contains a finite number of points, a sufficient condition for the existence of a unique accumulation point is

$$\lim_{k \to \infty} \|x_{k+1} - x_k\| = 0.$$

◇

Remark. The angle condition can be replaced by the following one. There exists $\eta > 0$ and $q > 0$, both independent from k, such that

$$\nabla f(x_k)' d_k \leq -\eta \|\nabla f(x_k)\|^q \|d_k\|.$$

◇

The result illustrated in Theorem 4 requires the fulfillment of the angle condition or of a similar one, *i.e.* of a condition involving ∇f. In many algorithms that do not make use of the gradient it may be difficult to check the validity of the angle condition, hence it is necessary to use different conditions on the research directions. For example, it is possible to replace the angle condition with a property of linear independence of the research directions.

Theorem 5 *Let $\{x_k\}$ be the sequence obtained by the algorithm*

$$x_{k+1} = x_k + \alpha_k d_k,$$

for $k \geq 0$. Assume that

- *$\nabla^2 f$ is continuous and the level set $\mathcal{L}(f(x_0))$ is compact.*
- *There exist $\sigma > 0$, independent from k, and $k_0 > 0$ such that, for any $k \geq k_0$ the matrix P_k composed of the columns*

$$\frac{d_k}{\|d_k\|}, \frac{d_{k+1}}{\|d_{k+1}\|}, \ldots, \frac{d_{k+n-1}}{\|d_{k+n-1}\|},$$

  is such that

$$|\det P_k| \geq \sigma.$$

- $\lim_{k \to \infty} \|x_{k+1} - x_k\| = 0.$

- *$f(x_{k+1}) < f(x_k)$ for any $k \geq 0$.*
- *The property*

$$\lim_{k\to\infty} \frac{\nabla f(x_k)' d_k}{\|d_k\|} = 0$$

  holds.

Then

- *$\{x_k\} \in \mathcal{L}(f(x_0))$ and any subsequence of $\{x_k\}$ has an accumulation point.*
- *$\{f(x_k)\}$ is monotonically decreasing and there exists $\bar{f}$ such that*

$$\lim_{k\to\infty} f(x_k) = \bar{f}.$$

- *Any accumulation point $\bar{x}$ of $\{x_k\}$ is such that $\nabla f(\bar{x}) = 0$.*

Moreover, if the set $\Omega \cap \mathcal{L}(f(x_0))$ is composed of a finite number of points, the sequence $\{x_k\}$ has a unique accumulation point.

2.3.4 Speed of convergence

Together with the property of convergence of the sequence $\{x_k\}$ it is important to study also the speed of convergence. To study such a notion it is convenient to assume that $\{x_k\}$ converges to a point $x_\star$.
If there exists a finite k such that $x_k = x_\star$ then we say that the sequence $\{x_k\}$ has finite convergence. Note that if $\{x_k\}$ is generated by an algorithm, there is a stopping condition that has to be satisfied at step k.
If $x_k \neq x_\star$ for any finite k, it is possible (and convenient) to study the asymptotic properties of $\{x_k\}$. One criterion to estimate the speed of convergence is based on the behavior of the error $\mathcal{E}_k = \|x_k - x_\star\|$, and in particular on the relation between $\mathcal{E}_{k+1}$ and $\mathcal{E}_k$.
We say that $\{x_k\}$ has speed of convergence of order p if

$$\lim_{k\to\infty} \left(\frac{\mathcal{E}_{k+1}}{\mathcal{E}_k^p}\right) = C_p$$

with $p \geq 1$ and $0 < C_p < \infty$. Note that if $\{x_k\}$ has speed of convergence of order p then

$$\lim_{k\to\infty} \left(\frac{\mathcal{E}_{k+1}}{\mathcal{E}_k^q}\right) = 0,$$

if $1 \leq q < p$, and

$$\lim_{k\to\infty} \left(\frac{\mathcal{E}_{k+1}}{\mathcal{E}_k^q}\right) = \infty,$$

if $q > p$. Moreover, from the definition of speed of convergence, it is easy to see that if $\{x_k\}$ has speed of convergence of order p then, for any $\epsilon > 0$ there exists k_0 such that

$$\mathcal{E}_{k+1} \leq (C_p + \epsilon)\mathcal{E}_k^p,$$

for any $k > k_0$.
In the cases $p = 1$ or $p = 2$ the following terminology is often used. If $p = 1$ and $0 < C_1 \leq 1$ the speed of convergence is linear; if $p = 1$ and $C_1 > 1$ the speed of convergence is sublinear; if

$$\lim_{k\to\infty} \left(\frac{\mathcal{E}_{k+1}}{\mathcal{E}_k} \right) = 0$$

the speed of convergence is superlinear, and finally if $p = 2$ the speed of convergence is quadratic.
Of special interest in optimization is the case of superlinear convergence, as this is the kind of convergence that can be established for the *efficient* minimization algorithms. Note that if x_k has superlinear convergence to $x_\star$ then

$$\lim_{k\to\infty} \frac{\|x_{k+1} - x_k\|}{\|x_k - x_\star\|} = 1.$$

Remark. In some cases it is not possible to establish the existence of the limit

$$\lim_{k\to\infty} \left(\frac{\mathcal{E}_{k+1}}{\mathcal{E}_k^q} \right).$$

In these cases an estimate of the speed of convergence is given by

$$Q_p = \limsup_{k\to\infty} \left(\frac{\mathcal{E}_{k+1}}{\mathcal{E}_k^q} \right).$$

◇

2.4 Line search

A line search is a method to compute the step α_k along a given direction d_k. The choice of α_k affects both the convergence and the speed of convergence of the algorithm. In any line search one considers the function of one variable $\phi : I\!R \to I\!R$ defined as

$$\phi(\alpha) = f(x_k + \alpha d_k) - f(x_k).$$

The derivative of $\phi(\alpha)$ with respect to α is given by

$$\dot{\phi}(\alpha) = \nabla f(x_k + \alpha d_k)' d_k$$

provided that ∇f is continuous. Note that $\nabla f(x_k + \alpha d_k)' d_k$ describes the slope of the tangent to the function $\phi(\alpha)$, and in particular

$$\dot{\phi}(0) = \nabla f(x_k)' d_k$$

coincides with the directional derivative of f at x_k along d_k.
From the general convergence results described, we conclude that the line search has to enforce the following conditions

$$f(x_{k+1}) < f(x_k)$$

$$\lim_{k\to\infty} \frac{\nabla f(x_k)' d_k}{\|d_k\|} = 0$$

and, whenever possible, also the condition

$$\lim_{k\to\infty} \|x_{k+1} - x_k\| = 0.$$

To begin with, we assume that the directions d_k are such that

$$\nabla f(x_k)' d_k < 0$$

for all k, *i.e.* d_k is a descent direction, and that it is possible to compute, for any fixed x, both f and ∇f. Finally, we assume that the level set $\mathcal{L}(f(x_0))$ is compact.

2.4.1 Exact line search

The exact line search consists in finding α_k such that

$$\phi(\alpha_k) = f(x_k + \alpha_k d_k) - f(x_k) \le f(x_k + \alpha d_k) - f(x_k) = \phi(\alpha)$$

for any $\alpha \ge 0$. Note that, as d_k is a descent direction and the set

$$\{\alpha \in I\!R^+ \ : \ \phi(\alpha) \le \phi(0)\}$$

is compact, because of compactness of $\mathcal{L}(f(x_0))$, there exists an α_k that minimizes $\phi(\alpha)$. Moreover, for such α_k one has

$$\dot{\phi}(\alpha_k) = \nabla f(x_k + \alpha_k d_k)' d_k = 0,$$

i.e. if α_k minimizes $\phi(\alpha)$ the gradient of f at $x_k + \alpha_k d_k$ is orthogonal to the direction d_k. From a geometrical point of view, if α_k minimizes $\phi(\alpha)$ then the level surface of f through the point $x_k + \alpha_k d_k$ is tangent to the direction d_k at such a point. (If there are several points of tangency, α_k is the one for which f has the smallest value).
The search of α_k that minimizes $\phi(\alpha)$ is very *expensive*, especially if f is not convex. Moreover, in general, the whole minimization algorithm does not gain any special advantage from the knowledge of such *optimal* α_k. It is therefore more convenient to use approximate methods, *i.e.* methods which are computationally simple and which guarantee particular convergence properties. Such methods are aimed at finding an interval of acceptable values for α_k subject to the following two conditions

- α_k has to guarantee a sufficient reduction of f;
- α_k has to be sufficiently distant from 0, *i.e.* $x_k + \alpha_k d_k$ has to be sufficiently away from x_k.

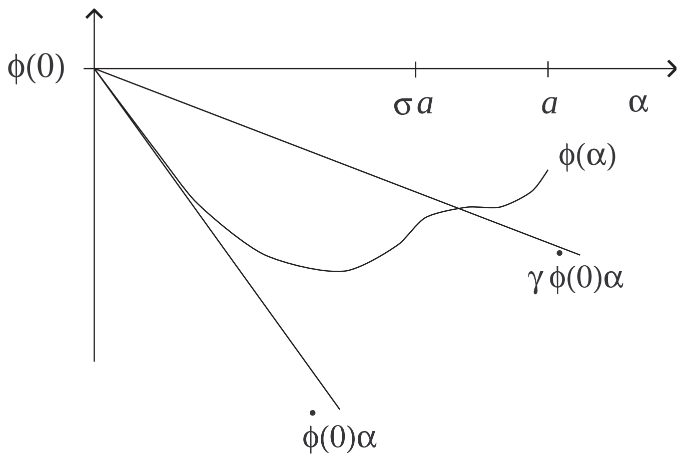


Figure 2.3: Geometrical interpretation of Armijo method.

2.4.2 Armijo method

Armijo method was the first non-exact linear search method.
Let $a > 0$, $\sigma \in (0, 1)$ and $\gamma \in (0, 1/2)$ be given and define the set of points

$$A = \{\alpha \in R \;:\; \alpha = a\sigma^j,\ j = 0, 1, \ldots\}.$$

Armijo method consists in finding the largest $\alpha \in A$ such that

$$\phi(\alpha) = f(x_k + \alpha d_k) - f(x_k) \le \gamma\alpha\nabla f(x_k)'d_k = \gamma\alpha\dot{\phi}(0).$$

Armijo method can be implemented using the following (conceptual) algorithm.

Step 1. Set $\alpha = a$.

Step 2. If

$$f(x_k + \alpha d_k) - f(x_k) \le \gamma\alpha\nabla f(x_k)'d_k$$

set $\alpha_k = \alpha$ and STOP. Else go to **Step 3.**

Step 3. Set $\alpha = \sigma\alpha$, and go to **Step 2.**

From a geometric point of view (see Figure 2.3) the condition in **Step 2** requires that α_k is such that $\phi(\alpha_k)$ is below the straight line passing through the point $(0, \phi(0))$ and with slope $\gamma\dot{\phi}(0)$. Note that, as $\gamma \in (0, 1/2)$ and $\dot{\phi}(0) < 0$, such a straight line has a slope smaller than the slope of the tangent at the curve $\phi(\alpha)$ at the point $(0, \phi(0))$.
For Armijo method it is possible to prove the following convergence result.

Theorem 6 *Let $f : I\!R^n \to I\!R$ and assume ∇f is continuous and $\mathcal{L}(f(x_0))$ is compact. Assume $\nabla f(x_k)'d_k < 0$ for all k and there exist $C_1 > 0$ and $C_2 > 0$ such that*

$$C_1 \ge \|d_k\| \ge C_2\|\nabla f(x_k)\|^q,$$

for some $q > 0$ and for all k.
Then Armijo method yields in a finite number of iterations a value of $\alpha_k > 0$ satisfying the condition in **Step 2***. Moreover, the sequence obtained setting $x_{k+1} = x_k + \alpha_k d_k$ is such that*

$$f(x_{k+1}) < f(x_k),$$

for all k, and

$$\lim_{k\to\infty} \frac{\nabla f(x_k)' d_k}{\|d_k\|} = 0.$$

Proof. We only prove that the method cannot loop indefinitely between **Step 2** and **Step 3**. In fact, if this is the case, then the condition in **Step 2** will never be satisfied, hence

$$\frac{f(x_k + a\sigma^j d_k) - f(x_k)}{a\sigma^j} > \gamma \nabla f(x_k)' d_k.$$

Note now that $\sigma^j \to 0$ as $j \to \infty$, and the above inequality for $j \to \infty$ is

$$\nabla f(x_k)' d_k > \gamma \nabla f(x_k)' d_k,$$

which is not possible since $\gamma \in (0, 1/2)$ and $\nabla f(x_k)' d_k \neq 0$. $\triangleleft$

Remark. It is interesting to observe that in Theorem 6 it is not necessary to assume that $x_{k+1} = x_k + \alpha_k d_k$. It is enough that x_{k+1} is such that

$$f(x_{k+1}) \leq f(x_k + \alpha_k d_k),$$

where α_k is generated using Armijo method. This implies that all acceptable values of α are those such that

$$f(x_k + \alpha d_k) \leq f(x_k + \alpha_k d_k).$$

As a result, Theorem 6 can be used to prove also the convergence of an algorithm based on the exact line search. $\diamond$

2.4.3 Goldstein conditions

The main disadvantage of Armijo method is in the fact that, to find α_k, all points in the set A, starting from the point $\alpha = a$, have to be tested till the condition in **Step 2** is fulfilled. There are variations of the method that do not suffer from this disadvantage. A criterion similar to Armijo's, but that allows to find an acceptable α_k in one step, is based on the so-called Goldstein conditions.
Goldstein conditions state that given $\gamma_1 \in (0, 1)$ and $\gamma_2 \in (0, 1)$ such that $\gamma_1 < \gamma_2$, α_k is any positive number such that

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \alpha_k \gamma_1 \nabla f(x_k)' d_k$$

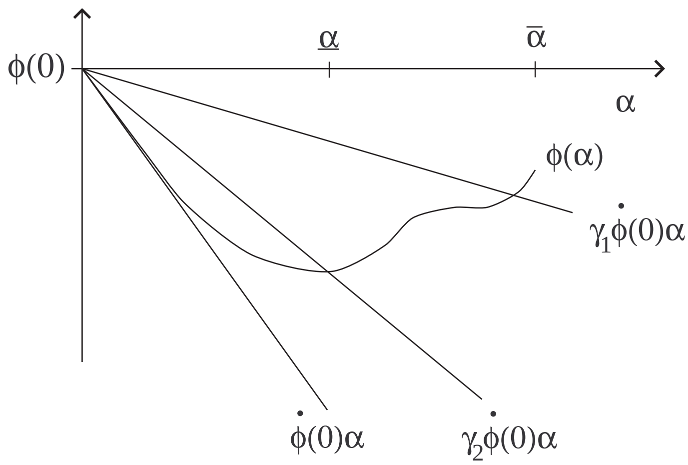


Figure 2.4: Geometrical interpretation of Goldstein method.

i.e. there is a sufficient reduction in f, and

$$f(x_k + \alpha_k d_k) - f(x_k) \geq \alpha_k \gamma_2 \nabla f(x_k)' d_k$$

i.e. there is a sufficient distance between x_k and x_{k+1}.
From a geometric point of view (see Figure 2.4) this is equivalent to select α_k as any point such that the corresponding value of f is included between two straight lines, of slope $\gamma_1 \nabla f(x_k)' d_k$ and $\gamma_2 \nabla f(x_k)' d_k$, respectively, and passing through the point $(0, \phi(0))$. As $0 < \gamma_1 < \gamma_2 < 1$ it is obvious that there exists always an interval $I = [\underline{\alpha}, \overline{\alpha}]$ such that Goldstein conditions hold for any $\alpha \in I$.
Note that, a result similar to Theorem 6, can be also established if the sequence $\{x_k\}$ is generated using Goldstein conditions.
The main disadvantage of Armijo and Goldstein methods is in the fact that none of them impose conditions on the derivative of the function $\phi(\alpha)$ in the point α_k, or what is the same on the value of $\nabla f(x_{k+1})' d_k$. Such extra conditions are sometimes useful in establishing convergence results for particular algorithms. However, for simplicity, we omit the discussion of these more general conditions (known as Wolfe conditions).

2.4.4 Line search without derivatives

It is possible to construct methods similar to Armijo's or Goldstein's also in the case that no information on the derivatives of the function f is available.
Suppose, for simplicity, that $\|d_k\| = 1$, for all k, and that the sequence $\{x_k\}$ is generated by

$$x_{k+1} = x_k + \alpha_k d_k.$$

If ∇f is not available it is not possible to decide *a priori* if the direction d_k is a descent direction, hence it is necessary to consider also negative values of α.

We now describe the simplest line search method that can be constructed with the considered hypothesis. This method is a modification of Armijo method and it is known as parabolic search.
Given $\lambda_0 > 0$, $\sigma \in (0, 1/2)$, $\gamma > 0$ and $\rho \in (0, 1)$. Compute α_k and λ_k such that one of the following conditions hold.

Condition (i)

- $\lambda_k = \lambda_{k-1}$;
- α_k is the largest value in the set
$$A = \{\alpha \in I\!R \ : \ \alpha = \pm\sigma^j, \ j = 0, 1, \ldots\}$$
such that
$$f(x_k + \alpha_k d_k) \leq f(x_k) - \gamma\alpha_k^2,$$
or, equivalently, $\phi(\alpha_k) \leq -\gamma\alpha_k^2$.

Condition (ii)

- $\alpha_k = 0$, $\lambda_k \leq \rho\lambda_{k-1}$;
- $\min\left(f(x_k + \lambda_k d_k), f(x_k - \lambda_k d_k)\right) \geq f(x_k) - \gamma\lambda_k^2$.

At each step it is necessary to satisfy either Condition (i) or Condition (ii). Note that this is always possible for any $d_k \neq 0$. Condition (i) requires that α_k is the largest number in the set A such that $f(x_k + \alpha_k d_k)$ is below the parabola $f(x_k) - \gamma\alpha^2$. If the function $\phi(\alpha)$ has a stationary point for $\alpha = 0$ then there may be no $\alpha \in A$ such that Condition (i) holds. However, in this case it is possible to find λ_k such that Condition (ii) holds. If Condition (ii) holds then $\alpha_k = 0$, *i.e.* the point x_k remains unchanged and the algorithms continues with a new direction $d_{k+1} \neq d_k$.
For the parabolic search algorithm it is possible to prove the following convergence result.

Theorem 7 *Let $f : I\!R^n \to I\!R$ and assume ∇f is continuous and $\mathcal{L}(f(x_0))$ is compact. If α_k is selected following the conditions of the parabolic search and if $x_{k+1} = x_k + \alpha_k d_k$, with $\|d_k\| = 1$ then the sequence $\{x_k\}$ is such that*
$$f(x_{k+1}) \leq f(x_k)$$
for all k,
$$\lim_{k\to\infty} \nabla f(x_k)' d_k = 0$$
and
$$\lim_{k\to\infty} \|x_{k+1} - x_k\| = 0.$$

Proof. (Sketch) Note that Condition (i) implies $f(x_{k+1}) < f(x_k)$, whereas Condition (ii) implies $f(x_{k+1}) = f(x_k)$. Note now that if Condition (ii) holds for all $k \geq \bar{k}$, then $\alpha_k = 0$ for all $k \geq \bar{k}$, *i.e.* $\|x_{k+1} - x_k\| = 0$. Moreover, as λ_k is reduced at each step, necessarily $\nabla f(x_{\bar{k}})'\bar{d} = 0$, where $\bar{d}$ is a limit of the sequence $\{d_k\}$. $\triangleleft$

2.4.5 Implementation of a line search algorithm

On the basis of the conditions described so far it is possible to construct algorithms that yield α_k in a finite number of steps. One such an algorithm can be described as follows. (For simplicity we assume that ∇f is known.)

- Initial data. x_k, $f(x_k)$, $\nabla f(x_k)$, $\underline{\alpha}$ and $\overline{\alpha}$.
- Initial guess for α. A possibility is to select α as the point in which a parabola through $(0, \phi(0))$ with derivative $\dot{\phi}(0)$ for $\alpha = 0$ takes a pre-specified minimum value $f_\star$. Initially, *i.e.* for $k = 0$, $f_\star$ has to be selected by the designer. For $k > 0$ it is possible to select $f_\star$ such that

  $$f(x_k) - f_\star = f(x_{k-1}) - f(x_k).$$

  The resulting α is

  $$\alpha_\star = -2\,\frac{f(x_k) - f_\star}{\nabla f(x_k)' d_k}.$$

  In some algorithms it is convenient to select $\alpha \leq 1$, hence the initial guess for α will be $\min(1, \alpha_\star)$.
- Computation of α_k. A value for α_k is computed using a line search method. If $\alpha_k \leq \underline{\alpha}$ the direction d_k may not be a descent direction. If $\alpha_k \geq \overline{\alpha}$ the level set $\mathcal{L}(f(x_k))$ may not be compact. If $\alpha_k \notin [\underline{\alpha}, \overline{\alpha}]$ the line search fails, and it is necessary to select a new research direction d_k. Otherwise the line search terminates and $x_{k+1} = x_k + \alpha_k d_k$.

2.5 The gradient method

The gradient method consists in selecting, as research direction, the direction of the anti-gradient at x_k, *i.e.*

$$d_k = -\nabla f(x_k),$$

for all k. This selection is justified noting that the direction[6]

$$-\frac{\nabla f(x_k)}{\|\nabla f(x_k)\|_E}$$

is the direction that minimizes the directional derivative, among all direction with unitary Euclidean norm. In fact, by Schwartz inequality, one has

$$|\nabla f(x_k)' d| \leq \|d\|_E \|\nabla f(x_k)\|_E,$$

and the equality sign holds if and only if $d = \lambda \nabla f(x_k)$, with $\lambda \in I\!R$. As a consequence, the problem

$$\min_{\|d\|_E = 1} \nabla f(x_k)' d$$

[6]We denote with $\|v\|_E$ the Euclidean norm of the vector v, *i.e.* $\|v\|_E = \sqrt{v'v}$.

has the solution $d_\star = -\frac{\nabla f(x_k)}{\|\nabla f(x_k)\|_E}$. For this reason, the gradient method is sometimes called the method of the steepest descent. Note however that the (local) optimality of the direction $-\nabla f(x_k)$ depends upon the selection of the norm, and that with a proper selection of the norm, any descent direction can be regarded as the steepest descent.
The real interest in the direction $-\nabla f(x_k)$ rests on the fact that, if ∇f is continuous, then the former is a continuous descent direction, which is zero only if the gradient is zero, *i.e.* at a stationary point.
The gradient algorithm can be schematized has follows.

Step 0. Given $x_0 \in I\!R^n$.

Step 1. Set $k = 0$.

Step 2. Compute $\nabla f(x_k)$. If $\nabla f(x_k) = 0$ STOP. Else set $d_k = -\nabla f(x_k)$.

Step 3. Compute a step α_k along the direction d_k with any line search method such that

$$f(x_k + \alpha_k d_k) \leq f(x_k)$$

and

$$\lim_{k\to\infty} \frac{\nabla f(x_k)' d_k}{\|d_k\|} = 0.$$

Step 4. Set $x_{k+1} = x_k + \alpha_k d_k$, $k = k + 1$. Go to **Step 2**.

By the general results established in Theorem 4, we have the following fact regarding the convergence properties of the gradient method.

Theorem 8 *Consider $f : I\!R^n \to I\!R$. Assume ∇f is continuous and the level set $\mathcal{L}(f(x_0))$ is compact. Then any accumulation point of the sequence $\{x_k\}$ generated by the gradient algorithm is a stationary point of f.*

To estimate the speed of convergence of the method we can consider the behavior of the method in the minimization of a quadratic function, *i.e.* in the case

$$f(x) = \frac{1}{2} x' Q x + c' x + d,$$

with $Q = Q' > 0$. In such a case it is possible to obtain the following estimate

$$\|x_{k+1} - x_\star\| \leq \sqrt{\frac{\lambda_M}{\lambda_m}} \frac{\sqrt{\frac{\lambda_M}{\lambda_m}} - 1}{\sqrt{\frac{\lambda_M}{\lambda_m}} + 1} \|x_k - x_\star\|,$$

where $\lambda_M \geq \lambda_m > 0$ are the maximum and minimum eigenvalue of Q, respectively. Note that the above estimate is exact for some initial points x_0. As a result, if $\lambda_M \neq \lambda_m$ the gradient algorithm has linear convergence, however, if λ_M/λ_m is large the convergence can be very slow (see Figure 2.5).

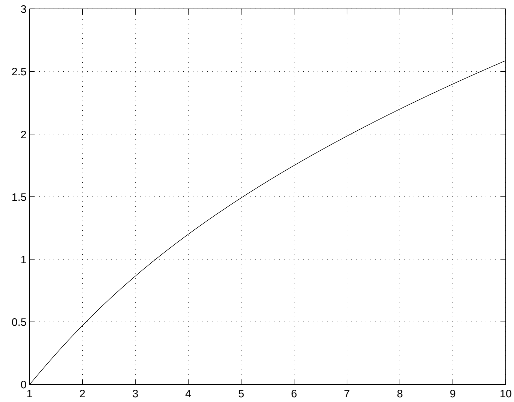


Figure 2.5: The function $\sqrt{\xi}\,\dfrac{\xi-1}{\xi+1}$.

Finally, if $\lambda_M/\lambda_m = 1$ the gradient algorithm converges in one step. From a geometric point of view the ratio λ_M/λ_m expresses the ratio between the lengths of the maximum and the minimum axes of the ellipsoids, that constitute the level surfaces of f. If this ratio is big there are points from which the gradient algorithm converges very slowly, see *e.g.* Figure 2.6.

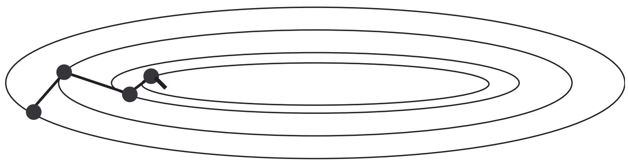


Figure 2.6: Behavior of the gradient algorithm.

In the non-quadratic case, the performance of the gradient method are unacceptable, especially if the level surfaces of f have high curvature.

2.6 Newton's method

Newton's method, with all its variations, is the most important method in unconstrained optimization. Let $f : I\!R^n \to I\!R$ be a given function and assume that $\nabla^2 f$ is continuous. Newton's method for the minimization of f can be derived assuming that, given x_k, the point x_{k+1} is obtained minimizing a quadratic approximation of f. As f is two times differentiable, it is possible to write

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

in which

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

For $\|s\|$ sufficiently small, it is possible to approximate $f(x_k + s)$ with its quadratic approximation

$$q(s) = f(x_k) + \nabla f(x_k)' s + \frac{1}{2} s' \nabla^2 f(x_k) s.$$

If $\nabla^2 f(x_k) > 0$, the value of s minimizing $q(s)$ can be obtained setting to zero the gradient of $q(s)$, *i.e.*

$$\nabla q(s) = \nabla f(x_k) + \nabla^2 f(x_k) s = 0,$$

yielding

$$s = -\left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k).$$

The point x_{k+1} is thus given by

$$x_{k+1} = x_k - \left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k).$$

Finally, Newton's method can be described by the simple scheme.

Step 0. Given $x_0 \in I\!R^n$.

Step 1. Set $k = 0$.

Step 2. Compute

$$s = -\left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k).$$

Step 3. Set $x_{k+1} = x_k + s$, $k = k + 1$. Go to **Step 2**.

Remark. An equivalent way to introduce Newton's method for unconstrained optimization is to regard the method as an algorithm for the solution of the system of n non-linear equations in n unknowns given by

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

For, consider, in general, a system of n equations in n unknown

$$F(x) = 0,$$

with $x \in I\!R^n$ and $F : I\!R^n \to I\!R^n$. If the Jacobian matrix of F exists and is continuous, then one can write

$$F(x + s) = F(x) + \frac{\partial F}{\partial x}(x) s + \gamma(x, s),$$

with

$$\lim_{\|s\|\to 0} \frac{\gamma(x, s)}{\|s\|} = 0.$$

Hence, given a point x_k we can determine $x_{k+1} = x_k + s$ setting s such that

$$F(x_k) + \frac{\partial F}{\partial x}(x_k)s = 0.$$

If $\frac{\partial F}{\partial x}(x_k)$ is invertible we have

$$s = -\left[\frac{\partial F}{\partial x}(x_k)\right]^{-1} F(x_k),$$

hence Newton's method for the solution of the system of equation $F(x) = 0$ is

$$x_{k+1} = x_k - \left[\frac{\partial F}{\partial x}(x_k)\right]^{-1} F(x_k), \tag{2.3}$$

with $k = 0, 1, \ldots$. Note that, if $F(x) = \nabla f$, then the above iteration coincides with Newton's method for the minimization of f. $\diamond$

To study the convergence properties of Newton's method we can consider the algorithm for the solution of a set of non-linear equations, summarized in equation (2.3). The following local convergence result, providing also an estimate of the speed of convergence, can be proved.

Theorem 9 *Let $F : I\!R^n \to I\!R^n$ and assume that F is continuously differentiable in an open set $\mathcal{D} \subset I\!R^n$. Suppose moreover that*

- *there exists $x_\star \in \mathcal{D}$ such that $F(x_\star) = 0$;*
- *the Jacobian matrix $\frac{\partial F}{\partial x}(x_\star)$ is non-singular;*
- *there exists $L > 0$ such that*[7]

$$\left\| \frac{\partial F}{\partial x}(z) - \frac{\partial F}{\partial x}(y) \right\| \leq L\|z - y\|,$$

for all $z \in \mathcal{D}$ and $y \in \mathcal{D}$.

Then there exists and open set $\mathcal{B} \subset \mathcal{D}$ such that for any $x_0 \in \mathcal{B}$ the sequence $\{x_k\}$ generated by equation (2.3) remains in $\mathcal{B}$ and converges to $x_\star$ with quadratic speed of convergence.

The result in Theorem 9 can be easily recast as a result for the convergence of Newton's method for unconstrained optimization. For, it is enough to note that all hypotheses on F and $\frac{\partial F}{\partial x}$ translate into hypotheses on ∇f and $\nabla^2 f$. Note however that the result is only local and does not allow to distinguish between local minimizers and local maximizers. To construct an algorithm for which the sequence $\{x_k\}$ does not converge to maxima, and for which global convergence, *i.e.* convergence from points outside the set $\mathcal{B}$, holds,

[7] This is equivalent to say that $\frac{\partial F}{\partial x}(x)$ is Lipschitz continuous in $\mathcal{D}$.

it is possible to modify Newton's method considering a line search along the direction $d_k = -\left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k)$. As a result, the modified Newton's algorithm

$$x_{k+1} = x_k - \alpha_k \left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k), \tag{2.4}$$

in which α_k is computed using any line search algorithm, is obtained. If $\nabla^2 f$ is uniformly positive definite, and this implies that the function f is convex, the direction $d_k = -\left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k)$ is a descent direction satisfying the condition of angle. Hence, by Theorem 4, we can conclude the (global) convergence of the algorithm (2.4). Moreover, it is possible to prove that, for k sufficiently large, the step $\alpha_k = 1$ satisfies the conditions of Armijo method, hence the sequence $\{x_k\}$ has quadratic speed of convergence.

Remark. If the function to be minimized is quadratic, *i.e.*

$$f(x) = \frac{1}{2}x'Qx + c'x + d,$$

and if $Q > 0$, Newton's method yields the (global) minimizer of f in one step. $\diamond$

In general, *i.e.* if $\nabla^2 f(x)$ is not positive definite for all x, Newton's method may be inapplicable because either $\nabla^2 f(x_k)$ is not invertible, or $d_k = -\left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k)$ is not a descent direction. In these cases it is necessary to further modify Newton's method. Diverse criteria have been proposed, most of which rely on the substitution of the matrix $\nabla^2 f(x_k)$ with a matrix $M_k > 0$ which is *close in some sense* to $\nabla^2 f(x_k)$. A simpler modification can be obtained using the direction $d_k = -\nabla f(x_k)$ whenever the direction $d_k = -\left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k)$ is not a descent direction. This modification yields the following algorithm.

Step 0. Given $x_0 \in I\!R^n$ and $\epsilon > 0$.

Step 1. Set $k = 0$.

Step 2. Compute $\nabla f(x_k)$. If $\nabla f(x_k) = 0$ STOP. Else compute $\nabla^2 f(x_k)$. If $\nabla^2 f(x_k)$ is singular set $d_k = -\nabla f(x_k)$ and go to **Step 6**.

Step 3. Compute Newton direction s solving the (linear) system

$$\nabla^2 f(x_k)s = -\nabla f(x_k).$$

Step 4. If

$$|\nabla f(x_k)'s| < \epsilon\|\nabla f(x_k)\|\|s\|$$

set $d_k = -\nabla f(x_k)$ and go to **Step 6**.

Step 5. If

$$\nabla f(x_k)'s < 0$$

set $d_k = s$; if

$$\nabla f(x_k)'s > 0$$

set $d_k = -s$.

Step 6. Make a line search along d_k assuming as initial estimate $\alpha = 1$. Compute $x_{k+1} = x_k + \alpha_k d_k$, set $k = k + 1$ and go to **Step 2**.

The above algorithm is such that the direction d_k satisfies the condition of angle, *i.e.*

$$\nabla f(x_k)' d_k \leq -\epsilon \|\nabla f(x_k)\| \|d_k\|,$$

for all k. Hence, the convergence is guaranteed by the general result in Theorem 4. Moreover, if ϵ is sufficiently small, if the hypotheses of Theorem 9 hold, and if the line search is performed with Armijo method and with the initial guess $\alpha = 1$, then the above algorithm has quadratic speed of convergence.
Finally, note that it is possible to modify Newton's method, whenever it is not applicable, without making use of the direction of the anti-gradient. We now briefly discuss two such modifications.

2.6.1 Method of the trust region

A possible approach to modify Newton's method to yield global convergence is to set the direction d_k and the step α_k in such a way to minimize the quadratic approximation of f on a sphere centered at x_k and of radius a_k. Such a sphere is called *trust region*. This name refers to the fact that, in a small region around x_k we are confident (we trust) that the quadratic approximation of f is a *good* approximation.
The method of the trust region consists in selecting $x_{k+1} = x_k + s_k$, where s_k is the solution of the problem

$$\min_{\|s\| \leq a_k} q(s), \tag{2.5}$$

with

$$q(s) = f(x_k) + \nabla f(x_k)' s + \frac{1}{2} s' \nabla^2 f(x_k) s,$$

and $a_k > 0$ the estimate at step k of the trust region. As the above (constrained) optimization problem has always a solution, the direction s_k is always defined. The computation of the estimate a_k is done, iteratively, in such a way to enforce the condition $f(x_{k+1}) < f(x_k)$ and to make sure that $f(x_k + s_k) \approx q(s_k)$, *i.e.* that the change of f and the estimated change of f are *close*.
Using these simple ingredients it is possible to construct the following algorithm.

Step 0. Given $x_0 \in I\!R^n$ and $a_0 > 0$.

Step 1. Set $k = 0$.

Step 2. Compute $\nabla f(x_k)$. If $\nabla f(x_k) = 0$ STOP. Else go to **Step 3**.

Step 3. Compute s_k solving problem (2.5).

Step 4. Compute[8]

$$\rho_k = \frac{f(x_k + s_k) - f(x_k)}{q(s_k) - f(x_k)}. \tag{2.6}$$

[8]If f is quadratic then $\rho_k = 1$ for all k.

Step 5. If $\rho_k < 1/4$ set $a_{k+1} = \|s_k\|/4$. If $\rho_k > 3/4$ and $\|s_k\| = a_k$ set $a_{k+1} = 2a_k$. Else set $a_{k+1} = a_k$.

Step 6. If $\rho_k \leq 0$ set $x_{k+1} = x_k$. Else set $x_{k+1} = x_k + s_k$.

Step 7. Set $k = k + 1$ and go to **Step 2**.

Remark. Equation (2.6) expresses the ratio between the actual change of f and the estimated change of f. ◇

It is possible to prove that, if $\mathcal{L}(f(x_0))$ is compact and $\nabla^2 f$ is continuous, any accumulation point resulting from the above algorithm is a stationary point of f, in which the second order necessary conditions hold.
The update of a_k is devised to enlarge or shrink the region of confidence on the basis of the number ρ_k. It is possible to show that if $\{x_k\}$ converges to a local minimizer in which $\nabla^2 f$ is positive definite, then ρ_k converges to one and the direction s_k coincides, for k sufficiently large, with the Newton direction. As a result, the method has quadratic speed of convergence.
In practice, the solution of the problem (2.5) cannot be obtained analytically, hence approximate problems have to be solved. For, consider s_k as the solution of the equation

$$\left(\nabla^2 f(x_k) + \nu_k I\right) s_k = -\nabla f(x_k), \tag{2.7}$$

in which $\nu_k > 0$ has to be determined with proper considerations. Under certain hypotheses, the s_k determined solving equation (2.7) coincides with the s_k computed using the method of the trust region.

Remark. A potential disadvantage of the method of the trust region is to reduce the step along Newton direction even if the selection $\alpha_k = 1$ would be feasible. ◇

2.6.2 Non-monotonic line search

Experimental evidence shows that Newton's method gives the best result if the step $\alpha_k = 1$ is used. Therefore, the use of $\alpha_k < 1$ along Newton direction, resulting *e.g.* from the application of Armijo method, results in a degradation of the performance of the algorithm. To avoid this phenomenon it has been suggested to relax the condition $f(x_{k+1}) < f(x_k)$ imposed on Newton algorithm, thus allowing the function f to increase for a certain number of steps. For example, it is possible to substitute the *reduction* condition of Armijo method with the condition

$$f(x_k + \alpha_k d_k) \leq \max_{0 \leq j \leq M} [f(x_{k-j})] + \gamma \alpha_k \nabla f(x_k)' d_k$$

for all $k \geq M$, where $M > 0$ is a fixed integer independent from k.

	Gradient method	Newton's method
Information required at each step	f and ∇f	f, ∇f and $\nabla^2 f$
Computation to find the research direction	$\nabla f(x_k)$	$\nabla f(x_k)$, $\nabla^2 f(x_k)$, $-[\nabla^2 f(x_k)]^{-1}\nabla f(x_k)$
Convergence	Global if $\mathcal{L}(f(x_0))$ compact and ∇f continuous	Local, but may be rendered global
Behavior for quadratic functions	Asymptotic convergence	Convergence in one step
Speed of convergence	Linear for quadratic functions	Quadratic (under proper hypotheses)

Table 2.1: Comparison between the gradient method and Newton's method.

2.6.3 Comparison between Newton's method and the gradient method

The gradient method and Newton's method can be compared from different point of views, as described in Table 2.1. From the table, it is obvious that Newton's method has better convergence properties but it is computationally more expensive. There exist methods which preserve some of the advantages of Newton's method, namely speed of convergence faster than the speed of the gradient method and finite convergence for quadratic functions, without requiring the knowledge of $\nabla^2 f$. Such methods are

- the conjugate directions methods;
- quasi-Newton methods.

2.7 Conjugate directions methods

Conjugate directions methods have been motivated by the need of improving the convergence speed of the gradient method, without requiring the computation of $\nabla^2 f$, as required in Newton's method.

A basic characteristic of conjugate directions methods is to find the minimizer of a quadratic function in a finite number of steps. These methods have been introduced for the solution of systems of linear equations and have later been extended to the solution of unconstrained optimization problems for non-quadratic functions.

Definition 5 *Given a matrix $Q = Q'$, the vectors d_1 and d_2 are said to be Q-conjugate if*

$$d_1' Q d_2 = 0.$$

Remark. If $Q = I$ then two vectors are Q-conjugate if they are orthogonal. $\diamond$

Theorem 10 *Let $Q \in I\!R^{n\times n}$ and $Q = Q' > 0$. Let $d_i \in I\!R^n$, for $i = 0, \cdots, k$, be non-zero vectors. If d_i are mutually Q-conjugate,* i.e.

$$d_i' Q d_j = 0,$$

for all $i \neq j$, then the vectors d_i are linearly independent.

Proof. Suppose there exists constants α_i, with $\alpha_i \neq 0$ for some i, such that

$$\alpha_0 d_0 + \cdots \alpha_k d_k = 0.$$

Then, left multiplying with Q and d_j' yields

$$\alpha_j d_j' Q d_j = 0,$$

which implies, as $Q > 0$, $\alpha_j = 0$. Repeating the same considerations for all $j \in [0, k]$ yields the claim. $\triangleleft$

Consider now a quadratic function

$$f(x) = \frac{1}{2} x' Q x + c' x + d,$$

with $x \in I\!R^n$ and $Q = Q' > 0$. The (global) minimizer of f is given by

$$x_\star = -Q^{-1} c,$$

and this can be computed using the procedure given in the next statement.

Theorem 11 *Let $Q = Q' > 0$ and let d_0, d_1, $\cdots$, d_{n-1} be n non-zero vectors mutually Q-conjugate. Consider the algorithm*

$$x_{k+1} = x_k + \alpha_k d_k$$

with

$$\alpha_k = -\frac{\nabla f(x_k)' d_k}{d_k' Q d_k} = -\frac{(x_k' Q + c') d_k}{d_k' Q d_k}.$$

Then, for any x_0, the sequence $\{x_k\}$ converges, in at most n steps, to $x_\star = -Q^{-1}c$, i.e. *it converges to the minimizer of the quadratic function f.*

Remark. Note that α_k is selected at each step to minimize the function $f(x_k + \alpha d_k)$ with respect to α, *i.e.* at each step an exact line search in the direction d_k is performed. $\diamond$

In the above statement we have assumed that the directions d_k have been preliminarily assigned. However, it is possible to construct a procedure in which the directions are computed iteratively. For, consider the quadratic function $f(x) = \frac{1}{2}x'Qx + c'x + d$, with $Q > 0$, and the following algorithm, known as conjugate gradient method.

Step 0. Given $x_0 \in I\!R^n$ and the direction

$$d_0 = -\nabla f(x_0) = -(Qx_0 + c).$$

Step 1. Set $k = 0$.

Step 2. Let

$$x_{k+1} = x_k + \alpha_k d_k$$

with

$$\alpha_k = -\frac{\nabla f(x_k)' d_k}{d_k' Q d_k} = -\frac{(x_k' Q + c') d_k}{d_k' Q d_k}.$$

Step 3. Compute d_{k+1} as follows

$$d_{k+1} = -\nabla f(x_{k+1}) + \beta_k d_k,$$

with

$$\beta_k = \frac{\nabla f(x_{k+1})' Q d_k}{d_k' Q d_k}.$$

Step 4. Set $k = k + 1$ and go to **Step 2**.

Remark. As already observed, α_k is selected to minimize the function $f(x_k + \alpha d_k)$. Moreover, this selection of α_k is also such that

$$\nabla f(x_{k+1})' d_k = 0. \tag{2.8}$$

In fact,

$$Qx_{k+1} = Qx_k + \alpha_k Q d_k$$

hence

$$\nabla f(x_{k+1}) = \nabla f(x_k) + \alpha_k Q d_k. \tag{2.9}$$

Left multiplying with d_k' yields

$$d_k' \nabla f(x_{k+1}) = d_k' \nabla f(x_k) + d_k' Q d_k \alpha_k = d_k' \nabla f(x_k) - d_k' Q d_k \frac{\nabla f(x_k)' d_k}{d_k' Q d_k} = 0.$$

$\diamond$

Remark. β_k is such that d_{k+1} is Q-conjugate with respect to d_k. In fact,

$$d_k'Qd_{k+1} = d_k'Q\left(-\nabla f(x_{k+1}) + \frac{\nabla f(x_{k+1})'Qd_k}{d_k'Qd_k}d_k\right) = d_k'Q\left(-\nabla f(x_{k+1}) + \nabla f(x_{k+1})\right) = 0.$$

Moreover, this selection of β_k yields also

$$\nabla f(x_k)'d_k = -\nabla f(x_k)'\nabla f(x_k). \tag{2.10}$$

$\diamond$

For the conjugate gradient method it is possible to prove the following fact.

Theorem 12 *The conjugate gradient method yields the minimizer of the quadratic function*

$$f(x) = \frac{1}{2}x'Qx + c'x + d,$$

with $Q = Q' > 0$, in at most n iterations, i.e. *there exists $m \leq n-1$ such that*

$$\nabla f(x_{m+1}) = 0.$$

Moreover

$$\nabla f(x_j)'\nabla f(x_i) = 0 \tag{2.11}$$

and

$$d_j'Qd_i = 0, \tag{2.12}$$

for all $[0, m+1] \ni i \neq j \in [0, m+1]$.

Proof. To prove the (finite) convergence of the sequence $\{x_k\}$ it is enough to show that the directions d_k are Q-conjugate, *i.e.* that equation (2.12) holds. In fact, if equation (2.12) holds the claim is a consequence of Theorem 11. $\triangleleft$

The conjugate gradient algorithm, in the form described above, cannot be used for the minimization of non-quadratic functions, as it requires the knowledge of the matrix Q, which is the Hessian of the function f. Note that the matrix Q appears at two levels in the algorithm: in the computation of the scalar β_k required to compute the new direction of research, and in the computation of the step α_k. It is therefore necessary to modify the algorithm to avoid the computation of $\nabla^2 f$, but at the same time it is reasonable to make sure that the modified algorithm coincides with the above one in the quadratic case.

2.7.1 Modification of β_k

To begin with note that, by equation (2.9), β_k can be written as

$$\beta_k = \frac{\nabla f(x_{k+1})'\dfrac{\nabla f(x_{k+1}) - \nabla f(x_k)}{\alpha_k}}{d_k'\dfrac{\nabla f(x_{k+1}) - \nabla f(x_k)}{\alpha_k}} = \frac{\nabla f(x_{k+1})'\left[\nabla f(x_{k+1}) - \nabla f(x_k)\right]}{d_k'\left[\nabla f(x_{k+1}) - \nabla f(x_k)\right]},$$

and, by equation (2.8),

$$\beta_k = -\frac{\nabla f(x_{k+1})' \left[\nabla f(x_{k+1}) - \nabla f(x_k)\right]}{d_k' \nabla f(x_k)}. \tag{2.13}$$

Using equation (2.13), it is possible to construct several expressions for β_k, all equivalent in the quadratic case, but yielding different algorithms in the general (non-quadratic) case. A first possibility is to consider equations (2.10) and (2.11) and to define

$$\beta_k = \frac{\nabla f(x_{k+1})' \nabla f(x_{k+1})}{\nabla f(x_k)' \nabla f(x_k)} = \frac{\|\nabla f(x_{k+1})\|^2}{\|\nabla f(x_k)\|^2}, \tag{2.14}$$

which is known as Fletcher-Reeves formula.
A second possibility is to write the denominator as in equation (2.14) and the numerator as in equation (2.13), yielding

$$\beta_k = \frac{\nabla f(x_{k+1})' \left[\nabla f(x_{k+1}) - \nabla f(x_k)\right]}{\|\nabla f(x_k)\|^2}, \tag{2.15}$$

which is known as Polak-Ribiere formula. Finally, it is possible to have the denominator as in (2.13) and the numerator as in (2.14), *i.e.*

$$\beta_k = -\frac{\|\nabla f(x_{k+1})\|^2}{d_k' \nabla f(x_k)}. \tag{2.16}$$

2.7.2 Modification of α_k

As already observed, in the quadratic version of the conjugate gradient method also the step α_k depends upon Q. However, instead of using the α_k given in **Step 2** of the algorithm, it is possible to use a line search along the direction α_k. In this way, an algorithm for non-quadratic functions can be constructed. Note that α_k, in the algorithm for quadratic functions, is also such that $d_k \nabla f(x_{k+1}) = 0$. Therefore, in the line search, it is reasonable to select α_k such that, not only $f(x_{k+1}) < f(x_k)$, but also d_k is approximately orthogonal to $\nabla f(x_{k+1})$.

Remark. The condition of approximate orthogonality between d_k and $\nabla f(x_{k+1})$ cannot be enforced using Armijo method or Goldstein conditions. However, there are more sophisticated line search algorithms, known as Wolfe conditions, which allow to enforce the above constraint. ◇

2.7.3 Polak-Ribiere algorithm

As a result of the modifications discussed in the last sections, it is possible to construct an algorithm for the minimization of general functions. For example, using equation (2.15) we obtain the following algorithm, due to Polak-Ribiere, which has proved to be one of the most efficient among the class of conjugate directions methods.

Step 0. Given $x_0 \in I\!R^n$.

Step 1. Set $k = 0$.

Step 2. Compute $\nabla f(x_k)$. If $\nabla f(x_k) = 0$ STOP. Else let

$$d_k = \begin{cases} -\nabla f(x_0), & \text{if } k = 0 \\ -\nabla f(x_k) + \dfrac{\nabla f(x_k)' \left[\nabla f(x_k) - \nabla f(x_{k-1})\right]}{\|\nabla f(x_{k-1})\|^2} d_{k-1}, & \text{if } k \geq 1. \end{cases}$$

Step 3. Compute α_k performing a line search along d_k.

Step 4. Set $x_{k+1} = x_k + \alpha_k d_k$, $k = k + 1$ and go to **Step 2**.

Remark. The line search has to be sufficiently accurate, to make sure that all directions generated by the algorithm are descent directions. A suitable line search algorithm is the so-called Wolfe method, which is a modification of Goldstein method. ◇

Remark. To guarantee global convergence of a subsequence it is possible to use, every n steps, the direction $-\nabla f$. In this case, it is said that the algorithm uses a *restart* procedure. For the algorithm with restart it is possible to have quadratic speed of convergence in n steps, *i.e*

$$\|x_{k+n} - x_\star\| \leq \gamma \|x_k - x_\star\|^2,$$

for some $\gamma > 0$. ◇

Remark. It is possible to modify Polak-Ribiere algorithm to make sure that at each step the angle condition holds. In this case, whenever the direction d_k does not satisfy the angle condition, it is sufficient to use the direction $-\nabla f$. Note that, enforcing the angle condition, yields a globally convergent algorithm. ◇

Remark. Even if the use of the direction $-\nabla f$ every n steps, or whenever the angle condition is not satisfied, allows to prove global convergence of Polak-Ribiere algorithm, it has been observed in numerical experiments that such modified algorithms do not perform as well as the original one. ◇

2.8 Quasi-Newton methods

Conjugate gradient methods have proved to be more efficient than the gradient method. However, in general, it is not possible to guarantee superlinear convergence. The main advantage of conjugate gradient methods is in the fact that they do not require to construct and store any matrix, hence can be used in large scale problems.

In small and medium scale problems, *i.e.* problems with less then a few hundreds decision variables, in which $\nabla^2 f$ is not available, it is convenient to use the so-called quasi-Newton methods.
Quasi Newton methods, as conjugate directions methods, have been introduced for quadratic functions. They are described by an algorithm of the form

$$x_{k+1} = x_k - \alpha_k H_k \nabla f(x_k),$$

with H_0 given. The matrix H_k is an approximation of $[\nabla^2 f(x_k)]^{-1}$ and it is computed iteratively at each step.
If f is a quadratic function, the gradient of f is given by

$$\nabla f(x) = Qx + c,$$

for some Q and c, hence for any $x \in I\!R^n$ and $y \in I\!R^n$ one has

$$\nabla f(y) - \nabla f(x) = Q(y - x),$$

or, equivalently,

$$Q^{-1}[\nabla f(y) - \nabla f(x)] = y - x.$$

It is then natural, in general, to construct the sequence $\{H_k\}$ such that

$$H_{k+1}[\nabla f(x_{k+1}) - \nabla f(x_k)] = x_{k+1} - x_k. \tag{2.17}$$

Equation (2.17) is known as quasi-Newton equation.
There exist several update methods satisfying the quasi-Newton equation. For simplicity, set

$$\gamma_k = \nabla f(x_{k+1}) - \nabla f(x_k),$$

and

$$\delta_k = x_{k+1} - x_k.$$

As a result, equation (2.17) can be rewritten as

$$H_{k+1}\gamma_k = \delta_k.$$

One of the first quasi-Newton methods has been proposed by Davidon, Fletcher and Powell, and can be summarized by the equations

$$\text{DFP}\left\{\begin{array}{lcl} H_0 & = & I \\ & & \\ H_{k+1} & = & H_k + \dfrac{\delta_k \delta_k'}{\delta_k' \gamma_k} - \dfrac{H_k \gamma_k \gamma_k' H_k}{\gamma_k' H_k \gamma_k}. \end{array}\right. \tag{2.18}$$

It is easy to show that the matrix H_{k+1} satisfies the quasi-Newton equation (2.17), *i.e.*

$$\begin{array}{rcl} H_{k+1}\gamma_k & = & H_k\gamma_k + \dfrac{\delta_k \delta_k'}{\delta_k' \gamma_k}\gamma_k - \dfrac{H_k \gamma_k \gamma_k' H_k}{\gamma_k' H_k \gamma_k}\gamma_k \\ & = & H_k\gamma_k + \dfrac{\delta_k' \gamma_k}{\delta_k' \gamma_k}\delta_k - \dfrac{\gamma_k' H_k \gamma_k}{\gamma_k' H_k \gamma_k} H_k\gamma_k \\ & = & \delta_k. \end{array}$$

Moreover, it is possible to prove the following fact, which gives conditions such that the matrices generated by DFP method are positive definite for all k.

Theorem 13 *Let $H_k = H_k' > 0$ and assume $\delta_k'\gamma_k > 0$. Then the matrix*

$$H_k + \frac{\delta_k\delta_k'}{\delta_k'\gamma_k} - \frac{H_k\gamma_k\gamma_k'H_k}{\gamma_k'H_k\gamma_k}$$

is positive definite.

DFP method has the following properties. In the quadratic case, if α_k is selected to minimize

$$f(x_k - \alpha H_k \nabla f(x_k)),$$

then

- the directions $d_k = -H_k\nabla f(x_k)$ are mutually conjugate;
- the minimizer of the (quadratic) function is found in at most n steps, moreover $H_n = Q^{-1}$;
- the matrices H_k are always positive definite.

In the non-quadratic case

- the matrices H_k are positive definite (hence $d_k = -H_k\nabla f(x_k)$ is a descent direction) if $\delta_k'\gamma_k > 0$;
- it is globally convergent if f is strictly convex and if the line search is exact;
- it has superlinear speed of convergence (under proper hypotheses).

A second, and more general, class of update formulae, including as a particular case DFP formula, is the so-called Broyden class, defined by the equations

$$\text{Broyden}\left\{\begin{array}{rcl} H_0 & = & I \\ H_{k+1} & = & H_k + \dfrac{\delta_k\delta_k'}{\delta_k'\gamma_k} - \dfrac{H_k\gamma_k\gamma_k'H_k}{\gamma_k'H_k\gamma_k} + \phi v_k v_k', \end{array}\right. \tag{2.19}$$

with $\phi \geq 0$ and

$$v_k = (\gamma_k'H_k\gamma_k)^{1/2}\left(\frac{\delta_k}{\delta_k'\gamma_k} - \frac{H_k\gamma_k}{\gamma_k'H_k\gamma_k}\right).$$

If $\phi = 0$ then we obtain DFP formula, whereas for $\phi = 1$ we have the so-called Broyden-Fletcher-Goldfarb-Shanno (BFGS) formula, which is one of the preferred algorithms in applications. From Theorem 13 it is easy to infer that, if $H_0 > 0$, $\gamma_k'\delta_k > 0$ and $\phi \geq 0$, then all formulae in the class of Broyden generate matrices $H_k > 0$.

Remark. Note that the condition $\delta_k'\gamma_k > 0$ is equivalent to

$$(\nabla f(x_{k+1}) - \nabla f(x_k))'\, d_k > 0,$$

and this can be enforced with a sufficiently precise line search. ◇

For the method based on BFGS formula, a global convergence result, for convex functions and in the case of non-exact (but sufficiently accurate) line search, has been proved. Moreover, it has been shown that the algorithm has superlinear speed of convergence. This algorithm can be summarized as follows.

Step 0. Given $x_0 \in I\!R^n$.

Step 1. Set $k = 0$.

Step 2. Compute $\nabla f(x_k)$. If $\nabla f(x_k) = 0$ STOP. Else compute H_k with BFGS equation and set

$$d_k = -H_k \nabla f(x_k).$$

Step 3. Compute α_k performing a line search along d_k.

Step 4. Set $x_{k+1} = x_k + \alpha_k d_k$, $k = k + 1$ and go to **Step 2**.

In the general case it is not possible to prove global convergence of the algorithm. However, this can be enforced verifying (at the end of **Step 2**), if the direction d_k satisfies an angle condition, and if not use the direction $d_k = -\nabla f(x_k)$. However, as already observed, this modification improves the convergence properties, but reduces (sometimes drastically) the speed of convergence.

2.9 Methods without derivatives

All the algorithms that have been discussed presuppose the knowledge of the derivatives (first and/or second) of the function f. There are, however, also methods which do not require such a knowledge. These methods can be divided in two classes: direct research methods and methods using finite difference approximations.
Direct search methods are based upon the direct comparison of the values of the function f in the points generated by the algorithm, without making use of the necessary condition of optimality $\nabla f = 0$. In this class, the most interesting methods, *i.e.* the methods for which it is possible to give theoretical results, are those that make use cyclically of n linearly independent directions. The simplest possible method, known as the method of the coordinate directions, can be described by the following algorithm.

Step 0. Given $x_0 \in I\!R^n$.

Step 1. Set $k = 0$.

Step 2. Set $j = 1$.

Step 3. Set $d_k = e_j$, where e_j is the j-th coordinate direction.

Step 4. Compute α_k performing a line search without derivatives along d_k.

Step 5. Set $x_{k+1} = x_k + \alpha_k d_k$, $k = k + 1$.

Step 6. If $j < n$ set $j = j + 1$ and go to **Step 3**. If $j = n$ go to **Step 2**.

It is easy to verify that the matrix

$$P_k = \left[\begin{array}{cccc} d_k & d_{k+1} & \cdots & d_{k+n-1} \end{array} \right]$$

is such that

$$|\mathrm{det} P_k| = 1,$$

hence, if the line search is such that

$$\lim_{k\to\infty} \frac{\nabla f(x_k)' d_k}{\|d_k\|} = 0$$

and

$$\lim_{k\to\infty} \|x_{k+1} - x_k\| = 0,$$

convergence to stationary points is ensured by the general result in Theorem 5. Note that, the line search can be performed using the parabolic line search method described in Section 2.4.4.
The method of the coordinate directions is not very efficient, in terms of speed of convergence. Therefore, a series of heuristics have been proposed to improve its performance. One such heuristics is the so-called method of Jeeves and Hooke, in which not only the search along the coordinate directions is performed, but also a search along directions joining pairs of points generated by the algorithm. In this way, the search is performed along what may be considered to be the most promising directions.
An alternative direct search method is the so-called simplex method (which should not be confused with the simplex method of linear programming). The method starts with $n + 1$ (equally spaced) points $x_{(i)} \in I\!R^n$ (these points give a simplex in $I\!R^n$). In each of these points the function f is computed and the vertex where the function f attains the maximum value is determined. Suppose this is the vertex $x_{(n+1)}$. This vertex is reflected with respect to the center of the simplex, *i.e.* the point

$$x_c = \frac{1}{n+1} \sum_{i=1}^{n+1} x_{(i)}.$$

As a result, the new vertex

$$x_{(n+2)} = x_c + \alpha(x_c - x_{(n+1)})$$

where $\alpha > 0$, is constructed, see Figure 2.7. The procedure is then repeated.

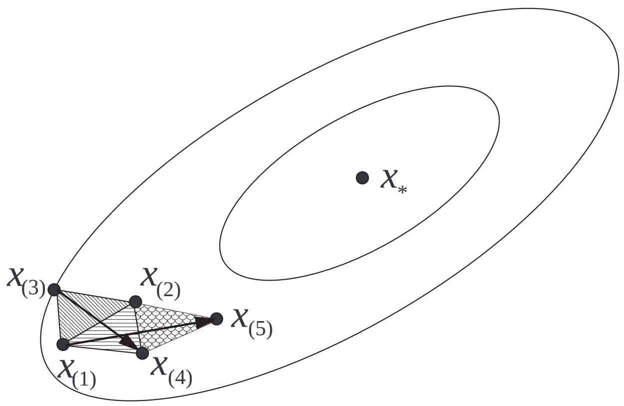


Figure 2.7: The simplex method. The points $x_{(1)}$, $x_{(2)}$ and $x_{(3)}$ yields the starting simplex. The second simplex is given by the points $x_{(1)}$, $x_{(2)}$ and $x_{(4)}$. The third simplex is given by the points $x_{(2)}$, $x_{(4)}$ and $x_{(5)}$.

It is possible that the vertex that is generated by one step of the algorithm is (again) the one where the function f has its maximum. In this case, the algorithm cycles, hence the next vertex has to be determined using a different strategy. For example, it is possible to construct the next vertex by reflecting another of the remaining n vertex, or to shrink the simplex.

As a stopping criterion it is possible to consider the condition

$$\frac{1}{n+1}\sum_{i=1}^{n+1}\left(f(x_{(i)})-\bar{f}\right)^2<\epsilon \tag{2.20}$$

where $\epsilon > 0$ is assigned by the designer, and

$$\bar{f}=\frac{1}{n+1}\sum_{i=1}^{n+1}f(x_{(i)}),$$

i.e. $\bar{f}$ is the mean value of the $f(x_{(i)})$. Condition (2.20) implies that the points $x_{(i)}$ are all in a region where the function f is *flat*.

As already observed, direct search methods are not very efficient, and can be used only for problems with a few decision variables and when approximate solutions are acceptable. As an alternative, if the derivatives of the function f are not available, it is possible to resort to numeric differentiation, *e.g* the entries of the gradient of f can be computed using the so-called forward difference approximation, *i.e.*

$$\frac{\partial f(x)}{\partial x_i}\approx\frac{f(x+te_i)-f(x)}{t},$$

where e_i is the i-th column of the identity matrix of dimension n, and $t > 0$ has to be fixed by the user. Note that there are methods for the computation of the optimal value of t, *i.e.* the value of t which minimizes the approximation error.

2.10 Exercises

This section contains a set of exercises related to the notions, concepts, algorithms and tools discussed in Chapter 2, together with exercises on topics which have not been covered in the book. In these cases, the specific topic is briefly illustrated in the text of the exercise. The objective is to draw the reader's attention to the fact that there are many more ideas, methods and algorithms which have been developed to solve unconstrained optimization problems and which are not covered in the book. The basic principles provided in Chapter 2 should however be sufficient to understand more advanced and involved methods. Not surprisingly, a significant number of exercises is devoted to Newton's method and its modifications: undoubtedly, Newton's method is one of the most important methods in optimization (and numerical analysis). All exercises have a brief worked out solution, which provides guidelines and checkpoints to help the reader assess their level of understanding and familiarity with the content covered. Note that the exercises are not ordered in any particular way: the order is the result of the history of my optimization course and of the exam papers I have set over the years.

Exercise 1 Consider the problem of minimizing the function

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

a) Compute the unique stationary point of the function, and show that the function is radially unbounded.

b) Using second order sufficient conditions show that the stationary point determined in part a) is a local minimizer. Also show that the point is a global minimizer.

c) Consider the minimization of the function f using the gradient algorithm. Express analytically the form of the generic iteration, *i.e.*

$$p_{k+1} = p_k - \alpha \nabla f,$$

where $p_i = [x_1^i, x_2^i]'$.

d) Consider the initial point $p_0 = [1, 1]'$ and apply one step of the gradient algorithm from part c) with exact line search. Verify that p_1 coincides with the stationary point determined in part a).

e) It is known that for quadratic functions, such as the function f above, the gradient algorithm is globally convergent, however the speed of convergence may be very slow. Discuss why, for the function f, the gradient algorithm with exact line search converges in one step.

Solution 1

a) The stationary points of the function f are computed solving the equation

$$0 = \nabla f = \left[\begin{array}{c} 2x_1 - 1 \\ 2x_2 \end{array} \right],$$

yielding the unique stationary point $x_1^\star = 1/2$ and $x_2^\star = 0$. The function f is of the form $x'Qx + c'x$ with $Q = \text{diag}(1, 1) > 0$, hence it is radially unbounded.

b) Note that $\nabla^2 f = \mathrm{diag}(2,2) > 0$, hence $x^\star$ is a local minimizer. It is also a global minimizer for the following reasons: the function f is C^1 and radially unbounded, therefore the global minimizer is a stationary point.

c) The generic iteration of the gradient algorithm for the considered function f is

$$x_1^{k+1} = x_1^k - \alpha(2x_1^k - 1) \qquad x_2^{k+1} = x_2^k - \alpha(2x_2^k).$$

d) Let $x_1^0 = x_2^0 = 1$. Hence

$$x_1^1 = 1 - \alpha \qquad x_2^1 = 1 - 2\alpha.$$

Note now that $f(x_1^1, x_2^1) = 1 - 5\alpha + 5\alpha^2$ and this is minimized by $\alpha^\star = 1/2$, yielding

$$x_1^1 = 1 - \alpha^\star = 1/2 = x_1^\star \qquad x_2^1 = 1 - 2\alpha^\star = 0 = x_2^\star.$$

e) For the considered function the gradient algorithm with exact line search converges in one step (from any initial point) because the function is quadratic and the minimum and maximum eigenvalues of $\nabla^2 f$ coincide.

Exercise 2 Consider the problem of minimizing the function

$$f(x) = x_1^4 + x_1 x_2 + \frac{1}{2} x_2^2.$$

a) Compute the stationary points of the function.

b) Using second order sufficient conditions *classify* the stationary points determined in part a), *i.e.* say which is a local minimizer, or a local maximizer, or a saddle point.

c) Consider the minimization of the function f using Newton's algorithm. Express analytically the form of the generic iteration, *i.e.*

$$p_{k+1} = p_k - [\nabla^2 f]^{-1} \nabla f,$$

where $p_i = [x_1^i, x_2^i]'$.

d) The equation in part c) defines a nonlinear discrete-time system with equilibria coinciding with the stationary points of the function f.

Consider the linear approximation of the system in part c) around the equilibrium corresponding to the local minimizer of the function f with $x_1 > 0$, and compute the eigenvalues associated with the linear approximation.

Interpret the result obtained in terms of convergence properties of sequences generated by Newton's algorithm and initialized close to a local minimizer.

e) Consider the initial point

$$p_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and, using the results in part c), apply four steps of Newton's algorithm to generate the points p_1, p_2, p_3, p_4. Comment on the speed of convergence of the sequence.

Solution 2

a) The stationary point of the function f are computed solving the equation

$$0 = \nabla f = \begin{bmatrix} 4x_1^3 + x_2 \\ x_1 + x_2 \end{bmatrix},$$

yielding the stationary points

$$p^\star = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \tilde{p}^\star = \begin{bmatrix} -1/2 \\ 1/2 \end{bmatrix}, \qquad \hat{p}^\star = \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix}.$$

b) Note that

$$\nabla^2 f(p^\star) = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \not\geq 0 \qquad \nabla^2 f(\tilde{p}^\star) = \nabla^2 f(\hat{p}^\star) = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} > 0.$$

Hence, $p^\star$ is a saddle point, and $\tilde{p}^\star$ and $\hat{p}^\star$ are local minimizers.

c) The generic iteration of Newton's algorithm for the considered function f is

$$x_1^{k+1} = \frac{8(x_1^k)^3}{12(x_1^k)^2 - 1}, \qquad x_2^{k+1} = -\frac{8(x_1^k)^3}{12(x_1^k)^2 - 1}.$$

d) The linear approximation of the above nonlinear discrete-time system around the point $\hat{p}^\star$ is the system

$$p_{k+1} = Ap_k = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} p_k.$$

This linear system is such that $p_1 = 0$ for any p_0, and this explains the local 'fast' speed of convergence of Newton's iteration.

e) A simple computation yields the sequence

$$p_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad p_1 = \begin{bmatrix} 0.7272727273 \\ -0.7272727273 \end{bmatrix}, \qquad p_2 = \begin{bmatrix} 0.57552339460 \\ -0.57552339460 \end{bmatrix},$$

$$p_3 = \begin{bmatrix} 0.51266296461 \\ -0.51266296461 \end{bmatrix}, \qquad p_4 = \begin{bmatrix} 0.5004542259 \\ -0.5004542259 \end{bmatrix},$$

and this shows the fast convergence (approximately two exact digits for each iteration) of Newton's algorithm.

Exercise 3 Consider the function

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

a) Compute all stationary points of the function.
(Hint: obtain first $(x_1 - x_2)^3$ in terms of x_1 from the necessary conditions of optimality.)

b) Using second order sufficient conditions, *classify* the stationary points determined in part a), *i.e.* say which is a local minimizer, or a local maximizer, or a saddle point.

c) Consider the point $p = (0,0)$ and the direction $d = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$. Using the definition of a descent direction, show that d is a descent direction for f at p.

d) Perform an exact line search along the direction $d = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ starting at $p = (0,0)$. Show that the point obtained as a result of the line search procedure is a local minimizer of the function f.

Solution 3

a) The stationary points of the function f are computed by solving the equation

$$0 = \nabla f = \begin{bmatrix} 2x_1 + x_2 + 4(x_1 - x_2)^3 \\ x_1 - 4(x_1 - x_2)^3 \end{bmatrix},$$

yielding

$$P_1 = (0,0), \qquad P_2 = \left(-\frac{1}{16}, \frac{3}{16}\right), \qquad P_3 = \left(\frac{1}{16}, -\frac{3}{16}\right).$$

b) Note that

$$\nabla^2 f = \begin{bmatrix} 2+12(x_1-x_2)^2 & 1-12(x_1-x_2)^2 \\ 1-12(x_1-x_2)^2 & 12(x_1-x_2)^2 \end{bmatrix}.$$

Thus

$$\nabla^2 f(P_1) = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix},$$

which is an indefinite matrix, and

$$\nabla^2 f(P_2) = \nabla^2 f(P_3) = \begin{bmatrix} 11/4 & 1/4 \\ 1/4 & 3/4 \end{bmatrix},$$

which is a positive definite matrix. As a result, P_1 is a saddle point, and P_2 and P_3 are local minimizers.

c) By definition, a direction d is a descent direction for f at p if there exists $\delta > 0$ such that

$$f(p + \lambda d) < f(p),$$

for all $\lambda \in (0, \delta)$. Consider now the given direction and note that $f(p) = 0$ and that

$$f(p + \lambda d) = -2\lambda^2 + 256\lambda^4.$$

Hence, $f(p + \lambda d) < f(p)$ for all $\lambda > 0$ and sufficiently small. (Note that $\nabla f(p)'d = 0$, hence this condition cannot be used to decide if d is a descent direction, or otherwise.)

d) To perform an exact line search along the direction d, starting from $p = (0, 0)$, we need to find the minimum of the function

$$\phi(\lambda) = f(p + \lambda d) - f(p) = -2\lambda^2 + 256\lambda^4$$

for $\lambda > 0$. Note that

$$\frac{d\phi}{d\lambda} = -4\lambda + 1024\lambda^3,$$

hence the minimum is achieved for $\lambda = 1/16$. The resulting point is $(1/16, -3/16)$ and this coincides with one of the local minimizers determined in part a).

Exercise 4 Consider the problem of minimizing the function

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

a) Compute the stationary points of the function.

b) Consider the minimization of the function f using the gradient algorithm. Express analytically the form of the generic iteration, *i.e.*

$$p_{k+1} = p_k - \alpha \nabla f,$$

where $p_i = [x_1^i, x_2^i]'$.

c) Compute three steps of the gradient algorithm with exact line search from the initial point $p_0 = [0, 0]'$, using the fact that, for this p_0 the exact line search parameter α is equal to $1/3$ for all k. Check that indeed $\alpha^\star = 1/3$ for the first iteration.

d) Exploit the results of part c) to show that the gradient iteration with exact line search for $p_0 = [0, 0]'$ gives

$$x_1^{k+1} = \frac{1}{3}x_1^k - \frac{4}{3},$$

$$x_2^{k+1} = -\frac{1}{3}x_2^k - \frac{4}{3},$$

and hence show that

$$(x_1^{k+1} + 2) = \frac{1}{3}(x_1^k + 2),$$

$$(x_2^{k+1} + 1) = -\frac{1}{3}(x_2^k + 1).$$

Hence, deduce that the sequence $\{p_k\}$ can be written as

$$p_{k+1} = \begin{bmatrix} \frac{2}{3^{k+1}} - 2 \\ \left(-\frac{1}{3}\right)^{k+1} - 1 \end{bmatrix}.$$

Show that the sequence $\{p_k\}$ converges to the stationary point determined in part a).

Solution 4

a) The stationary points of the function f are computed by solving the equation

$$0 = \nabla f = \begin{bmatrix} 2x_1 + 4 \\ 4x_2 + 4 \end{bmatrix},$$

yielding the stationary point

$$p^\star = (-2, 1).$$

b) The generic iteration of the gradient algorithm for the considered function f is

$$x_1^{k+1} = x_1^k - \alpha(2x_1^k + 4), \qquad x_2^{k+1} = x_2^k - \alpha(4x_2^k + 4).$$

c) Setting $(x_1^0, x_2^0) = (0, 0)$ one has $(x_1^1, x_2^1) = (-4\alpha, -4\alpha)$ and

$$f(-4\alpha, -4\alpha) - f(0, 0) = 48\alpha^2 - 32\alpha.$$

Minimizing this function yields $\alpha^\star = 1/3$ (as stated). Therefore, $(x_1^1, x_2^1) = (-4/3, -4/3)$. Repeating the same considerations, and setting always $\alpha = 1/3$, one has

$$(x_1^2, x_2^2) = (-16/9, -8/9)$$

and

$$(x_1^3, x_2^3) = (-52/27, -28/27).$$

d) Setting $\alpha = 1/3$ in the gradient iteration yields

$$x_1^{k+1} = \frac{1}{3}x_1^k - \frac{4}{3}, \qquad x_2^{k+1} = -\frac{1}{3}x_2^k - \frac{4}{3},$$

and this can be also written as

$$(x_1^{k+1} + 2) = \frac{1}{3}(x_1^k + 2), \qquad (x_2^{k+1} + 1) = -\frac{1}{3}(x_2^k + 1).$$

As a result

$$(x_1^{k+1} + 2) = \left(\frac{1}{3}\right)^k (x_1^0 + 2), \qquad (x_2^{k+1} + 1) = \left(-\frac{1}{3}\right)^k (x_2^0 + 1),$$

or, equivalently,

$$x_1^{k+1} = 2\left(\frac{1}{3}\right)^k - 2, \qquad x_2^{k+1} = \left(-\frac{1}{3}\right)^k - 1.$$

Finally, as $k \to \infty$ $x_1^k \to -2$ and $x_2^k \to -1$, *i.e.* the sequence converges to the stationary point determined in part a).

Exercise 5

a) An electrical engineer wants to maximize the current I between two points A and B of a complex network by adjusting the values x_1 and x_2 of two variable resistors. The engineer does not have a model of the network and decides to opt for this procedure.

- Keep the value x_2 fixed and adjust x_1 to maximize I.
- Keep the value x_1 fixed and adjust x_2 to maximize I.
- Repeat the above steps until no further improvement can be obtained.

Explain if this approach has sound theoretical basis, *i.e.* discuss under what assumptions the above procedure determines a stationary point of the function I.

b) Let $f : I\!R^n \to I\!R$ be a differentiable function. Suppose that $x_\star$ is a local minimizer of f along every line that passes through $x_\star$, *i.e.* the function

$$g(\alpha) = f(x_\star + \alpha d)$$

is minimized at $\alpha = 0$ for all $d \in I\!R^n$.

i) Show that $\nabla f(x_\star) = 0$.

ii) Is $x_\star$ a local minimizer of f?

iii) Consider the function

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

Show that the point $(0, 0)$ is a local minimizer of f along every line that passes through (0,0). Show that the point $(0, 0)$ is not a local minimizer of f.
(Hint: consider the values of f for $x_1 = y$ and $x_2 = my^2$ and $m \in I\!R$.)

Solution 5

a) The engineer is applying the so-called coordinate directions method with an exact line search (without derivatives), as described in Section 2.9, for the minimization of the function $-I = -I(x_1, x_2)$. This approach provides a sequence of points converging to a stationary point of the function I provided that the initial point is selected inside a compact level set of $-I(x_1, x_2)$.

b) i) Note that, by assumption, the function

$$\frac{dg}{d\alpha} = \nabla f(x_\star + \alpha d)' d$$

is zero for $\alpha = 0$ and for every d. This means that

$$\nabla f(x_\star)' d = 0$$

for every d, and this implies that

$$\nabla f(x_\star) = 0.$$

ii) Without further information on f it is not possible to draw any conclusion on $x_\star$, *i.e.* $x_\star$ is a stationary point of f, but it may be a local minimizer, a local maximizer or a saddle point.

iii) Consider a line that goes through zero, namely $x_2 = \gamma x_1$, and note that

$$f(x_1, \gamma x_1) = (\gamma x_1 - x_1^2)(\gamma x_1 - 2x_1^2) = \gamma^2 x_1^2 - 3\gamma x_1^3 + 2x_1^4$$

and this shows that for all γ the point $x_1 = 0$ is a local minimizer of $f(x_1, \gamma x_1)$. For completeness we have also to consider the line $x_1 = 0$ (which corresponds formally to $\gamma = \infty$). Note that

$$f(0, x_2) = x_2^2$$

hence the point $x_2 = 0$ is a minimizer of $f(0, x_2)$.
To show that $(0, 0)$ is not a local minimizer of f note first that $f(0, 0) = 0$ and then let $x_1 = y$ and $x_2 = my^2$. Note that

$$f(y, my^2) = y^4(m - 1)(m - 2).$$

Pick $m \in (1, 2)$ and note that for such values of m

$$f(y, my^2) = y^4(m - 1)(m - 2) < 0$$

for all $y \neq 0$. This shows that close to the point $(0, 0)$, where the function is zero, there are points in which the function takes negative values. Hence, $(0, 0)$ is not a local minimizer of f.

Exercise 6 Consider the problem of minimizing the function

$$f(x_1, x_2) = \frac{1}{3}x_1^2 - \alpha x_1^4 + \frac{1}{4}x_1^6 + x_1 x_2 + x_2^2,$$

where α is a constant.

a) Compute all stationary points of the function.

b) Let $\alpha = 5/12$. Using second order sufficient conditions *classify* the stationary points determined in part a), *i.e.* say which is a local minimizer, or a local maximizer, or a saddle point.

c) Let $\alpha = 5/12$. Show that the function f is radially unbounded and hence compute the global minimum of f. Is the global minimizer unique?

Solution 6

a) The stationary points of the function f are computed by solving the equation

$$0 = \nabla f = \begin{bmatrix} 2/3 x_1 - 4\alpha x_1^3 + 3/2 x_1^5 + x_2 \\ x_1 + 2x_2 \end{bmatrix}.$$

Solving the second equation for x_2 yields $x_2 = -1/2\, x_1$, and upon replacement in the first equation we obtain

$$\frac{1}{6}x_1 - 4\alpha x_1^3 + \frac{3}{2}x_1^5 = 0,$$

yielding

$$x_{1a} = 0, \qquad x_{1b} = \frac{1}{3}\sqrt{12\alpha + 3\sqrt{16\alpha^2 - 1}}, \qquad x_{1c} = -\frac{1}{3}\sqrt{12\alpha + 3\sqrt{16\alpha^2 - 1}},$$

$$x_{1d} = \frac{1}{3}\sqrt{12\alpha - 3\sqrt{16\alpha^2 - 1}}, \qquad x_{1e} = -\frac{1}{3}\sqrt{12\alpha - 3\sqrt{16\alpha^2 - 1}}.$$

b) For $\alpha = 5/12$ we obtain the stationary points

$$P_a = (0, 0), \qquad P_b = (1, -1/2), \qquad P_c = (-1, 1/2),$$

$$P_d = (1/3, -1/6), \qquad P_e = (-1/3, 1/6).$$

Note now that, for $\alpha = 5/12$,

$$\nabla^2 f = \begin{bmatrix} 2/3 - 5x_1^2 + 15/2 x_1^4 & 1 \\ 1 & 2 \end{bmatrix}$$

and that

$$\nabla^2 f(P_a) = \begin{bmatrix} 2/3 & 1 \\ 1 & 2 \end{bmatrix} > 0$$

$$\nabla^2 f(P_b) = \nabla^2 f(P_c) = \begin{bmatrix} 19/6 & 1 \\ 1 & 2 \end{bmatrix} > 0$$

$$\nabla^2 f(P_d) = \nabla^2 f(P_e) = \begin{bmatrix} 11/54 & 1 \\ 1 & 2 \end{bmatrix} \not\geq 0.$$

As a result, P_a, P_b and P_c are local minimizers, and P_d and P_e are saddle points.

c) Note that

$$-\frac{5}{12}x_1^4 + \frac{1}{4}x_1^6 = x_1^4\left(\frac{1}{4}x_1^2 - \frac{5}{12}\right)$$

is radially unbounded. Hence

$$f(x_1, x_2) = (\frac{1}{3}x_1^2 + x_1 x_2 + x_2^2) + x_1^4\left(\frac{1}{4}x_1^2 - \frac{5}{12}\right)$$

is also radially unbounded. The global minimum of f is also a local minimum of f. Note that

$$f(P_a) = 0 \qquad f(P_b) = f(P_c) = -0.833\cdots.$$

Hence, P_b and P_c are both global minimizers, therefore the global minimizer is not unique.

Exercise 7 Consider the problem of minimizing the function

$$f(x) = x - \log x,$$

with $x > 0$.

a) Compute analytically the minimizer of f.

b) Write Newton's iteration for the considered problem.

c) Consider the Newton's iteration in part b) with initial point $x_0 = 1.99$. Compute ten steps of the Newton's iteration. Argue that the resulting sequence converges to the minimizer of f. Show that the sequence converges to the minimizer of f with quadratic speed of convergence.

d) Consider the Newton's iteration in part b) with initial point $x_0 = 2.01$. Compute five steps of the Newton's iteration. Argue that the resulting sequence diverges.

e) Consider the Newton's iteration in part b). Show that

   i) if the initial point $x_0 = 2$ then $x_k = 0$, for all $k \geq 1$;

   ii) if the initial point $x_0 = 0$ then $x_k = 0$, for all $k \geq 1$;

   iii) if the initial point $x_0 > 2$ then $x_k < 0$, for all $k \geq 1$ and the sequence does not converge;

   iv) if the initial point $x_0 \in (0, 2)$ then $x_k \in (0, 2)$, for all $k \geq 1$ and the sequence converges to the minimizer determined in part a).

Solution 7

a) The minimizer of f is obtained solving $\nabla f = 1 - 1/x = 0$, yielding $x = 1$. Note that $x = 1$ is indeed a minimizer (a global one), because the function f is convex for all $x > 0$.

b) The Newton's iteration is

$$x_{k+1} = x_k - \frac{1}{\nabla^2 f(x_k)} \nabla f(x_k) = x_k - x_k^2(1 - \frac{1}{x_k}) = (2 - x_k)x_k.$$

c) Let $x_0 = 1.99$ then

$$\begin{aligned} x_1 &= 0.01990 \\ x_2 &= 0.03940399 \\ x_3 &= 0.07725530557208 \\ x_4 &= 0.14854222890512 \\ x_5 &= 0.27501966404215 \\ x_6 &= 0.47440351247444 \\ x_7 &= 0.72374833230079 \\ x_8 &= 0.92368501609341 \\ x_9 &= 0.99417602323134 \\ x_{10} &= 0.99996608129460. \end{aligned}$$

The sequence is converging to $x = 1$, *i.e.* to the local minimizer of f. To establish quadratic speed of convergence note that

$$\frac{\mathcal{E}_{k+1}}{\mathcal{E}_k^2} = \frac{|x_{k+1} - 1|}{|x_k - 1|^2} = \frac{|(2 - x_k)x_k - 1|}{(x_k - 1)^2} = 1.$$

d) Let $x_0 = 2.01$ then

$$\begin{array}{rcl} x_1 & = & -0.02010 \\ x_2 & = & -0.04060401 \\ x_3 & = & -0.08285670562808 \\ x_4 & = & -0.17257864492369 \\ x_5 & = & -0.37494067853109. \end{array}$$

We then infer that the sequence is monotonically decreasing and $\lim_{k\to\infty} x_k = -\infty$.

e) The first two points are trivial noting that

$$x_{k+1} = (2 - x_k)x_k$$

and the right hand side of this equation is zero for $x_k = 0$ or $x_k = 2$, that is $x = 0$ and $x = 2$ are equilibria of the above discrete-time system. Note now that if $x_0 > 2$ then $x_1 < 0$. Moreover if $x_k < 0$ then

$$x_{k+1} = (2 - x_k)x_k < x_k,$$

which proves the third claim. Finally, if $x_k \in (0, 2)$ then it is easy to verify that

$$0 < x_{k+1} = (2 - x_k)x_k < 2.$$

Moreover, if $x_k = 1$ then $x_{k+1} = 1$, hence $x = 1$ is an equilibrium of the discrete-time system $x_{k+1} = (2 - x_k)x_k$. Finally, if $x_k \in (1, 2)$

$$0 < x_{k+1} < 1,$$

and if $x_k \in (0, 1)$

$$x_k < x_{k+1} < 1,$$

which shows convergence of the sequence to $x = 1$.

Exercise 8 Consider the problem of minimizing the function

$$f(x_1, x_2) = \frac{1}{2n+2}x_1^{2n+2} - x_1x_2 + \frac{1}{2}x_2^2,$$

where n is a positive integer.

a) Compute all stationary points of the function.

b) Using second order sufficient conditions *classify* the stationary points determined in part a), *i.e.* say which is a local minimizer, or a local maximizer, or a saddle point.

c) Show that the function f is radially unbounded and hence compute the global minimum of f. Is the global minimizer unique?

d) Consider the point $P_0 = (0, 0)$ and the direction

$$d = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Show that the direction d is a descent direction for f at P_0.

Solution 8

a) The stationary points of the function f are computed by solving the equation

$$0 = \nabla f = \begin{bmatrix} x_1^{2n+1} - x_2 \\ -x_1 + x_2 \end{bmatrix}.$$

The second equation yields $x_2 = x_1$, hence the first equation becomes

$$0 = x_1^{2n+1} - x_1 = x_1(x_1^{2n} - 1).$$

The (real) solutions of this equation are $x_1 = 0$, $x_1 = 1$ and $x_1 = -1$. In summary, the function f has three stationary points

$$P_a = (0, 0), \qquad P_b = (1, 1), \qquad P_c = (-1, -1).$$

b) Note that (recall that n is a positive integer)

$$\nabla^2 f = \begin{bmatrix} (2n+1)x_1^{2n} & -1 \\ -1 & 1 \end{bmatrix}.$$

Hence

$$\nabla^2 f(P_a) = \begin{bmatrix} 0 & -1 \\ -1 & 1 \end{bmatrix}$$

which is an indefinite matrix, and

$$\nabla^2 f(P_b) = \nabla^2 f(P_c) = \begin{bmatrix} 2n+1 & -1 \\ -1 & 1 \end{bmatrix} > 0.$$

As a result P_a is a saddle point, and P_b and P_c are local minimizers.

c) Note that

$$f = \frac{1}{2n+2}x_1^{2n+2} - x_1x_2 + \frac{1}{2}x_2^2 = \frac{1}{2n+2}x_1^{2n+2} - x_1^2 + \left(x_1^2 - x_1x_2 + \frac{1}{2}x_2^2\right).$$

The function

$$\frac{1}{2n+2}x_1^{2n+2} - x_1^2 = x_1^2\left(\frac{1}{2n+2}x_1^{2n} - 1\right)$$

is radially unbounded, as a function of x_1 alone, and the function $x_1^2 - x_1x_2 + \frac{1}{2}x_2^2$ is radially unbounded as a function of x_1 and x_2. As a result the global minimum of f is also a local minimum. Note that (recall again that n is a positive integer)

$$f(P_b) = f(P_c) = -\frac{1}{2}\frac{n}{n+1} < 0,$$

hence both P_b and P_c are global minimizers.

d) The point P_0 coincides with the saddle point P_a. The function f along the direction d is given by

$$\phi(\alpha) = f(\alpha, \alpha) = \frac{1}{2n+2}\alpha^{2n+2} - \frac{1}{2}\alpha^2.$$

Note that $\phi(0) = 0$ and that $\phi(\alpha) < 0$ for $\alpha > 0$ and sufficiently small (namely for all $\alpha \in \left(0, (n+1)^{\frac{1}{2n}}\right)$, hence d is a descent direction for f at P_0.

(Note that $\phi(\alpha)$ is negative also for $\alpha \in \left(-(n+1)^{\frac{1}{2n}}, 0\right)$, i.e. $-d$ is also a descent direction for f at P_0.)

Exercise 9 Newton's method for the minimization of a function $f : I\!R \to I\!R$ is based on a quadratic approximation of the function at a given point. An alternative way to construct a quadratic approximation that does not require the computation of the second derivative is to consider an approximation based on the knowledge of two points x_k and x_{k-1} and of the values $f(x_k)$, $\frac{df(x_k)}{dx}$ and $\frac{df(x_{k-1})}{dx}$. Such an approximation is given by

$$q(x) = f(x_k) + \frac{df(x_k)}{dx}(x - x_k) + \frac{\dfrac{df(x_{k-1})}{dx} - \dfrac{df(x_k)}{dx}}{x_{k-1} - x_k}\frac{(x - x_k)^2}{2}.$$

a) Show that the function $q(x)$ is such that

$$q(x_k) = f(x_k), \qquad \frac{dq(x_k)}{dx} = \frac{df(x_k)}{dx}, \qquad \frac{dq(x_{k-1})}{dx} = \frac{df(x_{k-1})}{dx}.$$

b) Compute the stationary point $x_\star$ of $q(x)$.

c) Consider the algorithm, known as the method of the false position, obtained by setting $x_{k+1} = x_\star$, with $x_\star$ as in part b), and argue that this algorithm provides an approximation of Newton's method that does not require the computation of the second derivative of f.

d) Show that the method of the false position applied to the minimization of a quadratic function $f = ax^2 + bx + c$, with $a > 0$, coincides with Newton's method.

e) Consider the function $f = \frac{x^4}{4} + x$. This function has a global minimizer at $x = -1$.

i) Show that the method of the false position yields the iteration

$$x_{k+1} = x_k - (x_k^3 + 1)\frac{1}{x_{k-1}^2 + x_{k-1}x_k + x_k^2}.$$

ii) Evaluate

$$\frac{|\mathcal{E}_{k+1}|}{\mathcal{E}_k^2} = \frac{|x_{k+1} + 1|}{(x_k + 1)^2}$$

and show that if $\lim_{k\to\infty} x_k = -1$ then

$$\lim_{k\to\infty} \frac{|\mathcal{E}_{k+1}|}{\mathcal{E}_k^2} = 1.$$

Hence, quantify the speed of convergence of the method.

Solution 9

a) Setting $x = x_k$ in $q(x)$ yields $q(x_k) = f(x_k)$. Note that

$$\frac{dq(x)}{dx} = \frac{df(x_k)}{dx} + \frac{\frac{df(x_{k-1})}{dx} - \frac{df(x_k)}{dx}}{x_{k-1} - x_k}(x - x_k)$$

hence, setting $x = x_k$ and $x = x_{k-1}$ yields

$$\frac{dq(x_k)}{dx} = \frac{df(x_k)}{dx}, \qquad \frac{dq(x_{k-1})}{dx} = \frac{df(x_{k-1})}{dx}.$$

b) The stationary point $x_\star$ of $q(x)$ is obtained by solving the equation

$$\frac{dq(x)}{dx} = 0,$$

which yields

$$x_\star = x_k - \left(\frac{\frac{df(x_{k-1})}{dx} - \frac{df(x_k)}{dx}}{x_{k-1} - x_k}\right)^{-1} \frac{df(x_k)}{dx}.$$

c) The method of the false position is therefore given by

$$x_{k+1} = x_k - \left(\frac{\frac{df(x_{k-1})}{dx} - \frac{df(x_k)}{dx}}{x_{k-1} - x_k}\right)^{-1} \frac{df(x_k)}{dx}.$$

This algorithm is an approximation of Newton's method because the quantity

$$\frac{\frac{df(x_{k-1})}{dx} - \frac{df(x_k)}{dx}}{x_{k-1} - x_k}$$

is an approximation of $\frac{d^2 f(x)}{dx^2}$ at $x = x_k$. Note however that, unlike Newton's method, the method of the false position does not need the computation of the second derivative: it uses an approximation.

d) For quadratic functions one has

$$\frac{d^2 f(x)}{dx^2} = 2a$$

and

$$\frac{\dfrac{df(x_{k-1})}{dx} - \dfrac{df(x_k)}{dx}}{x_{k-1} - x_k} = \frac{(2ax_{k-1} + b) - (2ax_k + b)}{x_{k-1} - x_k} = 2a,$$

hence, for such functions, Newton's method and the method of the false position coincide.

e) If $f = \dfrac{x^4}{4} + x$ then $\dfrac{df(x)}{dx} = x^3 + 1$, and replacing in the expression of the considered method yields

$$x_{k+1} = x_k - \frac{x_{k-1} - x_k}{(x_{k-1}^3 + 1) - (x_k^3 + 1)}(x_k^3 + 1) = x_k - \frac{x_{k-1} - x_k}{x_{k-1}^3 - x_k^3}(x_k^3 + 1).$$

and, recalling that

$$x_{k-1}^3 - x_k^3 = (x_{k-1} - x_k)(x_{k-1}^2 + x_{k-1}x_k + x_k^2),$$

$$x_{k+1} = x_k - (x_k^3 + 1)\frac{1}{x_{k-1}^2 + x_{k-1}x_k + x_k^2}.$$

Note that

$$\begin{aligned} x_{k+1} + 1 = &\quad x_k + 1 - (x_k^3 + 1)\frac{1}{x_{k-1}^2 + x_{k-1}x_k + x_k^2} \\ = &\quad (x_k + 1)(x_{k-1} + 1)\frac{x_k + x_{k-1} - 1}{x_{k-1}^2 + x_{k-1}x_k + x_k^2}, \end{aligned}$$

hence

$$\frac{|\mathcal{E}_{k+1}|}{\mathcal{E}_k^2} = \left| \frac{x_{k-1} + 1}{x_k + 1} \frac{x_k + x_{k-1} - 1}{x_{k-1}^2 + x_{k-1}x_k + x_k^2} \right|.$$

If $\lim_{k\to\infty} x_k = -1$ then also $\lim_{k\to\infty} x_{k-1} = -1$, hence $\dfrac{|\mathcal{E}_{k+1}|}{\mathcal{E}_k^2} = 1$, which shows that the algorithm has quadratic speed of convergence (if it converges).

Exercise 10 Consider the problem of minimizing the function

$$f(x_1, x_2, \cdots, x_n, y) = \frac{1}{4}x_1^4 + \frac{1}{4}x_2^4 + \cdots + \frac{1}{4}x_n^4 - (x_1 + x_2 + \cdots + x_n)y + \frac{n}{2}y^2,$$

where n is a positive integer.

a) Compute all stationary points of the function.

b) Using second order sufficient conditions *classify* the stationary points determined in part a), *i.e.* say which is a local minimizer, or a local maximizer, or a saddle point.

c) Show that the function f is radially unbounded and hence compute the global minimum of f. Is the global minimizer unique?

d) Consider the points $P_p = (1, 1, \cdots, 1, 1)$ and $P_m = (-1, -1, \cdots, -1, -1)$ and the direction d from P_p to P_m. Show that this is an ascent direction for f at P_p.

Solution 10

a) The stationary points of the function f are computed by solving the equations

$$0 = \nabla f = \begin{bmatrix} x_1^3 - y \\ x_2^3 - y \\ \vdots \\ x_n^3 - y \\ -x_1 - x_2 - \cdots - x_n + ny \end{bmatrix}.$$

The first n equations yield $x_i = y^{1/3}$, hence the last equation becomes

$$0 = -ny^{1/3} + ny = n(y - y^{1/3}).$$

The solutions of this equation are $y = 0$, $y = 1$ and $y = -1$. In summary, the function f has three stationary points

$$P_a = (0, \cdots, 0, 0) \qquad P_b = (1, \cdots, 1, 1) \qquad P_c = (-1, \cdots, -1, -1).$$

b) Note that

$$\nabla^2 f = \begin{bmatrix} 3x_1^2 & 0 & \cdots & 0 & -1 \\ 0 & 3x_2^2 & \cdots & 0 & -1 \\ 0 & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 3x_n^2 & -1 \\ -1 & -1 & \cdots & -1 & n \end{bmatrix}.$$

Hence

$$\nabla^2 f(P_a) = \begin{bmatrix} 0 & 0 & \cdots & 0 & -1 \\ 0 & 0 & \cdots & 0 & -1 \\ 0 & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & -1 \\ -1 & -1 & \cdots & -1 & n \end{bmatrix},$$

which is an indefinite matrix, hence P_a is a saddle point. Finally,

$$\nabla^2 f(P_b) = \nabla^2 f(P_c) = \begin{bmatrix} 3I & -v \\ -v' & n \end{bmatrix},$$

where $v' = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}$. Exploiting the relation

$$\begin{bmatrix} I & 0 \\ v'/3 & 1 \end{bmatrix} \begin{bmatrix} 3I & -v \\ -v' & n \end{bmatrix} \begin{bmatrix} I & v/3 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3I & 0 \\ 0 & 2/3n \end{bmatrix},$$

we conclude that P_b and P_c are local minimizers.

c) The function f can be written as

$$f = \frac{1}{4}(x_1^2 - 1)^2 + \cdots + \frac{1}{4}(x_n^2 - 1)^2 + \frac{1}{2}(x_1 - y)^2 + \cdots + \frac{1}{2}(x_n - y)^2 - \frac{n}{4}.$$

Hence $f + n/4$ is a *sum of squares*, and all variables x_1, x_2, $\cdots$, x_n, y are present in one of the squares. As a result the function is radially unbounded and the local minimum of f is also a global minimum. Note that

$$f(P_b) = f(P_c) = -\frac{n}{4} < 0,$$

hence both P_b and P_c are global minimizers.

d) The direction from P_p to P_m is

$$d = P_m - P_p = -2 \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix}.$$

The function f along the direction d at P_p is given by

$$\phi(\alpha) = f(1 - 2\alpha, \cdots, 1 - 2\alpha, 1 - 2\alpha) = \frac{n}{4}(1 - 2\alpha)^4 - \frac{n}{2}(1 - 2\alpha^2) = -\frac{n}{4} + 4n\alpha^2 + \cdots$$

Note that $\phi(0) = -n/4$ and that $\phi(\alpha) > -n/4$ for $\alpha > 0$ and sufficiently small, hence d is an ascent direction for f at P_p.

Exercise 11 The problem of minimizing a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ can be solved with the so-called heavy ball algorithm, which is a modification of the gradient algorithm, and it is described (in its simplest form) by the iteration

$$x_{k+1} = x_k - \alpha \nabla f(x_k) + \beta(x_k - x_{k-1}),$$

where $\alpha > 0$ is a constant and $\beta \in [0, 1)$ is the *heavy ball* parameter.

a) Assume $x_{-1} = x_0$. Show that the iteration of the heavy ball algorithm can be written as

$$x_{k+1} = x_k - \alpha \left(\nabla f(x_k) + \beta \nabla f(x_{k-1}) + \beta^2 \nabla f(x_{k-2}) + \cdots + \beta^k \nabla f(x_0) \right),$$

for $k \geq 1$.

b) Consider the function

$$f(x_1, x_2) = 2x_1^2 + \frac{1}{2}x_2^2,$$

which has a unique (global) minimizer at $x_1 = x_2 = 0$.

i) Consider the heavy ball algorithm with $\beta = 0$, *i.e.* the gradient algorithm with a constant line search parameter α. Show that the sequence $\{x_k\} = \{(x_{k,1}, x_{k,2})\}$ generated by this algorithm converges to the minimizer of f if and only if $0 < \alpha < \frac{1}{2}$. Select $\alpha = 1/4$. Show that $x_{k,1} = 0$, for all $k \geq 1$, and determine the speed of convergence of the sequence $\{x_k\}$.

ii) Consider the heavy ball algorithm described above with $x_{-1} = x_0$, $\alpha = 1/4$ and $\beta = 3/4$. Show that the sequence $\{x_k\} = \{(x_{k,1}, x_{k,2})\}$ generated by this algorithm is such that $x_{k,1} = 0$ for all $k \geq 1$. Evaluate $x_{k,2}$ for $k = 1, \cdots, 4$. Estimate the speed of convergence of the sequence $\{x_k\}$.

Solution 11

a) Setting $x_{-1} = x_0$ yields

$$\begin{aligned}
k = 0 \quad &\Rightarrow \quad x_1 = x_0 - \alpha \nabla f(x_0), \\
k = 1 \quad &\Rightarrow \quad x_2 = x_1 - \alpha \nabla f(x_1) + \beta(x_1 - x_0) = x_1 - \alpha(\nabla f(x_1) + \beta \nabla f(x_0)), \\
k = 2 \quad &\Rightarrow \quad x_3 = x_2 - \alpha \nabla f(x_2) + \beta(x_2 - x_1) = x_2 - \alpha(\nabla f(x_2) + \beta \nabla f(x_1) + \beta^2 \nabla f(x_0))
\end{aligned}$$

from which we deduce the general expression

$$x_{k+1} = x_k - \alpha \left(\nabla f(x_k) + \beta \nabla f(x_{k-1}) + \beta^2 \nabla f(x_{k-2}) + \cdots + \beta^k \nabla f(x_0) \right).$$

b) i) For the considered function the gradient algorithm with constant α is described by the iteration

$$\begin{aligned}
x_{k+1,1} &= x_{k,1} - \alpha(4x_{k,1}) &= (1 - 4\alpha)x_{k,1}, \\
x_{k+1,2} &= x_{k,2} - \alpha(x_{k,2}) &= (1 - \alpha)x_{k,2}.
\end{aligned}$$

Both sequences $\{x_{k,1}\}$ and $\{x_{k,2}\}$ converge to 0 if, and only if, the conditions

$$-1 < 1 - 4\alpha < 1, \qquad -1 < 1 - \alpha < 1$$

hold simultaneously, which is equivalent to $\alpha \in (0, 1/2)$.
Setting $\alpha = 1/4$ yields

$$x_{k+1,1} = 0, \qquad x_{k+1,2} = \frac{3}{4}x_{k,2},$$

hence $x_{k,1} = 0$, for all $k \geq 1$.
To determine the speed of convergence note that we can consider only the sequence $\{x_{k,2}\}$, which is such that (recall that the sequence converges to 0)

$$\frac{x_{k+1,2}}{x_{k,2}} = \frac{3}{4},$$

which shows linear speed of convergence.

ii) For the considered function and under the stated conditions the heavy ball algorithm is described by the iteration

$$\begin{aligned} x_{k+1,1} &= x_{k,1} - \alpha(4x_{k,1}) + \beta(x_{k,1} - x_{k-1,1}), \\ x_{k+1,2} &= x_{k,2} - \alpha(x_{k,2}) + \beta(x_{k,2} - x_{k-1,2}). \end{aligned}$$

The first of the equations above, the condition $x_{1,0} = x_{1,-1}$, and $\alpha = 1/4$ imply $x_{1,1} = 0$ and $x_{k,1} = 0$, for all $k \geq 1$.
The second of the equations above and the results in part a) yield

$$x_{k+1,2} = x_{k,2} - \frac{1}{4}\left(x_{k,2} + \frac{3}{4}x_{k-1,2} + ...\right).$$

Hence

$$x_{1,2} = \frac{3}{4}x_{0,2},$$

$$x_{2,2} = x_{1,2} - \frac{1}{4}(x_{1,2} + 3/4x_{0,2}) = \frac{1}{2}x_{2,1},$$

$$x_{3,2} = x_{2,2} - \frac{1}{4}(x_{2,2} + \frac{3}{4}x_{1,2} + \frac{9}{16}x_{0,2}) = 0,$$

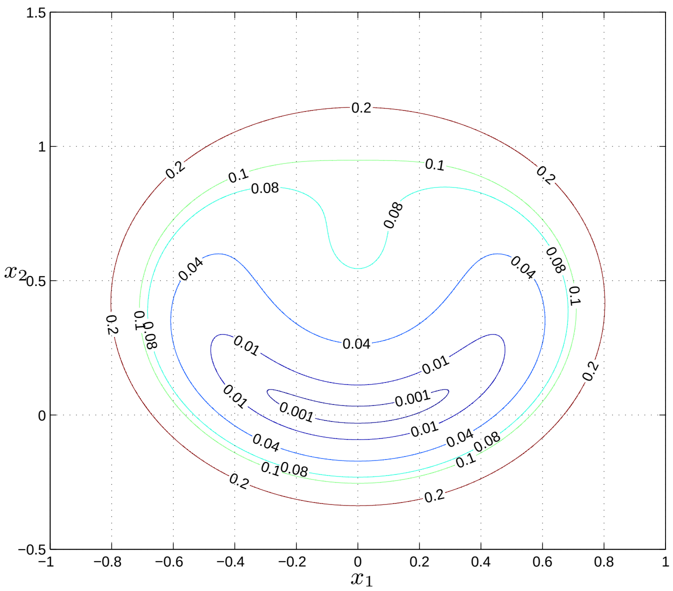
$$x_{4,2} = 0,$$

which shows that the sequence generated by the heavy ball algorithm converges in *finite time.*

Exercise 12 Consider the problem of minimizing the function

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

the level lines of which, for $\delta = \sqrt{32}/3$ are plotted in the figure below.



a) Compute all stationary points of the function as a function of δ.

b) Assume $\delta = \sqrt{32}/3$.

i) Determine the stationary points of the function f, indicate them the figure, and *classify* the stationary points *i.e.* say which is a local minimizer, or a local maximizer, or a saddle point, without computing the Hessian matrix of f.

ii) Determine, from inspection of the figure, a set of points such that the gradient algorithm with exact line search initialized at such points yields a sequence which converges to the global minimizer in one step. Sketch the obtained set on the figure.

iii) Determine, analytically, all points such that the gradient algorithm with exact line search initialized at such points yields a sequence which converges to the global minimizer in one step. Sketch the obtained set on the figure.

Solution 12

a) The stationary points of the function f are computed by solving the equations

$$0 = \nabla f = \begin{bmatrix} 2x_1(2x_1^2 - \delta x_2 + 2x_2^2) \\ 2x_2 - \delta x_1^2 - 3\delta x_2^2 + 4x_2x_1^2 + 4x_2^3 \end{bmatrix}.$$

From the first equation we have $x_1 = 0$ or $x_1^2 = -x_2^2 + \frac{\delta}{2}x_2$. Replacing $x_1 = 0$ in the second equation yields

$$0 = x_2(2 - 3\delta x_2 + 4x_2^2).$$

Replacing $x_1^2 = -x_2^2 + \dfrac{\delta}{2}x_2$ in the second equation yields

$$0 = -\frac{1}{2}x_2(\delta - 2)(\delta + 2).$$

In conclusion the function f has the following stationary points.

- $P_0 = (0, 0)$, for any value of δ.
- $P_1 = (0, \dfrac{3\delta + \sqrt{9\delta^2 - 32}}{8})$ and $P_2 = (0, \dfrac{3\delta - \sqrt{9\delta^2 - 32}}{8})$ if $\delta^2 \geq \frac{32}{9}$. Note that if $\delta = \pm\frac{\sqrt{32}}{3}$ then $P_1 = P_2$.
- If $\delta = \pm 2$ then all points in the set $x_1^2 + x_2^2 - \frac{\delta}{2}x_2 = x_1^2 + x_2^2 \mp x_2 = 0$ are stationary points.

b) Consider now the case in which $\delta = \dfrac{\sqrt{32}}{3}$.

i) The only stationary points are P_0 and $P_1 = P_2 = (0, \frac{\sqrt{2}}{2})$. From the figure we conclude that P_0 is a local minimizer, and $P_1 = P_2$ is a saddle point. (The Hessian matrix is singular at P_0 and P_1, hence it cannot be used to classify these points.)

ii) Note that the gradient of f on the x_2-axis is given by

$$\nabla f(0, x_2) = \begin{bmatrix} 0 \\ x_2(2 - \sqrt{32}x_2 + 4x_2^2) \end{bmatrix}.$$

The gradient of f on the x_2-axis is a direction of ascent which is parallel to the x_2-axis. Therefore, the gradient algorithm with exact line search yields the global minimizer in one step for all initial points on the x_2-axis.

iii) The set of points such that the gradient algorithm with exact line search yields a sequence which converges to the global minimizer in one step is obtained eliminating α, *i.e.* the line search parameter, from the equation

$$0 = x - \alpha\nabla f(x).$$

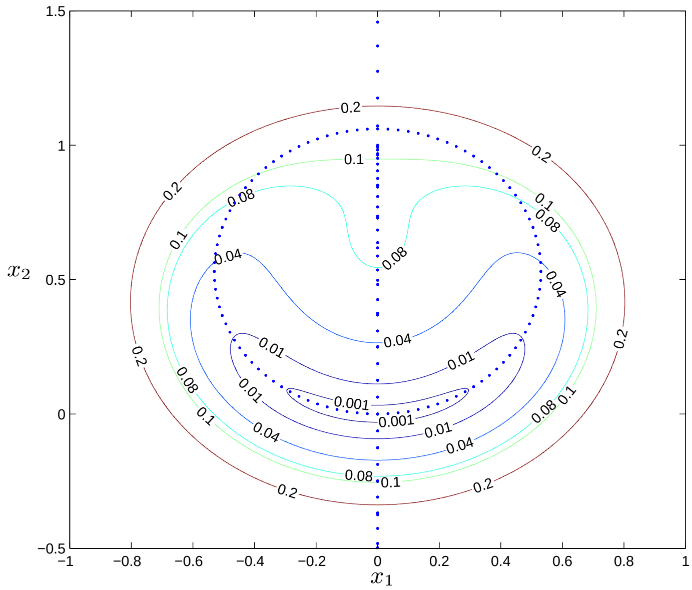
This yields the set of points described by

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

i.e. the x_2-axis and the circle

$$x_1^2 + x_2^2 - \frac{3}{4}\sqrt{2}x_2 = 0,$$

which is a circle centered at $P = (0, \frac{3}{8}\sqrt{2})$ and with radius equal to $\frac{3}{8}\sqrt{2})$. The set of all points with the requested property is indicated on the figure with "dots".



Exercise 13 Consider the problem of minimizing the function

$$f(x_1, x_2) = 4x_1^2 - 2x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 + \frac{1}{4}x_2^2.$$

a) Compute all stationary points of the function.

b) Using second order sufficient conditions *classify* the stationary points determined in part a), *i.e.* say which is a local minimizer, or a local maximizer, or a saddle point.

c) Show that the function f is radially unbounded and hence compute the global minimum of f. Is the global minimizer unique?

d) Using the results of parts a), b) and c) sketch the level lines of the function f.

Solution 13

a) The stationary points of the function f are computed by solving the equations

$$0 = \nabla f = \begin{bmatrix} 8x_1 - 8x_1^3 + 2x_1^5 + x_2 \\ x_1 + \frac{1}{2}x_2 \end{bmatrix}.$$

The second equation yields $x_2 = -2x_1$, which replaced in the first equation yields

$$0 = 2x_1(x_1 - 1)(x_1 + 1)(x_1^2 - 3).$$

As a result, the function f has five stationary points

$$P_1 = (0, 0), \qquad P_2 = (-1, 2), \qquad P_3 = (1, -2), \qquad P_4 = (\sqrt{3}, -2\sqrt{3}), \qquad P_5 = (-\sqrt{3}, 2\sqrt{3}).$$

b) Note that

$$\nabla^2 f = \begin{bmatrix} 8 - 24x_1^2 + 10x_1^4 & 1 \\ 1 & \dfrac{1}{2} \end{bmatrix}.$$

As a result

$$\nabla^2 f(P_1) = \begin{bmatrix} 8 & 1 \\ 1 & \frac{1}{2} \end{bmatrix},$$

which is a positive definite matrix, hence P_1 is a local minimizer;

$$\nabla^2 f(P_2) = \nabla^2 f(P_3) \begin{bmatrix} -6 & 1 \\ 1 & \frac{1}{2} \end{bmatrix},$$

which is an indefinite matrix, hence P_2 and P_3 are saddle points;

$$\nabla^2 f(P_4) = \nabla^2 f(P_5) \begin{bmatrix} 26 & 1 \\ 1 & \frac{1}{2} \end{bmatrix},$$

which is a positive definite matrix, hence P_4 and P_5 are local minimizers.

c) The function f can be written as

$$f = \left(x_1 + \frac{1}{2}x_2\right)^2 + \frac{x_1^2}{3}\left(x_1^2 - 3\right)^2.$$

Hence f is a *sum of squares*, and all variables x_1 and x_2 are present in one of the squares. As a result the function is radially unbounded and the local minimum of f is also a global minimum. Note that

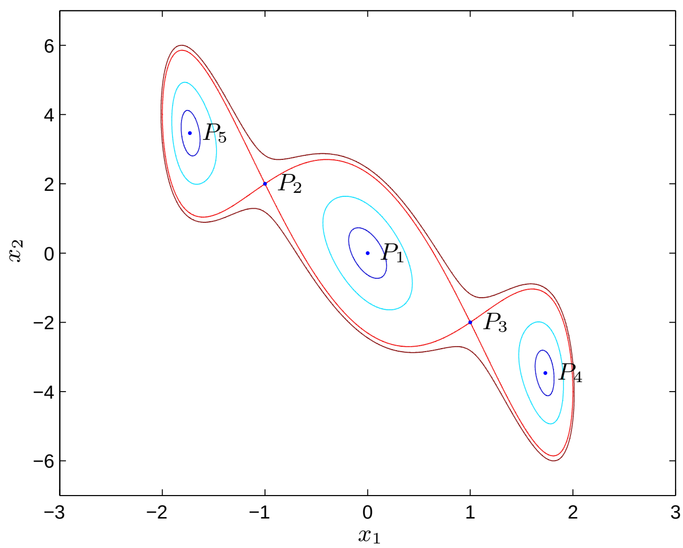
$$f(P_1) = f(P_4) = f(P_5) = 0$$

hence P_1, P_4 and P_5 are all global minimizers.

d) The level lines of f can be sketched using the following considerations.

- Around the minimizers the level lines are closed.
- The value of f at the saddle points P_2 and P_3 is $4/3$. There is a level line that *connects* the saddle points. Close to the saddle points this level line is composed of two curves.

A sketch of the level lines is in the figure below.



Exercise 14 Consider the problem of minimizing the function

$$f(x_1, x_2) = \frac{1}{2}x_1^2\left(\frac{1}{6}x_1^2 + 1\right) + x_2 \arctan x_2 - \frac{1}{2}\ln(x_2^2 + 1).$$

a) Compute the unique stationary point of the function.

b) Using second order sufficient conditions show that the stationary point determined in part a) is a local minimizer.

c) Consider now the minimization of the function using Newton's method.

   i) Write Newton's iteration for the considered problem.

   ii) Show that Newton's direction is a descent direction for f at any point which is not a stationary point.

   iii) Compute four steps of Newton's algorithm from the initial point $(1, 0.5)$. Compute four steps of Newton's algorithm from the initial point $(1, 2)$.

   iv) Discuss why the second sequence computed in part c.iii) does not converge to the global minimizer, despite the fact that Newton's direction is always a descent direction. Propose a simple modification of Newton's iteration that would guarantee global convergence to the minimizer.

Solution 14

a) The stationary points of the function f are computed by solving the equations

$$0 = \nabla f = \begin{bmatrix} \frac{1}{3}x_1^3 + x_1 \\ \arctan x_2 \end{bmatrix}.$$

These equations have the unique solution $x_1 = x_2 = 0$, which is therefore the unique stationary point of f.

b) Note that

$$\nabla^2 f = \begin{bmatrix} x_1^2 + 1 & 0 \\ 0 & \dfrac{1}{1 + x_2^2} \end{bmatrix}.$$

Hence $\nabla^2 f(0, 0) = \mathrm{diag}(1, 1)$, which is a positive definite matrix. The stationary point is a local minimizer.

c) i) Newton's iteration is

$$x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1}\nabla f(x_k)$$

hence

$$x_{k+1,1} = \frac{2x_{k,1}^3}{3(x_{k,1}^2 + 1)}, \qquad x_{k+1,2} = x_{k,2} - (1 + x_{k,2}^2)\arctan x_{k,2}.$$

   ii) Newton's direction is

$$d = -[\nabla^2 f(x)]^{-1}\nabla f(x).$$

   Note that

$$\nabla f'\ d = -\nabla f'[\nabla^2 f(x)]^{-1}\nabla f(x) < 0,$$

   for all points such that $\nabla f(x) \neq 0$, since $\nabla^2 f$ is positive definite. As a result, d is a descent direction for f for all $x \neq 0$.

   iii) A direct computation yields

$$x_0 = (1, 1/2), \qquad x_1 = (1/3, -0.079), \qquad x_2 = (0.022, 0.00033),$$
$$x_3 = (0.000007, -2.5\ 10^{-11}), \qquad x_4 = (2.6\ 10^{-16}, 0),$$

   and

$$x_0 = (1, 2), \qquad x_1 = (1/3, -3.53), \qquad x_2 = (0.022, 13.95),$$
$$x_3 = (0.000007, -279.34), \qquad x_4 = (2.6\ 10^{-16}, 1.2\ 10^5).$$

iv) The second sequence does not converge since Newton's method guarantee only local convergence properties. To achieve global convergence, since Newton's direction is a descent direction for f at any $x \neq 0$, it is enough to introduce a line search parameter, *i.e.* to consider the iteration

$$x_{k+1} = x_k - \alpha[\nabla^2 f(x_k)]^{-1} \nabla f(x_k),$$

with $\alpha > 0$, and determined using a line search algorithm.

Exercise 15 Consider the problem of minimizing a function of n variables x_1, x_2, $\cdots$, x_n, defined as

$$f(x_1, \cdots, x_n) = f_1(x_1)\ f_2(x_2)\ \cdots\ f_n(x_n),$$

that is the function f is the product of the n functions f_i, each of the variable x_i only.

a) Assume that all functions f_i are such that

$$f_i(x_i) > 0$$

for all x_i and that there exist unique $x_i^\star$ such that $x_i^\star$ is a stationary point of f_i.

i) Compute the stationary point $x^\star$ of the function f.

ii) Using second order sufficient conditions show that the stationary point $x^\star$ of the function f is a strict local minimizer if and only if all $x_i^\star$ are strict local minimizers of the functions f_i.

b) Assume $n = 3$, that is consider the function

$$f(x_1, x_2, x_3) = f_1(x_1)\ f_2(x_2)\ f_3(x_3).$$

Assume that the functions f_i do not have stationary points but that there exists, for $i = 1, 2, 3$, a unique point x_i° such that

$$f_i(x_i^\circ) = 0$$

and

$$f_i(x_i) \neq 0$$

for all $x_i \neq x_i^\circ$.

i) Compute all stationary points of the function f.

ii) Show that the Hessian matrix of f at any stationary point is either identically zero or it has positive and negative eigenvalues. Hence argue that none of the stationary point can be a strict local minimizer.
(Hint: recall that a symmetric matrix has real eigenvalues and that the trace of a matrix, that is the sum of its diagonal entries, is equal to the sum of its eigenvalues.)

Solution 15

a) Consider the function f.

i) The stationary points of the function f are computed by solving the equations

$$0 = \nabla f = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}\ f_2\ \cdots\ f_n \\ f_1 \dfrac{\partial f_2}{\partial x_2}\ f_3\ \cdots\ f_n \\ \vdots \\ f_1\ f_2\ \cdots\ \dfrac{\partial f_n}{\partial x_n} \end{bmatrix}.$$

Since all f_i's are positive, and have a unique stationary point, the only stationary point of f is the point

$$x^\star = (x_1^\star, x_2^\star, \cdots, x_n^\star).$$

ii) Note that, for $i \neq j$,

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial f_i}{\partial x_i}\frac{\partial f_j}{\partial x_j} M$$

where M is a positive function, hence

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x^\star) = 0.$$

As a result, the Hessian matrix of the function f at $x^\star$ is

$$\nabla^2 f(x^\star) = \mathrm{diag}\left(\frac{\partial^2 f_1}{\partial x_1^2}(x_1^\star)\ f_2(x_2^\star)\ \cdots\ f_n(x_n^\star),\ \cdots,\ f_1(x_1^\star)\ f_2(x_2^\star)\ \cdots\ \frac{\partial^2 f_n}{\partial x_n^2}(x_n^\star) \right).$$

This implies that the function f has a strict local minimizer at $x^\star$ if and only if all functions f_i have a strict local minimizer at $x_i^\star$.

b) Consider the function f with $n = 3$.

i) The stationary points of the functions f are the solution of the equations

$$0 = \nabla f = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} f_2 f_3 \\ f_1 \frac{\partial f_2}{\partial x_2} f_3 \\ f_1 f_2 \frac{\partial f_3}{\partial x_3} \end{bmatrix}.$$

These equations admit infinitely many solutions given by

$$x_{12}^\circ = (x_1^\circ, x_2^\circ, \bar{x}_3), \qquad x_{13}^\circ = (x_1^\circ, \bar{x}_2, x_3^\circ), \qquad x_{23}^\circ = (\bar{x}_1, x_2^\circ, x_3^\circ),$$

where $\bar{x}_1$, $\bar{x}_2$ and $\bar{x}_3$ are arbitrary values.

ii) The Hessian matrix of f is

$$\nabla^2 f = \begin{bmatrix} \frac{\partial^2 f_1}{\partial x_1^2}\ f_2\ f_3 & \frac{\partial f_1}{\partial x_1}\ \frac{\partial f_2}{\partial x_2}\ f_3 & \frac{\partial f_1}{\partial x_1}\ \frac{\partial f_3}{\partial x_3}\ f_2 \\ \frac{\partial f_1}{\partial x_1}\ \frac{\partial f_2}{\partial x_2}\ f_3 & \frac{\partial^2 f_2}{\partial x_2^2}\ f_1\ f_3 & \frac{\partial f_2}{\partial x_2}\ \frac{\partial f_3}{\partial x_3}\ f_1 \\ \frac{\partial f_1}{\partial x_1}\ \frac{\partial f_3}{\partial x_3}\ f_2 & \frac{\partial f_2}{\partial x_2}\ \frac{\partial f_3}{\partial x_3}\ f_1 & \frac{\partial^2 f_3}{\partial x_3^2}\ f_1\ f_2 \end{bmatrix}.$$

Hence

$$\nabla^2 f(x_{12}^\circ) = \begin{bmatrix} 0 & \alpha & 0 \\ \alpha & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where

$$\alpha = \frac{\partial f_1}{\partial x_1}(x_1^\circ)\ \frac{\partial f_2}{\partial x_2}(x_2^\circ)\ f_3(\bar{x}_3).$$

The function α is zero for $\bar{x}_3 = x_3^\circ$ and it is non-zero otherwise. Hence, $\nabla^2 f(x_{12}^\circ)$ is either identically zero or has trace zero, which means that it has a positive and a negative eigenvalue. In both cases, the points x_{12}° cannot be local strict minimizers.
Similar considerations apply to x_{13}° and x_{23}°.

Exercise 16 An alternative way to introduce Newton's method for the solution of a nonlinear equation is to consider the evaluation of the integral

$$f(x) = f(x_k) + \int_{x_k}^{x} \dot{f}(t)dt,$$

where $\dot{f}$ denotes the derivative of the function f, by means of the so-called Newton-Cotes quadrature formula of order zero (the rectangular rule) yielding

$$f(x) \approx f(x_k) + (x - x_k)\dot{f}(x_k),$$

setting $x = x_{k+1}$ and replacing the $\approx$ sign with an $=$ sign, thus yielding

$$f(x_{k+1}) = f(x_k) + (x_{k+1} - x_k)\dot{f}(x_k),$$

and setting $f(x_{k+1}) = 0$, thus obtaining the iteration

$$x_{k+1} = x_k - \frac{f(x_k)}{\dot{f}(x_k)}.$$

a) Consider the evaluation of the integral by means of the Newton-Cotes quadrature formula of order one (the trapezoidal rule), that is

$$\int_{x_k}^{x} \dot{f}(t)dt \approx \frac{x - x_k}{2}\left(\dot{f}(x_k) + \dot{f}(x)\right).$$

   i) Determine a new iteration for the solution of the nonlinear equation $f(x) = 0$. Note that the obtained iteration, which is a modified Newton's iteration, is implicitly defined, that is x_{k+1} is a function of x_k and of $\dot{f}(x_{k+1})$.

   ii) An explicit iteration can be obtained replacing $\dot{f}(x_{k+1})$ with $\dot{f}(x^\star)$, where

$$x^\star = x_k - \frac{f(x_k)}{\dot{f}(x_k)}.$$

   Write the expression of the resulting modified Newton's iteration.

b) Consider the problem of determining the square root of 2.

   i) Write Newton's iteration for the solution of this problem. Let $x_0 = 1$ and apply three steps of Newton's iteration, that is compute the values x_1, x_2, and x_3 resulting from the application of Newton's iteration with the given initial point. Evaluate the absolute error $e_k = |\sqrt{2} - x_k|$.

   ii) Write the modified Newton's iteration for the solution of this problem. Let $x_0 = 1$ and apply three steps of the modified Newton's iteration, that is compute the values x_1, x_2, and x_3 resulting from the application of the modified Newton's iteration with the given initial point. Evaluate the absolute error $e_k = |\sqrt{2} - x_k|$.

   iii) Compare the Newton's iteration and the modified Newton's iteration in terms of convergence speed and computational complexity.

Solution 16

a) i) Consider the relation

$$f(x) = f(x_k) + \frac{x - x_k}{2}(\dot{f}(x) + \dot{f}(x_k)).$$

   Setting $x = x_{k+1}$ and $f(x) = 0$ yields

$$0 = f(x_k) + \frac{x_{k+1} - x_k}{2}(\dot{f}(x_{k+1}) + \dot{f}(x_k)),$$

   hence solving for x_{k+1} provides the iteration

$$x_{k+1} = x_k - 2\frac{f(x_k)}{\dot{f}(x_{k+1}) + \dot{f}(x_k)}$$

   ii) The modified Newton's iteration is

$$x_{k+1} = x_k - 2\frac{f(x_k)}{\dot{f}(x_k - f(x_k)/\dot{f}(x_k)) + \dot{f}(x_k)}.$$

b) To determine the square root of 2 consider the equation $x^2 - 2 = 0$.

i) Newton's iteration is given by

$$x_{k+1} = x_k - \frac{1}{2}\frac{x_k^2 - 2}{x_k}.$$

The sequence generated by Newton's iteration is

$$x_0 = 1, \qquad x_1 = 1.5, \qquad x_2 = 1.416666667, \qquad x_3 = 1.414215686,$$

and this yields the sequence of the absolute error

$$e_0 = 0.414213562, \quad e_1 = 0.085786438, \quad e_2 = 0.002453105, \quad e_3 = 0.000002124.$$

ii) The modified Newton's iteration is

$$x_{k+1} = x_k - \frac{2(x_k^2 - 2)x_k}{3x_k^2 + 2}.$$

The sequence generated by the modified Newton's iteration is

$$x_0 = 1, \qquad x_1 = 1.4, \qquad x_2 = 1.414213198, \qquad x_3 = 1.414213563,$$

and this yields the sequence of the absolute error

$$e_0 = 0.414213562, \quad e_1 = 0.014213562, \quad e_2 = 3.64 \times 10^{-7}, \quad e_3 = 1 \times 10^{-9}.$$

iii) The modified Newton's iteration is much faster (this is a general conclusion) and has similar complexity than the (classical) Newton's iteration.

Exercise 17 Consider the function

$$f(x) = x_1^4 + x_1 x_2 + \frac{1}{2}x_2^2.$$

a) Compute the stationary points of the function.

b) Using second order sufficient conditions classify the stationary points determined in part a), that is say which is a local minimizer, or a local mazimizer, or a saddle point.

c) Sketch on the (x_1, x_2)-plane the level lines of the function f.

d) Consider the point $P_0 = (0, 0)$.

i) Determine a direction d_0 which is a descent direction for f at P_0.

ii) Consider the problem of performing an exact line search along the direction d_0 starting from P_0. Determine a solution to such a problem.

Solution 17

a) The stationary points of the function f are computed by solving the equations

$$0 = \nabla f = \begin{bmatrix} 4x_1^3 + x_2 \\ x_1 + x_2 \end{bmatrix}.$$

Replacing the second equation in the first yields $x_1(4x_1^2 - 1) = 0$. Hence, the stationary points are

$$P_1 = (0, 0), \qquad P_2 = \left(\frac{1}{2}, -\frac{1}{2}\right), \qquad P_2 = \left(-\frac{1}{2}, \frac{1}{2}\right).$$

b) The Hessian matrix of the function f is

$$\nabla^2 f(x) = \begin{bmatrix} 12x_1^2 & 1 \\ 1 & 1 \end{bmatrix}.$$

Note that

$$\nabla^2 f(P_1) = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

is indefinite and

$$\nabla^2 f(P_2) = \nabla^2 f(P_3) = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$$

is positive definite. Hence P_2 and P_3 are local minimizers, and P_1 is a saddle point.

c) The level lines of f can be sketched using the following considerations.

- Around the minimizers the level lines are closed.
- The value of f at the saddle point P_1 is 0.
- The value of f at the local minimizers P_2 and P_3 is $-1/16$. There is a closed level line which goes through the saddle point and encircles both local minimizers.

A sketch of the level lines is in the figure below.

d) Note that $\nabla f(P_0) = 0$, hence for any direction d the scalar product $\nabla' f \ d$ is zero, *i.e.* it is not possible to use first order sufficient conditions to establish if a direction is a descent direction.

i) Let, for example, $d_0 = [1, -1]'$ and consider the restriction of the function f along d_0, with *initial point* P_0, namely

$$f(P_0 + \alpha d_0) = \alpha^2 \left(-\frac{1}{2} + \alpha^2 \right).$$

For any $\alpha > 0$ and sufficiently small (namely $\alpha \in (0, 1/\sqrt{2})$)

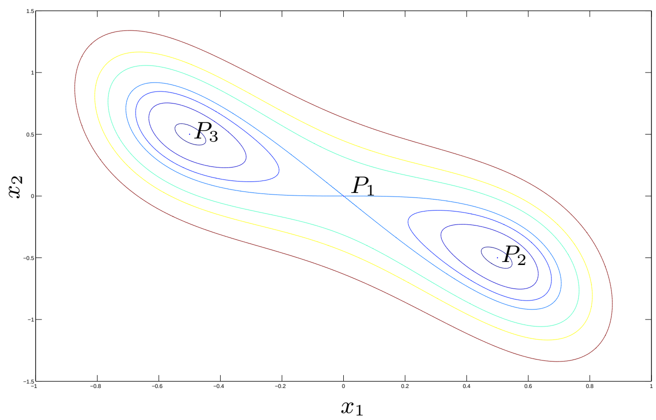
$$f(P_0) > f(P_0 + \alpha d_0),$$

hence d_0 is a descent direction for f at P_0.

ii) To solve an exact line search problem along d_0 at P_0 one has to find the global minimizer, if it exists, of $f(P_0 + \alpha d_0)$. Note that the function

$$f(P_0 + \alpha d_0) = \alpha^2(-1/2 + \alpha^2)$$

is radially unbounded (and bounded from below), hence possesses a global minimizer, which is a stationary point. The stationary points of this function are $\alpha = 0$ (local maximizer) and $\alpha = \pm 1/2$ (local minimizer). Hence, an exact line search along d_0, starting at P_0, gives either the point P_2 or the point P_3.



Exercise 18 Consider the function

$$f(x_1, x_2) = \sin(x_1^2 + x_2^2).$$

a) Sketch on the (x_1, x_2)-plane the level lines of the function f.

b) Compute the stationary points of the function.

c) Explain why second order sufficient conditions of optimality are inadequate to classify some of the stationary points of the functions.

d) Consider the change of variable

$$x_1 = \rho \cos\theta, \qquad x_2 = \rho \sin\theta,$$

with $\rho \geq 0$ and $\theta \in (-\pi, \pi]$.

   i) Rewrite the function f in the new variables. Note that the function depends only upon the variable ρ.

   ii) Compute the stationary points of the function f as a function of ρ and classify these stationary points.

   iii) Exploiting the results in part d.ii) classify the stationary points of the function f.

Solution 18

a) Note that the function is constant on any circle centered at the origin, *i.e.* on any set of the form $x_1^2 + x_2^2 = R^2$. A sketch of the level lines is therefore as in the figure below.

b) The stationary points of the function f are computed by solving the equations

$$0 = \nabla f = \begin{bmatrix} 2x_1 \cos(x_1^2 + x_2^2) \\ 2x_2 \cos(x_1^2 + x_2^2) \end{bmatrix}.$$

Hence, the point $(0, 0)$ is a stationary point and all points such that

$$x_1^2 + x_2^2 = \frac{\pi}{2} + k\pi,$$

with k integer, are stationary points.

c) The Hessian matrix of the function f is

$$\nabla^2 f(x) = 2\cos(x_1^2 + x_2^2)I - 4\sin(x_1^2 + x_2^2)\begin{bmatrix} x_1^2 & x_1x_2 \\ x_1x_2 & x_2^2 \end{bmatrix}.$$

Note that

$$\nabla^2 f(0) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

is positive definite, hence the point $(0, 0)$ is a local minimizer. To classify the stationary points such that $x_1^2 + x_2^2 = \frac{\pi}{2} + k\pi$, note that at such points P_k

$$\nabla^2 f(P_k) = \mp 4 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \end{bmatrix},$$

hence $\nabla^2 f(P_k)$ is singular, and this does not enable the use of second order sufficient conditions of optimality (which require the Hessian to be non-singular).

d) i) The function f in the new variables is given by

$$f(\rho, \theta) = \sin\rho^2,$$

hence it is a function of ρ only.

ii) The stationary points of the function $\sin \rho^2$ are all points such that

$$\frac{df}{d\rho} = 2\rho \cos \rho^2 = 0.$$

These are given by

$$\rho = 0 \qquad\qquad \rho^2 = \frac{\pi}{2} + k\pi,$$

with k any non-negative integer.

iii) Note that

$$\frac{d^2 f}{d\rho^2} = 2\cos\rho^2 - 4\rho^2 \sin\rho^2,$$

hence the point $\rho = 0$ is a local minimizer, the points

$$\rho^2 = \frac{\pi}{2} + 2k\pi,$$

with k any non-negative integer, are local maximizers, and the points

$$\rho^2 = \frac{\pi}{2} + (2k+1)\pi,$$

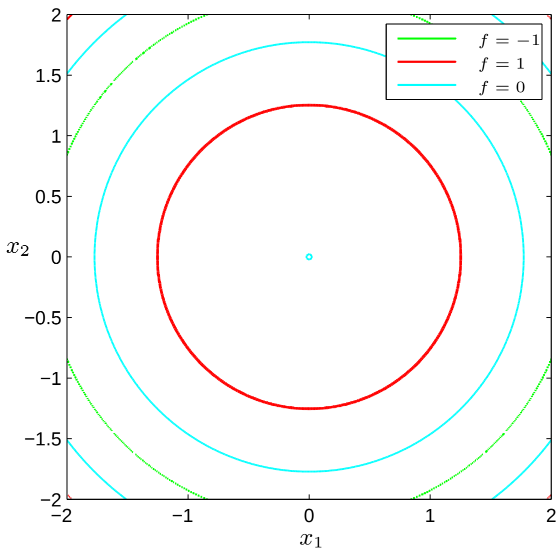
with k any non-negative integer, are local minimizers. This implies that the point $(x_1, x_2) = (0, 0)$ is a local strict minimizer, the points (x_1, x_2) such that

$$x_1^2 + x_2^2 = \frac{\pi}{2} + 2k\pi,$$

with k any non-negative integer, are local non-strict maximizers, and the points (x_1, x_2) such that

$$x_1^2 + x_2^2 = \frac{\pi}{2} + (2k+1)\pi,$$

with k any non-negative integer, are local non-strict minimizers.



Exercise 19 A nonlinear least-squares problem is an unconstrained optimization problem of the form

$$\min_x \frac{1}{2} \sum_{i=1}^{m} r_i^2(x),$$

where $x \in I\!R^n$. The functions $r_1, r_2, \cdots, r_m$ are called residuals and the objective function can be rewritten as $\frac{1}{2}r'(x)r(x)$, with

$$r(x) = \begin{bmatrix} r_1(x) \\ \cdots \\ r_m(x) \end{bmatrix}.$$

a) Write Newton's iteration for the solution of the considered least-square problem.

b) Gauss-Newton's iteration for the solution of the considered least-square problem is given by

$$x_{(k+1)} = x_{(k)} - [J'(x_{(k)})J(x_{(k)})]^{-1}J'(x_{(k)})r(x_{(k)}),$$

where

$$J(x) = \begin{bmatrix} \frac{\partial r_1}{\partial x_1} & \cdots & \frac{\partial r_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial r_m}{\partial x_1} & \cdots & \frac{\partial r_m}{\partial x_n} \end{bmatrix}$$

and $x_{(k)} = \begin{bmatrix} x_{k,1}, & \cdots & x_{k,n} \end{bmatrix}'$. Discuss the differences between Newton's iteration and Gauss-Newton's iteration.
(Hint: consider the *difference* between the Hessian of $\frac{1}{2}r'(x)r(x)$ and the matrix $J'(x)J(x)$.)
Discuss under what conditions the Gauss-Newton direction

$$d_{GN} = -[J'(x)J(x)]^{-1}J'(x)r(x)$$

is a descent direction.

c) Assume $m = 2$, $x = (x_1, x_2)$ and

$$r_1(x) = x_1 + x_2 - x_1x_2 + 2, \qquad r_2(x) = x_1 - e^{x_2}.$$

i) Sketch on the (x_1, x_2)-plane the set of points $r_1(x) = 0$ and $r_2(x) = 0$, hence argue that the considered least-square problem has two (global) solutions. Find an approximation of these global solutions using graphical considerations.

ii) Write explicitly Gauss-Newton's iteration for the considered problem.

iii) Compute three iterations of Gauss-Newton's methods from the initial conditions $(0, 0)$. Evaluate the residuals at $(0, 0)$ and at the last iteration.

iv) Comment on the convergence speed and complexity of Gauss-Newton's method.

Solution 19

a) Newton's method for the minimization of the function

$$f(x) = \frac{1}{2}r'(x)r(x)$$

is described by the iteration

$$x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1}\nabla f(x_k),$$

where

$$\nabla f = \left[\frac{\partial r}{\partial x}\right]' r = J'(x)r$$

(with $J = \dfrac{\partial r}{\partial x}$, as defined above) and

$$\nabla^2 f = \left[\frac{\partial r}{\partial x}\right]' \left[\frac{\partial r}{\partial x}\right] + \sum_{i=1}^{m} r_i \nabla^2 r_i.$$

b) The difference between Newton's method and Gauss-Newton's method is in the matrix that it is inverted. In Newton's method this is the Hessian of the function to be minimized, in Gauss-Newton's method this is one term of the Hessian, which can be computed using only first derivatives. Gauss-Newton's direction is a descent direction if

$$\nabla' f d_{GN} = -r'(x)J(x)[J'(x)J(x)]^{-1}J'(x)r(x) < 0,$$

which holds at all points in which $J(x)$ is full rank and $r(x) \neq 0$.

c) i) The sets $r_1(x) = r_2(x) = 0$ are displayed in the figure below. These sets have two points of intersection, hence the least square problem has only two solutions, which are both global minimizers of the function $\frac{1}{2}r'r$. From the graph one sees that the minimizers are approximately given by the points $(0.1, -2.3)$ and $(5.4, 1.7)$.

ii) Note that

$$J(x) = \begin{bmatrix} 1 - x_2 & 1 - x_1 \\ 1 & -e^{x_2} \end{bmatrix}$$

and

$$d_{GN}(x) = -\frac{1}{(x_1 - 1) + e^{x_2}(x_2 - 1)} \begin{bmatrix} -x_2e^{x_2} + x_1x_2e^{x_2} - 2ex_2x_1 - e^{x_2} + x_1^2 - x_1 \\ -e^{x_2} + x_2e^{x_2} - 2 - x_2 \end{bmatrix}.$$

Hence, Gauss-Newton iteration can be written as

$$x_{k+1} = x_k + d_{GN}(x_k).$$

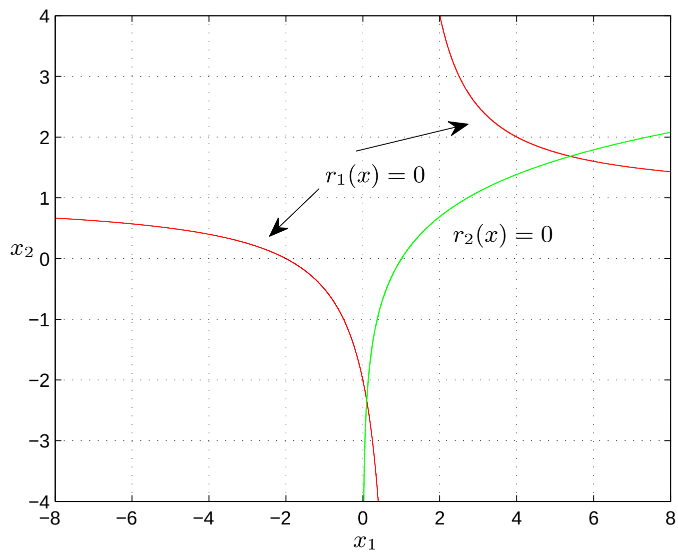
iii) Let $x_{(0)} = (0, 0)$. The residuals at $(0, 0)$ are $r_1(0) = 2$ and $r_2(0) = -1$. The first three elements of the sequence generated by Gauss-Newton's iteration are

$$x_{(1)} = (-0.5, -1.2), \qquad x_{(2)} = (0.0974, -2.1345), \qquad x_{(3)} = (0.096347, -2.319927),$$

and the value of the residuals after three iterations are

$$r_1(x_{(3)}) = -0.000061740, \qquad r_2(_{(3)}) = -0.001933.$$

iv) It is worth noting the fast convergence rate despite the fact that the iteration does not use second derivatives and a line search parameter.



Exercise 20 Consider the function

$$f(x_1, x_2) = \frac{1}{2}x_1^2\left(\frac{1}{6}x_1^2 + 1\right) + x_2 \arctan x_2 - \frac{1}{2}\ln(x_2^2 + 1).$$

a) Compute the unique stationary point of the function.

b) Using second order sufficient conditions of optimality show that the stationary point determined in part a) is a local minimizer. Show, in addition, that the function is convex. Finally, show that the local minimizer is a global minimizer.
(Hint: convexity of a function f is implied by the condition $\nabla^2 f(x) > 0$ for all x.)

c) Consider the problem of minimizing the function f using Newton's method.

   i) Write Newton's iteration for the minimization of the function f.

   ii) Perform 4 steps of Newton's iteration with starting point

$$(x_1, x_2) = (1, 2).$$

d) Consider the function

$$f_2(x_2) = x_2 \arctan x_2 - \frac{1}{2}\ln(x_2^2 + 1).$$

   i) Using the iteration derived in part c.i) write Newton's iteration for the minimization of the function f_2.

   ii) Write the Newton's iteration in part d.i) in the form

$$x_2(k + 1) = \psi(x_2(k)).$$

   Write explicitly the function ψ.

   iii) Plot on the same graph the functions x_2 and $\psi(x_2)$. Exploiting the graph explain why Newton's iteration for the minimization of f_2 converges for initial conditions sufficiently close to zero, and diverges otherwise.
   (Hint: use the graph to *execute* Newton's iteration graphically.)

e) Exploiting the results in part d) and the fact that the function f is the sum of two functions of one variable each, determine (qualitatively) for which initial points the Newton's iteration for the minimization of f converges to the minimizer.

Solution 20

a) The stationary points of the function f are computed by solving the equations

$$0 = \nabla f = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} \\ \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \frac{1}{3}x_1(x_1^2 + 3) \\ \arctan x_2 \end{bmatrix}.$$

Hence, the point $(0, 0)$ is the unique stationary point.

b) The Hessian matrix of the function f is

$$\nabla^2 f(x) = \begin{bmatrix} x_1^2 + 1 & 0 \\ 0 & \frac{1}{1+x_2^2} \end{bmatrix}.$$

Note that

$$\nabla^2 f(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

is positive definite, hence the point $(0, 0)$ is a local minimizer. In addition, $\nabla^2 f > 0$ for all (x_1, x_2), hence the function is convex. For convex function, a stationary point is a global minimizer, hence $(0, 0)$ is a global minimizer.

c) i) Newton's iteration, considering that the function f is the sum of a function of x_1 and of a function of x_2, gives two *decoupled* equations, namely

$$x_1^{k+1} = \frac{2}{3}\frac{x_1^3}{1+x_1^2} \qquad x_2^{k+1} = x_2 - (1+x_2^2)\arctan x_2.$$

ii) The first five elements of the sequences $\{x_1^k\}$ and $\{x_2^k\}$ are

$$x_1^0 = 1,\ x_1^1 = 1/3,\ x_1^2 = 1/45,\ x_1^3 = 1/136755,\ x_1^4 = 1/3836373661058445 \approx 0,$$

and

$$x_2^0 = 2,\ x_2^1 = -3.5357,\ x_2^2 = 13.95095909,\ x_2^3 = -279.3440667,\ x_2^4 = 122016.9990.$$

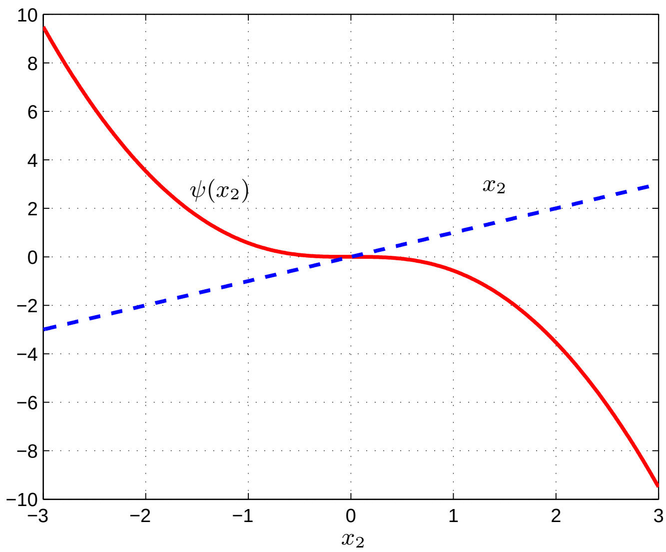
d) i) The iteration is the same as the "x_2" iteration in part c.i).

ii) The function ψ is given by

$$\psi(x_2) = x_2 - (1+x_2^2)\arctan(x_2).$$

iii) The graphs of the considered functions are displayed in the figure below. One can use the graph to show how Newton's iteration works. In fact, pick a point x_2^k on the x_2-axis, and *lift* it on the graph of the function ψ. Then *move* the point horizontally on the graph of the function x_2, and then vertically on the x_2-axis. This is the point x_2^{k+1}. Iterating the procedure one can construct the sequence $\{x_2^k\}$. Using this approach, one concludes that if x_2^0 is sufficiently close to zero the iteration yields a sequence converging to $x_2 = 0$. If $|x_2|$ is large, then the sequence diverges.

e) As shown in part c.i), Newton's iteration is composed of two decoupled iterations. The iteration for x_1 yields a globally converging sequence, whereas the iteration for x_2 converges only for $|x_2^0|$ sufficiently small (to be precise, for $|x_2^0| < 1.39...$). Hence, for all initial points (x_1^0, x_2^0) such that $|x_2^0| < 1.39...$, the iteration yields a sequence converging to the global minimizer.



Exercise 21 The company XYZ has invested £20000 to develop a new product. The product can be manufactured for £2 per unit. The company then performs a marketing research. The conclusion of the research is that if the company spends £a on advertising then it can sell the product at price £p per unit and it will sell $2000 + 4\sqrt{a} - 20p$ units.

a) Compute the revenue for sales as a function of a and p.

b) Compute the overall costs associated to the production and commercialization of the product, that is the development cost plus the production cost and the advertising cost, as a function of a and p.

c) Compute the company's profit as a function of a and p.

d) The company wishes to select a and p to maximize the profit. Pose this problem as an unconstrained optimization problem (disregard the non-negativity conditions on a and p).

e) Compute the unique stationary point of the profit. Using second order sufficient conditions of optimality show that the stationary point is a local maximizer.

f) Assume that the company is forced to fix the sale price of the product to $p = \tilde{p}$, with $\tilde{p} > 2$.

   i) Determine the optimal advertising cost as a function of $\tilde{p}$.

   ii) Determine the optimal profit as a function of $\tilde{p}$.

   iii) Plot the optimal profit as a function of the fixed price $\tilde{p}$ and show that as $\tilde{p}$ increases the profit becomes negative.

Solution 21

a) The revenue for sales is given by

$$\text{revenue} = p(2000 + 4\sqrt{a} - 20p).$$

b) The costs are

$$\begin{aligned} \text{production cost} &= 2(2000 + 4\sqrt{a} - 20p), \\ \text{development cost} &= 20000, \\ \text{advertising cost} &= a. \end{aligned}$$

Hence

$$\text{total cost} = 24000 + 8\sqrt{a} - 40p + a.$$

c) The profit is given by

$$\text{profit} = p(2000 + 4\sqrt{a} - 20p) - (24000 + 8\sqrt{a} - 40p + a).$$

d) The optimization problem is

$$\max_{a,p} = p(2000 + 4\sqrt{a} - 20p) - (24000 + 8\sqrt{a} - 40p + a).$$

e) The statiory points of the profit are the solutions of the equations

$$0 = \frac{\partial \text{profit}}{\partial a} = 2\frac{p}{\sqrt{a}} - \frac{4}{\sqrt{a}} - 1, \qquad 0 = \frac{\partial \text{profit}}{\partial p} = 2\frac{p}{\sqrt{a}} - \frac{4}{\sqrt{a}} - 1.$$

The only solution is

$$a^\star = \frac{60025}{4} = 15006.25, \qquad p^\star = \frac{253}{4} = 63.25.$$

The Hessian of the profit at the stationary point is

$$H(a^\star, p^\star) = -\begin{bmatrix} \dfrac{2}{60025} & -\dfrac{4}{245} \\ -\dfrac{4}{245} & 40 \end{bmatrix},$$

which is negative definite, hence the point $(a^\star, p^\star)$ is a local maximizer.

f) The profit for fixed price is

$$\text{profit fix price} = \tilde{p}(2000 + 4\sqrt{a} - 20\tilde{p}) - (24000 + 8\sqrt{a} - 40\tilde{p} + a).$$

i) The optimal advertising cost $\tilde{a}^\star$ is given by the solution of the equation

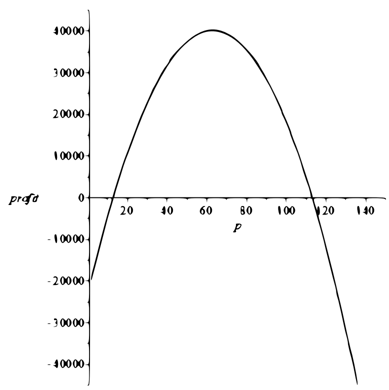
$$0 = \frac{\partial \text{profit fix price}}{\partial a},$$

which gives $\tilde{a}^\star = 4(\tilde{p} - 2)^2$.

ii) The resulting optimal profit is

$$\text{profit fix price}^\star = 2024\tilde{p} - 16\tilde{p}^2 - 23984.$$

iii) The optimal profit as a function of the fixed price $\tilde{p}$ is displayed in the graph below. Note that as $\tilde{p}$ increases the optimal profit becomes negative (because of the term $-16\tilde{p}^2$).



Exercise 22 Consider the function

$$f(x) = (x_1 - 2)^4 + (x_1 - 2)^2 x_2^2 + (x_2 + 1)^2.$$

a) Compute the unique stationary point $x_\star$ of the function f.

b) Using second order sufficient conditions of optimality show that the stationary point determined in part a) is a local minimizer. Hence, show that f is radially unbounded and that the stationary point determined in part a) is the global minimizer of f.

c) Write the modified Newton's iteration for the minimization of the function f given by

$$x_{k+1} = x_k - \left[\nabla^2 f(x_\star)\right]^{-1} \nabla f(x_k).$$

d) Run five steps of the modified Newton's iteration in part c) from the starting point $(1.5, 0)$.

e) Run four steps of the modified Newton's iteration in part c) from the starting point $(1, 0)$.

f) Show that the research directions generated by the modified Newton's iteration in part c) are descent directions satisfying the condition of angle. Explain why the iteration is not globally convergent.

Solution 22

a) The stationary points of the function f are computed by solving the equations

$$0 = \nabla f = \begin{bmatrix} 2(x_1 - 2)(2(x_1 - 2)^2 + x_2^2) \\ 2x_1^2 x_2 - 8x_1 x_2 + 10x_2 + 2 \end{bmatrix}.$$

As a result, the point $x_\star = (2, -1)$ is the unique stationary point.

b) The Hessian matrix of the function f is

$$\nabla^2 f(x) = \begin{bmatrix} 12(x_1-2)^2+2x_2^2 & 4(x_1-2)x_2 \\ 4(x_1-2)x_2 & 2(x_1-2)^2+2 \end{bmatrix},$$

hence

$$\nabla^2 f(x_\star) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Since $\nabla^2 f(x_\star) > 0$, $x_\star$ is a minimizer of f. Note now that

$$0 \leq (x_1-2)^4 + (x_2+1)^2 \leq f,$$

and the function $(x_1-2)^4+(x_2+1)^2$ takes non-negative values and it is radially unbounded (it is the sum of two squares, one involving x_1 and one involving x_2). Hence, f is radially unbounded, and since $f(x_\star) = 0$, $x_\star$ is the global minimizer of f.

c) The modified Newton's iteration is given by

$$x_{k+1} = x_k - \frac{1}{2}\nabla f(x_k) = \begin{bmatrix} x_{k,1} - (x_{k,1}-2)(2(x_{k,1}-2)^2 + x_{k,2}^2) \\ -4x_{k,2} - x_{k,2}x_{k,1}^2 + 4x_{k,2}x_{k,1} - 1 \end{bmatrix}.$$

d) The points generated by the modified Newton's iteration from the starting point $x_0 = (3/2, 0)$ are

$$x_1 = (1.75, -1), \quad x_2 = (2.03125, -0.9375), \quad x_3 = (2.003723145, -.9990844731),$$

$$x_4 = (2.000006711, -.9999861507), \quad x_5 = (2.000000000, -1.000000000).$$

e) The points generated by the modified Newton's iteration from the starting point $x_0 = (1, 0)$ are

$$x_1 = (3, -1), \quad x_2 = (0, 0), \quad x_3 = (16, -1), \quad x_4 = (-5486, 195).$$

f) The research direction used in the modified Newton's iteration is $-1/2\ \nabla f(x_k)$, which is nothing else than the direction of the anti-gradient, hence it is a descent direction satisfying the condition of angle. The reason why the method is not globally convergent is that the line search parameter is *fixed* to $\alpha = 1/2$, and this may not yield a descent algorithm at each step.

Exercise 23 Consider the function

$$f(x) = \frac{1}{2}x_1^2 + \frac{m}{2}x_2^2,$$

with $m > 0$. The function has a global minimizer at $x_\star = 0$.

a) Show that the gradient algorithm with exact line search for the function f can be written as

$$x_{k+1} = x_k - \frac{x_{k,1}^2 + m^2x_{k,2}^2}{x_{k,1}^2 + m^3x_{k,2}^2}\begin{bmatrix} x_{k,1} \\ m\ x_{k,2} \end{bmatrix}$$

b) Let $m = 9$ and $x_0 = [9, 1]'$. Show that the sequence of points generated by the gradient algorithm is given by

$$x_k = \begin{bmatrix} 9 \\ (-1)^k \end{bmatrix}(0.8)^k.$$

(Hint: assume that for the given values of m and x_0 the quantity

$$\frac{x_1^2 + m^2x_2^2}{x_1^2 + m^3x_2^2}$$

remains constant for all iterations of the algorithm.)

c) Compute the speed of convergence of the sequence generated by the algorithm and in particular show that
$$\frac{\|x_{k+1} - x_\star\|}{\|x_k - x_\star\|} = \text{constant}$$
for every k, where $\|v\| = \sqrt{v'v}$.

Solution 23

a) Note that
$$\nabla f = \begin{bmatrix} x_1 \\ mx_2 \end{bmatrix},$$
hence the gradient algorithm is described by the iteration
$$x_{k+1,1} = x_{k,1} - \alpha x_{k,1}, \qquad x_{k+1,2} = x_{k,2} - \alpha m\, x_{k,2}.$$
Replacing x_{k+1} in f yields
$$f(x_{k+1}) = \frac{1}{2}\left(x_{k,1}^2 + m\, x_{k,2}^2\right) - \alpha\left(x_{k,2}^2 + m^2 x_{k,2}^2\right) + \frac{1}{2}\left(x_{k,2}^2 + m^3\, x_{k,2}^2\right)\alpha^2.$$
To obtain the exact linear search parameter one has to compute the stationary point of $f(x_{k+1})$ as a function of α (since $f(x_{k+1})$ is convex in α), that is
$$\alpha_\star = \frac{x_{k,1}^2 + m^2 x_{k,2}^2}{x_{k,2}^2 + m^3\, x_{k,2}^2}.$$
As a result, the gradient algorithm with exact line search is given by
$$x_{k+1} = x_k - \alpha_\star \nabla f(x_k),$$
as given in the question.

b) As indicated in the question, for the considered initial condition and value of m the value of $\alpha_\star$ is constant, namely
$$\alpha_\star = \frac{x_{0,1}^2 + m^2 x_{0,2}^2}{x_{0,1}^2 + m^3\, x_{0,2}^2} = 1/5.$$
As a result, the gradient iteration is given by
$$x_{k+1,1} = \frac{4}{5} x_{k,1}, \qquad x_{k+1,2} = -\frac{4}{5} x_{k,2}.$$
This yields
$$x_{k,1} = x_{0,1}\left(\frac{4}{5}\right)^k = 9\left(\frac{4}{5}\right)^k, \qquad x_{k,2} = x_{0,2}\left(-\frac{4}{5}\right)^k = (-1)^k\left(\frac{4}{5}\right)^k.$$

c) Note that $x_\star = 0$, hence
$$\|x_{k+1}\|^2 = \left(9\left(\frac{4}{5}\right)^{k+1}\right)^2 + \left((-1)^{k+1}\left(\frac{4}{5}\right)^{k+1}\right)^2 = 82\left(\frac{4}{5}\right)^{2(k+1)},$$
$$\|x_k\|^2 = 82\left(\frac{4}{5}\right)^{2k},$$
thus
$$\frac{\|x_{k+1} - x_\star\|}{\|x_k - x_\star\|} = \frac{4}{5}.$$
The sequence thus converges with linear speed of convergence.

Exercise 24 Consider the problem of computing the average of four numbers, a_1, a_2, a_3 and a_4. This problem can be posed as an unconstrained optimization problem as follows

$$\min_x f(x),$$

with

$$f(x) = (x - a_1)^2 + (x - a_2)^2 + (x - a_3)^2 + (x - a_4)^2.$$

a) Compute the unique stationary point $x_\star$ of the function f and show that $x_\star$ is indeed the average of a_1, a_2, a_3 and a_4.

b) Using second order sufficient conditions of optimality show that the stationary point determined in part a) is a local minimizer. Hence, show that f is radially unbounded and that the stationary point determined in part a) is the global minimizer of f.

c) Assume $a_1 = 1$, $a_2 = 2$, $a_3 = 3$ and $a_4 = -6$.

   i) Write the gradient method for the minimization of the function f and determine the exact line search parameter $\alpha^\star$.

   ii) Consider the gradient method with line search parameter $\alpha = \gamma\, \alpha^\star$, with $\gamma \in [0, 3]$. Determine for which values of γ the iteration yields a converging sequence and, for these values of γ, determine the speed of convergence of the sequence.

Solution 24 Note that

$$\begin{aligned} f(x) &= (x - a_1)^2 + (x - a_2)^2 + (x - a_3)^2 + (x - a_4)^2 \\ &= x^2 - 2a_1 + a_1^2 + x^2 - 2a_2 + a_2^2 + x^2 - 2a_3 + a_3^2 + x^2 - 2a_4 + a_4^2 \\ &= 4x^2 - 2(a_1 + a_2 + a_3 + a_4)x + a_1^2 + a_2^2 + a_3^2 + a_4^2. \end{aligned}$$

a) The first order necessary condition of optimality is

$$0 = \nabla f = 8x - 2(a_1 + a_2 + a_3 + a_4),$$

which yields

$$x^* = \frac{a_1 + a_2 + a_3 + a_4}{4}.$$

Clearly, x^* is the average of a_1, a_2, a_3, a_4.

b) The second order sufficient condition of optimality is

$$\nabla^2 f = 8 > 0,$$

hence x^* is a local minimizer. Note that f is strictly convex and this implies that x^* is global minimizer.

c) Note that $f(x) = 4x^2 + c^2$, with $c^2 = a_1^2 + a_2^2 + a_3^2 + a_4^2$.

   i) The gradient is

$$\nabla f = 8x$$

and the gradient algorithm gives

$$x_{k+1} = x_k - 8\alpha_k x_k = (1 - 8\alpha_k)x_k.$$

Note that

$$f(x_{k+1}) = 4(1 - 16\alpha_k + 64\alpha_k^2)x_k^2 + c^2$$

and

$$f(x_k) = 4x_k^2 + c^2$$

yields

$$f(x_{k+1}) - f(x_k) = 4(64\alpha_k^2 - 16\alpha_k)x_k^2.$$

To find the exact line search parameter solve

$$0 = \frac{\partial[f(x_{k+1}) - f(x_k)]}{\partial \alpha} = 4(128\alpha_k - 16)x_k^2,$$

obtaining

$$\alpha^* = \frac{1}{8}.$$

ii) Note now that

$$x_{k+1} = x_k - 8\gamma\alpha^* x_k = (1-\gamma)x_k.$$

To have convergence we need

$$|1-\gamma| < 1.$$

Hence, $\gamma \in (0,2)$. For $\gamma = 0$ or $\gamma = 2$, $|x_{k+1}| = |x_k|$, and the sequence does not converge. For $\gamma \in (2,3]$ $|x_{k+1}| > |x_k|$, hence the sequence diverges. For $\gamma \in (0,2)$ the speed of convergence is linear, since

$$\frac{|x_{k+1} - x^*|}{|x_k - x^*|} = \frac{|x_{k+1}|}{|x_k|} = |1-\gamma|.$$

Exercise 25 Consider the problem of minimizing the function

$$\min_x f(x),$$

with

$$f(x) = \frac{1}{2}x_1^2 - x_1x_2 + \frac{1}{4}x_2^4 - \frac{1}{3}x_2^3.$$

a) Compute the stationary points of the function f.

b) Using second order sufficient conditions of optimality classify the stationary points determined in part a). Hence, determine the global minimizer of f.

c) Consider the problem of minimizing the function using the so-called gradient method with extrapolation, that is the method defined by the iteration

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k) + \beta_k(x_k - x_{k-1}),$$

with $\alpha_k > 0$ and $\beta_k \in [0,1)$, for all $k \geq 0$, and $x_{-1} = x_0$. Let $x_0 = (1,1)$.

i) Argue that the first step of the gradient method with extrapolation coincides with the first step of the gradient method.

ii) Run one iteration of the gradient method with extrapolation and determine the point x_1. Note that x_1 is a function of α_0 hence write a condition on α_0 such that the algorithm is a descent algorithm, that is $f(x_1) < f(x_0)$. Explain why β_0 does not appear in the descent condition $f(x_1) < f(x_0)$.

iii) Pick $\alpha_k = 1/2$ for all k. Run one more iteration of the gradient method with extrapolation (using as initial condition the point x_1 determined in part c.ii), that is compute the point x_2. Determine a condition on β_1 yielding a descent algorithm. Explain why $\beta_1 = 0$ is a feasible selection of β_1 and argue that it is not the best selection.

Solution 25

a) The first order necessary condition of optimality is

$$0 = \nabla f = \begin{bmatrix} x_1 - x_2 \\ -x_1 + x_2^3 - x_2^2 \end{bmatrix},$$

which gives the equations

$$\begin{aligned} &x_1 = x_2, \\ &x_2(x_2^2 - x_2 - 1) = 0. \end{aligned}$$

The stationary points are therefore $P_1 = (0,0)$, $P_2 = \left(\frac{1+\sqrt{5}}{2}, \frac{1+\sqrt{5}}{2}\right)$ and $P_3 = \left(\frac{1-\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2}\right)$.

b) The Hessian matrix of the function f is

$$\nabla^2 f(x) = \begin{bmatrix} 1 & -1 \\ -1 & 3x_2^2 - 2x_2 \end{bmatrix}.$$

Evaluating the matrix at the stationary points yields

$$\nabla^2 f(P_1) = \begin{bmatrix} 1 & -1 \\ -1 & 0 \end{bmatrix},$$

which is indefinite,

$$\nabla^2 f(P_2) = \begin{bmatrix} 1 & -1 \\ -1 & \frac{7+\sqrt{5}}{2} \end{bmatrix},$$

which is positive definite, and

$$\nabla^2 f(P_3) = \begin{bmatrix} 1 & -1 \\ -1 & \frac{7-\sqrt{5}}{2} \end{bmatrix},$$

which is also positive definite. Hence, P_1 is a saddle point, whereas P_2 and P_3 are two local minimizers. Computing the function at this points yields $f(P_2) = -1.0075$ and $f(P_3) = -0.0758$. Since f is radially unbounded, *i.e.*

$$\lim_{|x|\to+\infty} f(x) = +\infty,$$

the point P_2 is the global minimizer of f.

c) i) In the first step of the gradient method with extrapolation the term multiplied by β_0 is zero (because of the way the algorithm is initialized), hence the iteration coincides with the gradient iteration.

ii) Running one iteration of the algorithm from the indicated initial conditions yields the point

$$x_1 = \begin{bmatrix} 1 \\ 1+\alpha_0 \end{bmatrix}.$$

Note that, consistently with the answer to c.i), the parameter β_0 does not contribute to the point x_1. To check the descent condition note that

$$f(x_1) - f(x_0) = \frac{1}{12}\alpha_0(3\alpha_0^3 + 8\alpha_0^2 + 6\alpha_0 - 12),$$

hence α_0 should be selected such that

$$\alpha_0(3\alpha_0^3 + 8\alpha_0^2 + 6\alpha_0 - 12) < 0$$

which is the case for α_0 strictly positive and sufficiently small (approximately smaller than 0.82).

iii) Using $\alpha_k = 1/2$ and running one more iteration of the algorithm yields

$$x_2 = \begin{bmatrix} 5/4 \\ 23/16 + 1/2\beta_1 \end{bmatrix}.$$

The descent condition is now

$$f(x_2) - f(x_1) = -0.0788 - 0.1729\beta_1 + 0.415\beta_1^2 + 0.138\beta_1^3 + 0.015625\beta_1^4 < 0,$$

which shows that β_1 should be non-negative and smaller than approximately 0.6. The selection $\beta_1 = 0$ gives a descent condition because for such value of β_1 one has essentially the gradient iteration, for which the descent condition holds for the given selection of α_1. However, $\beta_1 = 0$ is not *optimal* since one could have a greater decrease selecting a strictly positive value of β_1. The optimal selection for this particular case is approximately $\beta_1 = 0.2$.

Exercise 26 The proximal method is a descent method in which the problem

$$\min_x f(x)$$

is replaced by the sequence of modified problems

$$\min_x \left(f(x) + \frac{1}{2\gamma_k}\|x - x_k\|^2 \right),$$

where x_k is the current estimate of the solution of the problem and x_{k+1} is the solution of the modified minimization problem, and $\gamma_k > 0$.
Consider the quadratic function

$$f(x) = \frac{1}{2}x'Qx + c'x + d,$$

with $Q = Q' > 0$. Recall that the function has a global minimizer at $x^\star = -Q^{-1}c$.

a) Write the optimization problem used in the proximal method and state under what conditions the problem has a unique solution.

b) Solve explicitly the optimization problem arising from the proximal method, that is determine x_{k+1} as a function of x_k. In particular, write the relation between x_{k+1} and x_k in the form

$$x_{k+1} = Ax_k + b,$$

in which A is a matrix and b is a vector. (Note that A and b are functions of k.) Write explicitly the matrix A and the vector b as a function of Q, c and γ_k.

c) Determine the fixed point $\bar{x}$ of the equation $x_{k+1} = Ax_k + b$, that is the point $\bar{x}$ such that

$$\bar{x} = A\bar{x} + b,$$

and show that the point is the global minimizer of the quadratic function.

d) Show that the iteration arising from the proximal method is globally convergent for all $\gamma_k > 0$. This can be achieved using the following steps.

   i) Show that $A = (\gamma_k Q + I)^{-1}$ and that $A'A < I$.

   ii) Write the iteration arising from the proximal method in the form $x_{k+1} - x^\star = A(x_k - x^\star) + \tilde{b}$ and show that $\tilde{b} = 0$.

   iii) Exploit the results in parts d.i) and d.ii) to demonstrate the global convergence claim. Discuss also the effect of the parameter γ_k on the speed of convergence of the algorithm.

Solution 26

a) The optimization problem used in the proximal method is

$$\begin{aligned} x_{k+1} &= \text{argmin}_x \left(\tfrac{1}{2}x'Qx + c'x + d + \tfrac{1}{2\gamma_k}(x - x_k)'(x - x_k) \right) \\ &= \text{argmin}_x \tfrac{1}{2}x' \left(Q + \tfrac{1}{\gamma_k} I \right) x + \left(c' - \tfrac{1}{\gamma_k} x_k' \right) x + d + \tfrac{1}{2\gamma_k} x_k' x_k. \end{aligned}$$

Note that the function to be minimized is again a quadratic function. Hence, it has a unique minimizer if and only if the matrix $\left(Q + \dfrac{1}{\gamma_k} I \right)$ is positive definite (which is always the case since $Q > 0$ and $\gamma_k > 0$).

b) The solution of the optimization problem is

$$x^* = x_{k+1} = -\left(Q + \frac{1}{\gamma_k} I \right)^{-1} \left(c - \frac{1}{\gamma_k} x_k \right).$$

This can be rewritten as

$$x_{k+1} = \frac{1}{\gamma_k}\left(Q + \frac{1}{\gamma_k} I\right)^{-1} x_k - \left(Q + \frac{1}{\gamma_k} I\right)^{-1} c.$$

Hence, $A = \frac{1}{\gamma_k}\left(Q + \frac{1}{\gamma_k} I\right)^{-1}$ and $b = -\left(Q + \frac{1}{\gamma_k} I\right)^{-1} c$.

c) The fixed point $\bar{x}$ of the equation above is such that

$$\bar{x} = \frac{1}{\gamma_k}\left(Q + \frac{1}{\gamma_k} I\right)^{-1} \bar{x} - \left(Q + \frac{1}{\gamma_k} I\right)^{-1} c.$$

Multiplying on left by $\left(Q + \frac{1}{\gamma_k} I\right)$ yields

$$\left(Q + \frac{1}{\gamma_k} I\right)\bar{x} = \frac{1}{\gamma_k}\bar{x} - c,$$

which gives

$$\bar{x} = -Q^{-1}c,$$

i.e. the global minimizer of the quadratic function.

d) i) Trivially

$$A = \frac{1}{\gamma_k}\left(Q + \frac{1}{\gamma_k} I\right)^{-1} = (\gamma_k)^{-1}\left(Q + \frac{1}{\gamma_k} I\right)^{-1} = (\gamma_k Q + I)^{-1}.$$

Observe now that $A'A < I$. Multiplying both sides with the matrix $(A'A)^{-1}$ (which exists because $A'A$ is positive definite) yields

$$I < (\gamma_k Q + I)(\gamma_k Q + I)',$$

hence

$$\gamma_k^2 Q'Q + \gamma_k(Q' + Q) > 0,$$

which holds by positivity of Q and γ_k.

ii) We add and subtract x^* and Ax^* to the equation in part b) obtaining

$$x_{k+1} - x^* = A(x_k - x^*) + b - x^* + Ax^*.$$

Defining $\tilde{b} = b - x^* + Ax^*$, it remains to prove that $\tilde{b} = 0$. Multiplying the expression of $\tilde{b}$ on the left-hand side by $\left(Q + \frac{1}{\gamma_k} I\right)$ yields

$$\begin{aligned}\left(Q + \frac{1}{\gamma_k} I\right)\tilde{b} &= \left(Q + \frac{1}{\gamma_k} I\right)\left[\frac{1}{\gamma_k}\left(Q + \frac{1}{\gamma_k} I\right)^{-1} x^* - \left(Q + \frac{1}{\gamma_k} I\right)^{-1} c - x^*\right] \\ &= -c - \left(Q + \frac{1}{\gamma_k} I\right) x^* + \frac{1}{\gamma_k} I x^* = -c - Qx^*.\end{aligned}$$

The claim is proved substituting the minimizer $x^* = -Q^{-1}c$ in the last equation.

iii) The equation

$$x_{k+1} - x^* = A(x_k - x^*)$$

is a linear difference equation in which all the eigenvalues of the dynamic matrix A have modulus strictly smaller than one. Hence, the state $x_k - x^*$ converges globally to zero, *i.e.* x_k converges to the optimal solution x^*. The greater the value of γ_k, the smaller the modulus of the eigenvalues of A. Thus, increasing γ_k corresponds to a faster convergence of the algorithm.

Exercise 27 The Levenberg-Marquardt algorithm is a modification of Newton's method for the solution of nonlinear equations. In the case of the scalar equation

$$f(x) = 0,$$

with $x \in I\!R$ and f differentiable, the Levenberg-Marquardt algorithm can be written as (note that f' denotes the first derivative of f with respect to x)

$$x_{k+1} = x_k - 2\frac{f(x_k)}{f'(x_k) + f'(\bar{x})}, \qquad \bar{x} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

Consider now the problem of minimizing the function

$$q(x) = \frac{x^4}{4} + \frac{4}{3}x^3 - 10x$$

(note that the global minimizer of the function q is $x = 1.365230013$).

a) Re-cast the considered minimization problem as the problem of finding the solution of a scalar equation.

b) Write Newton's iteration for the solution of the equation determined in part a).

c) Run four iterations of the Newton's iteration in part b) with $x_0 = 3$ and evaluate the first four values of the sequence of the relative errors

$$RE^N_{k+1} = \frac{x_{k+1} - x^\star}{(x_k - x^\star)^3},$$

that is evaluate RE^N_1, RE^N_2, RE^N_3 and RE^N_4. Hence argue that Newton's method does not have speed of convergence of order three. Explain why this is not un-expected.

d) Write now the Levenberg-Marquardt algorithm for the solution of the equation determined in part a).
(Hint: write the algorithm as two equations, that is do not substitute $\bar{x}$ into the first equation of the algorithm.)

e) Run four iterations of the Levenberg-Marquardt algorithm in part d) with $x_0 = 3$ and evaluate the first four values of the sequence of the relative errors

$$RE^{LM}_{k+1} = \frac{x_{k+1} - x^\star}{(x_k - x^\star)^3}.$$

Hence argue that the Levenberg-Marquardt algorithm is faster than Newton's algorithm. (It is well-known that the Levenberg-Marquardt algorithm has, under similar assumptions o those required by Newton's method, speed of convergence of order three.)

Solution 27

a) The minimization problem can be re-cast as the problem of finding the stationary points of the function q, that is as the problem of solving the scalar equation

$$k(x) = x^3 + 4x^2 - 10 = 0.$$

As noted above, this equation has a solution at $x = 1.365230013$, which is actually the only solution. Note also that the second derivative of f at $x = 1.365230013$ is positive, hence the point is a local minimizer.

b) Newton's iteration for the solution of the equation $k(x) = 0$ is

$$x_{k+1} = x_k - \frac{x_k^3 + 4x_k^2 - 10}{x_k(3x_k + 8)} = \frac{2x_k^3 + 4x_k^2 + 10}{x_k(3x_k + 8)}.$$

c) Setting $x_0 = 3$ yields

$$
\begin{aligned}
x_1 &= 1.960784314, & RE_1 &= 0.1363174326,\\
x_2 &= 1.486238507, & RE_2 &= 0.5728643018,\\
x_3 &= 1.371823522, & RE_3 &= 3.721080242,\\
x_4 &= 1.365251224, & RE_4 &= 73.99652652.
\end{aligned}
$$

Since the sequence of the relative errors diverges the speed of convergence of the method is not of order three (note the *cube* in the denominator of the definition of the relative error). This is not un-expected, since under the given conditions one can only claim quadratic speed of convergence of Newton's method.

d) The Levenberg-Marquardt iteration is given by the two equations

$$x_{k+1} = x_k - 2\frac{x_k^3 + 4x_k^2 - 10}{(x_k(3x_k + 8)) + (\bar{x}(3\bar{x} + 8))}, \qquad \bar{x} = x_k - \frac{x_k^3 + 4x_k^2 - 10}{x_k(3x_k + 8)}.$$

e) Setting $x_0 = 3$ yields

$$
\begin{aligned}
x_1 &= 1.644853060, & RE_1 &= 0.06400339279,\\
x_2 &= 1.369582249, & RE_2 &= 0.1990643754,\\
x_3 &= 1.365230035, & RE_3 &= 0.2668611603,\\
x_4 &= 1.365230013, & RE_4 &\approx 0.
\end{aligned}
$$

Since the sequence of the relative errors converges to zero, the speed of convergence of the Levenberg-Marquardt iteration is at least of order three, definitely faster than Newton's iteration.

Chapter 3

Nonlinear programming

3.1 Introduction

In this chapter we discuss the basic tools for the solution of optimization problems of the form

$$P_0 \begin{cases} \min\limits_x f(x) \\ g(x) = 0 \\ h(x) \leq 0. \end{cases} \tag{3.1}$$

In the problem P_0 there are both equality and inequality constraints[1]. However, sometimes for simplicity, or because a method has been developed for problems with special structure, we will refer to problems with only equality constraints, *i.e.* to problems of the form

$$P_1 \begin{cases} \min\limits_x f(x) \\ g(x) = 0, \end{cases} \tag{3.2}$$

or to problems with only inequality constraints, *i.e.* to problems of the form

$$P_2 \begin{cases} \min\limits_x f(x) \\ h(x) \leq 0. \end{cases} \tag{3.3}$$

In all the above problems we have $x \in I\!R^n$, $f : I\!R^n \to I\!R$, $g : I\!R^n \to I\!R^m$, and $h : I\!R^n \to I\!R^p$. From a formal point of view it is always possible to transform the equality constraint $g_i(x) = 0$ into a pair of inequality constraints, *i.e.* $g_i(x) \leq 0$ and $-g_i(x) \leq 0$. Hence, problem P_1 can be (equivalently) described by

$$\tilde{P}_1 \begin{cases} \min\limits_x f(x) \\ g(x) \leq 0 \\ -g(x) \leq 0, \end{cases}$$

which is a special case of problem P_2. In the same way, it is possible to transform the inequality constraint $h_i(x) \leq 0$ into the equality constraint $h_i(x) + y_i^2 = 0$, where y_i is an auxiliary variable (also called *slack* variable). Therefore, defining the extended vector $z = [x', y']'$, problem P_2 can be rewritten as

$$\tilde{P}_2 \begin{cases} \min\limits_z f(x) \\ h(x) + Y = 0, \end{cases}$$

with

$$Y = \begin{bmatrix} y_1^2 \\ y_2^2 \\ \vdots \\ y_p^2 \end{bmatrix},$$

which is a special case of problem P_1.

[1]The condition $h(x) \leq 0$ has to be understood element-wise, *i.e.* $h_i(x) \leq 0$ for all i.

Note however, that the transformation of equality constraints into inequality constraints yields an increase in the number of constraints, whereas the transformation of inequality constraints into equality constraints results in an increased number of variables.
Given problem P_0 (or P_1, or P_2), a point x satisfying the constraints is said to be an admissible point, and the set of all admissible points is called the admissible set and it is denoted with $\mathcal{X}$. Note that the problem makes sense only if $\mathcal{X} \neq \emptyset$.
In what follows it is assumed that the functions f, g and h are two times differentiable, however we do not make any special hypothesis on the form of such functions. Note however, that if g and h are linear there are special algorithms, and linear/quadratic programming algorithms are used if f is linear/quadratic and g and h are linear. We do not discuss these special algorithms, and concentrate mainly on algorithms suitable for general problems.

3.2 Definitions and existence conditions

Consider the problem P_0 (or P_1, or P_2). The following definitions are instrumental to provide a necessary condition and a sufficient condition for the existence of local minima.

Definition 6 *An open ball with center $x^\star$ and radius $\theta > 0$ is the set*

$$B(x^\star, \theta) = \{x \in I\!R^n \mid \|x - x^\star\| < \theta\}.$$

Definition 7 *A point $x^\star \in \mathcal{X}$ is a constrained local minimum if there exists $\theta > 0$ such that*

$$f(y) \geq f(x^\star), \tag{3.4}$$

for all $y \in \mathcal{X} \cap B(x^\star, \theta)$.
A point $x^\star \in \mathcal{X}$ is a constrained global minimum if

$$f(y) \geq f(x^\star), \tag{3.5}$$

for all $y \in \mathcal{X}$.
If the inequality (3.4) (or (3.5)) holds with a strict inequality sign for all $y \neq x^\star$ then the minimum is said to be strict.

Definition 8 *The i-th inequality constraints $h_i(x)$ is said to be active at $\tilde{x}$ if $h_i(\tilde{x}) = 0$. The set $I_a(\tilde{x})$ is the set of all indexes i such that $h_i(\tilde{x}) = 0$,* i.e.

$$I_a(\tilde{x}) = \{i \in \{1, 2, \cdots, p\} \mid h_i(\tilde{x}) = 0\}.$$

The vector $h_a(\tilde{x})$ is the subvector of $h(x)$ corresponding to the active constraints, i.e.

$$h_a(\tilde{x}) = \{h_i(\tilde{x}) \mid i \in I_a(\tilde{x}).$$

Definition 9 *A point $\tilde{x}$ is a regular point for the constraints if at $\tilde{x}$ the gradients of the active constraints,* i.e. *the vectors $\nabla g_i(\tilde{x})$, for $i = 1, \cdots, m$ and $\nabla h_i(\tilde{x})$, for $i \in I_a(\tilde{x})$, are linearly independent.*

The definition of regular point is given because, the necessary and the sufficient conditions for optimality, in the case of regular points are relatively simple. To state these conditions, and with reference to problem P_0, consider the Lagrangian function

$$L(x, \lambda, \rho) = f(x) + \lambda' g(x) + \rho' h(x) \tag{3.6}$$

with $\lambda \in I\!R^m$ and $\rho \in I\!R^p$. The vectors λ and ρ are called multipliers.
With the above ingredients and definitions it is now possible to provide a necessary condition and a sufficient condition for local optimality.

Theorem 14 *[First order necessary condition] Consider problem P_0. Suppose $x^\star$ is a local solution of the problem P_0, and $x^\star$ is a regular point for the constraints.*
Then there exist (unique) multipliers $\lambda^\star$ and $\rho^\star$ such that[2]

$$\begin{aligned}
&\nabla_x L(x^\star, \lambda^\star, \rho^\star) = 0 \\
&g(x^\star) = 0 \\
&h(x^\star) \leq 0 \\
&\rho^\star \geq 0 \\
&(\rho^\star)' h(x^\star) = 0.
\end{aligned} \tag{3.7}$$

Conditions (3.7) are known as Kuhn-Tucker conditions.

Definition 10 *Let $x^\star$ be a local solution of problem P_0 and let $\rho^\star$ be the corresponding (optimal) multiplier. At $x^\star$ the condition of strict complementarity holds if $\rho_i^\star > 0$ for all $i \in I_a(x^\star)$.*

Theorem 15 *[Second order sufficient condition] Consider the problem P_0. Assume that there exist $x^\star$, $\lambda^\star$ and $\rho^\star$ satisfying conditions (3.7). Suppose moreover that $\rho^\star$ is such that the condition of strict complementarity holds at $x^\star$. Suppose finally that*

$$s' \nabla^2_{xx} L(x^\star, \lambda^\star, \rho^\star) s > 0 \tag{3.8}$$

for all $s \neq 0$ such that

$$\begin{bmatrix} \dfrac{\partial g(x^\star)}{\partial x} \\ \dfrac{\partial h_a(x^\star)}{\partial x} \end{bmatrix} s = 0.$$

Then $x^\star$ is a strict constrained local minimum of problem P_0.

Remark. Necessary and sufficient conditions for a global minimum can be given under proper convexity hypotheses, *i.e.* if the function f is convex in $\mathcal{X}$, and if $\mathcal{X}$ is a convex set. This is the case, for example if there are no inequality constraints and if the equality constraints are linear. $\diamond$

[2]We denote with $\nabla_x f$ the vector of the partial derivatives of f with respect to x.

Remark. If all points in $\mathcal{X}$ are regular points for the constraints then conditions (3.7) yield a set of points $\mathcal{P}$, *i.e.* the points satisfying conditions (3.7), and among these points there are all constrained local minima (and also the constrained global minimum, if it exists). However, if there are points in $\mathcal{X}$ which are not regular points for the constraints, then the set $\mathcal{P}$ may not contain all constrained local minima. These have to be searched in the set $\mathcal{P}$ and in the set of non-regular points. ⋄

Remark. In what follows, we will always tacitly assume that the conditions of regularity and of strict complementarity hold. ⋄

3.2.1 A simple proof of Kuhn-Tucker conditions for equality constraints

Consider problem P_1, *i.e.* a minimization problem with only equality constraints, and a point $x^\star$ such that $g(x^\star) = 0$, *i.e.* $x^\star \in \mathcal{X}$. Suppose that[3]

$$\text{rank}\frac{\partial g}{\partial x}(x^\star) = m$$

i.e. $x^\star$ is a regular point for the constraints, and that $x^\star$ is a constrained local minimum. By the implicit function theorem, there exist a neighborhood of $x^\star$, a partition of the vector x, *i.e.*

$$x = \begin{bmatrix} u \\ v \end{bmatrix},$$

with $u \in I\!R^m$ and $v \in I\!R^{n-m}$, and a function ϕ such that the constrains $g(x) = 0$ can be (locally) rewritten as

$$u = \phi(v).$$

As a result (locally)

$$\begin{cases} \min\limits_x f(x) \\ g(x) = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} \min\limits_{u,v} f(u,v) \\ u = \phi(v) \end{cases} \quad \Leftrightarrow \quad \min_v f(\phi(v), v),$$

i.e. problem P_1 is (locally) equivalent to a unconstrained minimization problem. Therefore

$$0 = \nabla f(\phi(v^\star), v^\star) = \left(\frac{\partial f}{\partial u}\frac{\partial \phi}{\partial v} + \frac{\partial f}{\partial v}\right)_{x^\star} = \left(-\frac{\partial f}{\partial u}\left(\frac{\partial g}{\partial u}\right)^{-1}\frac{\partial g}{\partial v} + \frac{\partial f}{\partial v}\right)_{x^\star}.$$

Setting

$$\lambda^\star = \left(-\frac{\partial f}{\partial u}\left(\frac{\partial g}{\partial u}\right)^{-1}\right)'_{x^\star}$$

yields

$$\left(\frac{\partial f}{\partial v} + (\lambda^\star)'\frac{\partial g}{\partial v}\right)_{x^\star} = 0 \tag{3.9}$$

[3]Note that m is the number of the equality constraints, and that, to avoid trivial cases, $m < n$.

and

$$\left(\frac{\partial f}{\partial u} + (\lambda^\star)' \frac{\partial g}{\partial u}\right)_{x^\star} = 0. \tag{3.10}$$

Finally, let

$$L = f + \lambda' g,$$

note that equations (3.9) and (3.10) can be rewritten as

$$\nabla_x L(x^\star, \lambda^\star) = 0,$$

and this, together with $g(x^\star) = 0$, is equivalent to equations (3.7).

3.2.2 Quadratic cost function with linear equality constraints

Consider the function

$$f(x) = \frac{1}{2} x' Q x,$$

with $x \in I\!R^n$ and $Q = Q' > 0$, the equality constraints

$$g(x) = Ax - b = 0,$$

with $b \in I\!R^m$ and $m < n$, and the Lagrangian function

$$L(x, \lambda) = \frac{1}{2} x' Q x + \lambda'(Ax - b).$$

A simple application of Theorem 14 yields the necessary conditions of optimality

$$\begin{array}{rcl} \nabla_x L(x^\star, \lambda^\star) & = & Qx^\star + A'\lambda^\star = 0 \\ g(x^\star) & = & Ax^\star - b = 0. \end{array} \tag{3.11}$$

Suppose now that the matrix A is such that $AQ^{-1}A'$ is invertible[4]. As a result, the only solution of equations (3.11) is

$$x^\star = Q^{-1}A'(AQ^{-1}A')^{-1}b \qquad \lambda^\star = -(AQ^{-1}A')^{-1}b.$$

Finally, by Theorem 15, it follows that $x^\star$ is a strict constrained (global) minimum.

3.3 Nonlinear programming methods: introduction

The methods of non-linear programming that have been mostly studied in recent years belong to two categories. The former includes all methods based on the transformation of a constrained problem into one or more unconstrained problems, in particular the so-called (exact or sequential) penalty function methods and (exact or sequential) augmented Lagrangian methods. Sequential methods are based on the solution of a sequence of problems, with the property that the sequence of the solutions of the subproblems converge

[4] This is the case if rank$A = m$.

to the solution of the original problem. Exact methods are based on the fact that, under suitable assumptions, the optimal solutions of an unconstrained problem coincides with the optimal solution of the original problem.
The latter includes the methods based on the transformation of the original problem into a sequence of constrained quadratic problems.
From the above discussion it is obvious that, to construct algorithms for the solution of non-linear programming problems, it is necessary to use efficient unconstrained optimization routines.
Finally, in any practical implementation, it is also important to quantify the complexity of the algorithms in terms of number and type of operations (inversion of matrices, differentiation, ...), and the speed of convergence. These issues are still largely open, and will not be addressed in these notes.

3.4 Sequential and exact methods

3.4.1 Sequential penalty functions

In this section we study the so-called external sequential penalty functions. This name is based on the fact that the solutions of the resulting unconstrained problems are in general not admissible. There are also internal penalty functions (known as barrier functions) but this can be used only for problems in which the admissible set has a non-empty interior. As a result, such functions cannot be used in the presence of equality constraints.
The basic idea of external sequential penalty functions is very simple. Consider problem P_0, the function

$$q(x) = \begin{cases} 0, & \text{if } x \in \mathcal{X} \\ +\infty, & \text{if } x \notin \mathcal{X} \end{cases} \tag{3.12}$$

and the function

$$F = f + q. \tag{3.13}$$

It is obvious that the unconstrained minimization of F yields a solution of problem P_0. However, because of its discontinuous nature, the minimization of F cannot be performed. Nevertheless, it is possible to construct a sequence of continuously differentiable functions, converging to F, and it is possible to study the convergence of the minima of such a sequence of functions to the solutions of problem P_0.
For, consider a continuously differentiable function p such that

$$p(x) = \begin{cases} 0, & \text{if } x \in \mathcal{X} \\ > 0, & \text{if } x \notin \mathcal{X}, \end{cases} \tag{3.14}$$

and the function

$$F_\epsilon = f + \frac{1}{\epsilon} p,$$

with $\epsilon > 0$. It is obvious that[5]

$$\lim_{\epsilon \to 0} F_\epsilon = F.$$

The function F_ϵ is known as external penalty function. The attribute external is due to the fact that, if $\bar{x}$ is a minimum of F_ϵ in general $p(\bar{x}) \neq 0$, *i.e.* $\bar{x} \notin \mathcal{X}$. The term $\frac{1}{\epsilon}p$ is called penalty term, as it penalizes the violation of the constraints. In general, the function p has the following form

$$p = \sum_{i=1}^{m} (g_i)^2 + \sum_{i=1}^{p} (\max(0, h_i))^2. \tag{3.15}$$

Consider now a strictly decreasing sequence $\{\epsilon_k\}$ such that $\lim_{k\to\infty} \epsilon_k = 0$. The sequential penalty function method consists in solving the sequence of unconstrained problems

$$\min_x F_{\epsilon_k}(x),$$

with $x \in I\!R^n$. The most important convergence results for this methods are summarized in the following statements.

Theorem 16 *Consider the problem P_0. Suppose that for all $\sigma > 0$ the set*[6]

$$\mathcal{X}^\sigma = \{x \in I\!R^n \mid |g_i(x)| \leq \sigma, i = 1, \cdots, m\} \cap \{x \in I\!R^n \mid h_i(x) \leq \sigma, i = 1, \cdots, p\}$$

is compact. Suppose moreover that for all k the function $F_{\epsilon_k}(x)$ has a global minimum x_k. Then the sequence $\{x_k\}$ has (at least) one converging subsequence, and the limit of any converging subsequence is a global minimum for problem P_0.

Theorem 17 *Let $x^\star$ be a strict constrained local minimum for problem P_0. Then there exist a sequence $\{x_k\}$ and an integer $\bar{k} > 0$ such that $\{x_k\}$ converges to $x^\star$ and, for all $k \geq \bar{k}$, x_k is a local minimum of $F_{\epsilon_k}(x)$.*

The construction of the function F_ϵ is apparently very simple, and this is the main advantage of the method. However, the minimization of the function F_ϵ may be difficult, especially for small values of ϵ. In fact, it is possible to show, even via simple examples, that as ϵ tends to zero the Hessian matrix of the function F_ϵ becomes ill conditioned. As a result, any unconstrained minimization algorithm used to minimize F_ϵ has a very slow convergence rate. To alleviate this problem, it is possible to use, in the minimization of $F_{\epsilon_{k+1}}$, as initial point the point x_k. However, this is close to the minimum of $F_{\epsilon_{k+1}}$ only if ϵ_{k+1} is close to ϵ_k, *i.e.* only if the sequence $\{\epsilon_k\}$ converges slowly to zero.
We conclude that, to avoid the ill conditioning of the Hessian matrix of F_ϵ, hence the slow convergence of each unconstrained optimization problem, it is necessary to slow down the convergence of the sequence $\{x_k\}$, *i.e.* slow convergence is an intrinsic property of the method. This fact has motivated the search for alternatives methods, as described in the next sections.

[5]Because of the discontinuity of F, the limit has to be considered with proper *care*.

[6]The set $\mathcal{X}^\sigma$ is sometimes called the relaxed admissible set.

Remark. It is possible to show that the local minima of F_ϵ describe (continuous) trajectories that can be extrapolated. This observation is exploited in some sophisticated methods for the selection of initial estimate for the point x_k. However, even with the addition of this extrapolation procedure, the convergence of the method remains slow. ◇

Remark. Note that, if the function p is defined as in equation (3.15), then the function F_ϵ is not two times differentiable everywhere, *i.e.* it is not differentiable in all points in which an inequality constraints is active. This property restricts the class of minimization algorithms that can be used to minimize F_ϵ. ◇

3.4.2 Sequential augmented Lagrangian functions

Consider problem P_1, *i.e.* an optimization problem with only equality constraints. For such a problem the Lagrangian function is

$$L = f + \lambda' g,$$

and the first order necessary conditions require the existence of a multiplier $\lambda^\star$ such that, for any local solution $x^\star$ of problem P_1 one has

$$\begin{array}{l} \nabla_x L(x^\star, \lambda^\star) = 0 \\ \nabla_\lambda L(x^\star, \lambda^\star) = g(x^\star) = 0. \end{array} \tag{3.16}$$

The first of equations (3.16) is suggestive of the fact that the function $L(x, \lambda^\star)$ has a unconstrained minimum in $x^\star$. This is actually not the case, as $L(x, \lambda^\star)$ is not convex in a neighborhood of $x^\star$. However it is possible to render the function $L(x, \lambda^\star)$ convex with the addition of a penalty term, yielding the new function, known as augmented Lagrangian function[7],

$$L_a(x, \lambda^\star) = L(x, \lambda^\star) + \frac{1}{\epsilon}\|g(x)\|^2, \tag{3.17}$$

which, for ϵ sufficiently small, but such that $1/\epsilon$ is finite, has a unconstrained minimum in $x^\star$. This intuitive discussion can be given a formal justification, as shown in the next statement.

Theorem 18 *Suppose that at $x^\star$ and $\lambda^\star$ the sufficient conditions for a strict constrained local minimum for problem P_1 hold. Then there exists $\bar{\epsilon} > 0$ such that for any $\epsilon \in (0, \bar{\epsilon})$ the point $x^\star$ is a unconstrained local minimum for the function $L_a(x, \lambda^\star)$.*
Vice-versa, if for some $\bar{\epsilon}$ and $\lambda^\star$, at $x^\star$ the sufficient conditions for a unconstrained local minimum for the function $L_a(x, \lambda^\star)$ hold, and $g(x^\star) = 0$, then $x^\star$ is a strict constrained local minimum for problem P_1.

The above theorem highlights the fact that, under the stated assumptions, the function $L_a(x, \lambda^\star)$ is an (external) penalty function, with the property that, to give local minima

[7] To be precise we should write $L_a(x, \lambda^\star, \epsilon)$, however we omit the argument ϵ.

for problem P_1 it is not necessary that $\epsilon \to 0$. Unfortunately, this result is not of practical interest, because it requires the knowledge of $\lambda^\star$. To obtain a useful algorithm, it is possible to make use of the following considerations.
By the implicit function theorem, applied to the first of equation (3.16), we infer that there exist a neighborhood of $\lambda^\star$, a neighborhood of $x^\star$, and a continuously differentiable function $x(\lambda)$ such that (locally)

$$\nabla_x L_a(x(\lambda), \lambda) = 0.$$

Moreover, for any $\epsilon \in (0, \bar{\epsilon})$, as $\nabla^2_{xx} L_a(x^\star, \lambda^\star)$ is positive definite also $\nabla^2_{xx} L_a(x, \lambda)$ is locally positive definite. As a result, $x(\lambda)$ is the only value of x that, for any fixed λ, minimizes the function $L_a(x, \lambda)$. It is therefore reasonable to assume that if λ_k is a good estimate of $\lambda^\star$, then the minimization of $L_a(x, \lambda_k)$ for a sufficiently small value of ϵ, yields a point x_k which is a good approximation of $x^\star$.
On the basis of the above discussion it is possible to construct the following minimization algorithm for problem P_1.

Step 0. Given $x_0 \in I\!R^n$, $\lambda_1 \in I\!R^m$ and $\epsilon_1 > 0$.

Step 1. Set $k = 1$.

Step 2. Find a local minimum x_k of $L_a(x, \lambda_k)$ using any unconstrained minimization algorithm, with starting point x_{k-1}.

Step 3. Compute a new estimate λ_{k+1} of $\lambda^\star$.

Step 4. Set $\epsilon_{k+1} = \beta\epsilon_k$, with $\beta = 1$ if $\|g(x_{k+1})\| \le \frac{1}{4}\|g(x_k)\|$ or $\beta < 1$ otherwise.

Step 5. Set $k = k + 1$ and go to **Step 2**.

In **Step 3** it is necessary to construct a new estimate λ_{k+1} of λ_k. This can be done with proper considerations on the function $L_a(x(\lambda), \lambda)$, introduced in the above discussion. However, without providing the formal details, we mention that one of the most used update laws for λ are described by the equations

$$\lambda_{k+1} = \lambda_k + \frac{2}{\epsilon_k} g(x_k), \tag{3.18}$$

or

$$\lambda_{k+1} = \lambda_k - \left[\nabla^2 L_a(x(\lambda_k), \lambda_k)\right]^{-1} g(x_k), \tag{3.19}$$

whenever the indicated inverse exists.
Note that the convergence of the sequence $\{x_k\}$ to $x^\star$ is limited by the convergence of the sequence $\{\lambda_k\}$ to $\lambda^\star$. It is possible to prove that, if the update law (3.18) is used then the algorithm as linear convergence, whereas if (3.19) is used the convergence is superlinear.

Remark. Similar considerations can be done for problem P_2. For, recall that problem P_2 can be recast, increasing the number of variables, as an optimization problem with equality constraints, *i.e.* problem $\tilde{P}_2$. For such an *extended* problem, consider the augmented Lagrangian

$$L_a(x, y, \rho) = f(x) + \sum_{i=1}^{p} \rho_i \left(h_i(x) + y_i^2\right) + \frac{1}{\epsilon} \sum_{i=1}^{p} \left(h_i(x) + y_i^2\right)^2,$$

and note that, in principle, it would be possible to make use of the results developed with reference to problem P_1. However, the function L_a can be analytically minimized with respect to the variables y_i. In fact, a simple computation shows that the (global) minimum of L_a as a function of y is attained at

$$y_i(x, \rho) = \sqrt{-\min\left(0, h_i(x) + \frac{\epsilon}{2}\rho_i\right)}.$$

As a result, the augmented Lagrangian function for problem P_2 is given by

$$L_a(x, \rho) = f(x) + \rho' h(x) + \frac{1}{\epsilon}\|h(x)\|^2 - \frac{1}{\epsilon} \sum_{i=1}^{p} \left(\min(0, h_i(x) + \frac{\epsilon}{2}\rho_i)\right)^2.$$

◇

3.4.3 Exact penalty functions

An exact penalty function, for a given constrained optimization problem, is a function of the same variables of the problem with the property that its unconstrained minimization yields a solution of the original problem. The term *exact* as opposed to *sequential* indicates that only one, instead of several, minimization is required.
Consider problem P_1, let $x^\star$ be a local solution and let $\lambda^\star$ be the corresponding multiplier. The basic idea of exact penalty functions methods is to determine the multiplier λ appearing in the augmented Lagrangian function as a function of x, *i.e.* $\lambda = \lambda(x)$, with $\lambda(x^\star) = \lambda^\star$. With the use of this function one has[8]

$$L_a(x, \lambda(x)) = f(x) + \lambda(x)' g(x) + \frac{1}{\epsilon}\|g(x)\|^2.$$

The function $\lambda(x)$ is obtained considering the necessary condition of optimality

$$\nabla_x L_a(x^\star, \lambda^\star) = \nabla f(x^\star) + \frac{\partial g(x^\star)'}{\partial x}\lambda^\star = 0 \qquad (3.20)$$

and noting that, if $x^\star$ is a regular point for the constraints then equation (3.20) can be solved for $\lambda^\star$ yielding

$$\lambda^\star = -\left(\frac{\partial g(x^\star)}{\partial x}\frac{\partial g(x^\star)'}{\partial x}\right)^{-1} \frac{\partial g(x^\star)}{\partial x} \nabla f(x^\star).$$

[8] As in previous sections we omit the argument ϵ.

This equality suggests to define the function $\lambda(x)$ as

$$\lambda(x) = -\left(\frac{\partial g(x)}{\partial x}\frac{\partial g(x)'}{\partial x}\right)^{-1}\frac{\partial g(x)}{\partial x}\nabla f(x),$$

and this is defined at all x where the indicated inverse exists, in particular at $x^\star$.
It is possible to show that this selection of $\lambda(x)$ yields and exact penalty function for problem P_1. For, consider the function

$$G(x) = f(x) - g(x)'\left(\frac{\partial g(x)}{\partial x}\frac{\partial g(x)'}{\partial x}\right)^{-1}\frac{\partial g(x)}{\partial x}\nabla f(x) + \frac{1}{\epsilon}\|g(x)\|^2,$$

which is defined and differentiable in the set

$$\tilde{\mathcal{X}} = \{x \in I\!R^n \parallel \text{rank}\frac{\partial g(x)}{\partial x} = m\}, \tag{3.21}$$

and the following statements.

Theorem 19 *Let $\bar{\mathcal{X}}$ be a compact subset of $\tilde{\mathcal{X}}$. Assume that $x^\star$ is the only global minimum of f in $\mathcal{X} \cap \bar{\mathcal{X}}$ and that $x^\star$ is in the interior of $\bar{\mathcal{X}}$. Then there exists $\bar{\epsilon} > 0$ such that, for any $\epsilon \in (0, \bar{\epsilon})$, $x^\star$ is the only global minimum of G in $\bar{\mathcal{X}}$.*

Theorem 20 *Let $\bar{\mathcal{X}}$ be a compact subset of $\tilde{\mathcal{X}}$. Then there exists $\bar{\epsilon} > 0$ such that, for any $\epsilon \in (0, \bar{\epsilon})$, if $x^\star$ is a unconstrained minimum of $G(x)$ and $x^\star \in \bar{\mathcal{X}}$, then $x^\star$ is a constrained local minimum for problem P_1.*

Theorems 19 and 20 show that it is legitimate to search solutions of problem P_1 minimizing the function G for sufficiently small values of ϵ. Note that it is possible to prove stronger results, showing that there is (under certain hypotheses) a one to one correspondence between the minima of problem P_1 and the minima of the function G, provided ϵ is sufficiently small.
For problem P_2 it is possible to proceed as discussed in Section 3.4.2, *i.e.* transforming problem P_2 into problem $\tilde{P}_2$ and then minimizing analytically with respect to the new variables y_i. However, a different and more direct route can be taken. Consider problem P_2 and the necessary conditions

$$\nabla_x L(x^\star, \rho^\star) = \nabla f(x^\star) + \frac{\partial h(x^\star)'}{\partial x}\rho^\star = 0 \tag{3.22}$$

and

$$\rho_i^\star h_i(x^\star) = 0, \tag{3.23}$$

for $i = 1, \cdots, p$. Premultiplying equation (3.22) by $\frac{\partial h(x^\star)}{\partial x}$ and equation (3.23) by $\gamma^2 h_i(x^\star)$, with $\gamma > 0$, and adding, yields

$$\left(\frac{\partial h(x^\star)}{\partial x}\frac{\partial h(x^\star)'}{\partial x} + \gamma^2 H^2(x^\star)\right)\rho^\star + \frac{\partial h(x^\star)}{\partial x}\nabla f(x^\star) = 0,$$

where

$$H(x^\star) = \mathrm{diag}(h_1(x^\star), \cdots, h_p(x^\star)).$$

As a result, a natural candidate for the function $\rho(x)$ is

$$\rho(x) = -\left(\frac{\partial h(x)}{\partial x}\frac{\partial h(x)'}{\partial x} + \gamma^2 H^2(x)\right)^{-1} \frac{\partial h(x)}{\partial x}\nabla f(x),$$

which is defined whenever the indicated inverse exists, in particular in the neighborhood of any regular point. With the use of this function, it is possible to define an exact penalty function for problem P_2 and to establish results similar to those illustrated in Theorems 19 and 20.

The exact penalty functions considered in this section provide, in principle, a theoretically sound way of solving constrained optimization problem. However, in practice, they have two major drawbacks. Firstly, at each step, it is necessary to invert a matrix with dimension equal to the number of constraint. This operation is numerically ill conditioned if the number of constraints is large. Secondly, the exact penalty functions may not be sufficiently regular to allow the use of unconstrained minimization methods with fast speed of convergence, *e.g.* Newton method.

3.4.4 Exact augmented Lagrangian functions

An exact augmented Lagrangian function, for a given constrained optimization problem, is a function, defined on an augmented space with dimension equal to the number of variables plus the number of constraint, with the property that its unconstrained minimization yields a solution of the original problem. Moreover, in the computation of such a function it is not necessary to invert any matrix.

To begin with, consider problem P_1 and recall that, for such a problem, a sequential augmented Lagrangian function has been defined adding to the Lagrangian function a term, namely $\frac{1}{\epsilon}\|g(x)\|^2$, which penalizes the violation of the necessary condition $g(x) = 0$. This term, for ϵ sufficiently small, renders the function L_a convex in a neighborhood of $x^\star$. A *complete* convexification can be achieved adding a further term that penalizes the violation of the necessary condition $\nabla_x L(x, \lambda) = 0$. More precisely, consider the function

$$S(x, \lambda) = f(x) + \lambda' g(x) + \frac{1}{\epsilon}\|g(x)\|^2 + \eta\|\frac{\partial g(x)}{\partial x}\nabla_x L(x, \lambda)\|^2, \tag{3.24}$$

with $\epsilon > 0$ and $\eta > 0$. The function (3.24) is continuously differentiable and it is such that, for ϵ sufficiently small, the solutions of problem P_1 are in a one to one correspondence with the points (x, λ) which are local minima of S, as detailed in the following statements.

Theorem 21 *Let $\bar{\mathcal{X}}$ be a compact set. Suppose $x^\star$ is the unique global minimum of f in the set $\mathcal{X} \cap \bar{\mathcal{X}}$ and $x^\star$ is an interior point of $\bar{\mathcal{X}}$. Let $\lambda^\star$ be the multiplier associated to $x^\star$. Then, for any compact set $\Lambda \subset I\!R^m$ such that $\lambda^\star \in \Lambda$ there exists $\bar{\epsilon}$ such that, for all $\epsilon \in (0, \bar{\epsilon})$, $(x^\star, \lambda^\star)$ is the unique global minimum of S in $\mathcal{X} \times \Gamma$.*

Theorem 22 *Let*[9] $\mathcal{X} \times \Lambda \subset \tilde{\mathcal{X}} \times I\!R^m$ *be a compact set. Then there exists* $\bar{\epsilon} > 0$ *such that, for all* $\epsilon \in (0, \bar{\epsilon})$*, if* $(x^\star, \lambda^\star)$ *is a unconstrained local minimum of* S*, then* $x^\star$ *is a constrained local minimum for problem* P_1 *and* $\lambda^\star$ *is the corresponding multiplier.*

Theorems 21 and 22 justify the use of a unconstrained minimization algorithm, applied to the function S, to find local (or global) solutions of problem P_1.
Problem P_2 can be dealt with using the same considerations done in Section 3.4.2.

3.5 Recursive quadratic programming

Recursive quadratic programming methods have been widely studied in the past years. In this section we provide a preliminary description of such methods. For, consider problem P_1 and suppose that $x^\star$ and $\lambda^\star$ are such that the necessary conditions (3.7) hold. Consider now a series expansion of the function $L(x, \lambda^\star)$ in a neighborhood of $x^\star$, *i.e.*

$$L(x, \lambda^\star) = f(x^\star) + \frac{1}{2}(x - x^\star)' \nabla^2_{xx} L(x^\star, \lambda^\star)(x - x^\star) + ...$$

a series expansion of the constraint, *i.e.*

$$0 = g(x) = g(x^\star) + \frac{\partial g(x^\star)}{\partial x}(x - x^\star) + ...$$

and the problem

$$\widetilde{PQ}_1 \quad \begin{cases} \min_x f(x^\star) + \frac{1}{2}(x - x^\star)' \nabla^2_{xx} L(x^\star, \lambda^\star)(x - x^\star) \\ \frac{\partial g(x^\star)}{\partial x}(x - x^\star) = 0. \end{cases}$$

Note that problem $\widetilde{PQ}_1$ has the solution $x^\star$, and the corresponding multiplier is $\lambda = 0$, which is not equal (in general) to $\lambda^\star$. This phenomenon is called *bias* of the multiplier, and can be avoided by modifying the objective function and considering the new problem

$$PQ_1 \quad \begin{cases} \min_x f(x^\star) + \nabla f(x^\star)'(x - x^\star) + \frac{1}{2}(x - x^\star)' \nabla^2_{xx} L(x^\star, \lambda^\star)(x - x^\star) \\ \frac{\partial g(x^\star)}{\partial x}(x - x^\star) = 0, \end{cases} \tag{3.25}$$

which has solution $x^\star$ with multiplier $\lambda^\star$. This observation suggests to consider the sequence of quadratic programming problems

$$PQ_1^{k+1} \quad \begin{cases} \min_\delta f(x_k) + \nabla f(x_k)'\delta + \frac{1}{2}\delta' \nabla^2_{xx} L(x_k, \lambda_k)\delta \\ g(x_k) + \frac{\partial g(x_k)}{\partial x}\delta = 0, \end{cases} \tag{3.26}$$

[9] The set $\tilde{\mathcal{X}}$ is defined as in equation (3.21).

where $\delta = x - x_k$, and x_k and λ_k are the current estimates of the solution and of the multiplier. The solution of problem PQ_1^{k+1} yields new estimates x_{k+1} and λ_{k+1}. To assess the local convergence of the method, note that the necessary conditions for problem PQ_1^{k+1} yields the system of equations

$$\begin{bmatrix} \nabla^2_{xx} L(x_k, \lambda_k) & \dfrac{\partial g(x_k)'}{\partial x} \\ \dfrac{\partial g(x_k)}{\partial x} & 0 \end{bmatrix} \begin{bmatrix} \delta \\ \lambda \end{bmatrix} = - \begin{bmatrix} \nabla f(x_k) \\ g(x_k) \end{bmatrix}, \tag{3.27}$$

and this system coincides with the system arising from the application of Newton method to the solution of the necessary conditions for problem P_1. As a consequence, the solutions of the problems PQ_1^{k+1} converge to a solution of problem P_1 under the same hypotheses that guarantee the convergence of Newton method.

Theorem 23 *Let $x^\star$ be a strict constrained local minimum for problem P_1, and let $\lambda^\star$ be the corresponding multiplier. Suppose that for $x^\star$ and $\lambda^\star$ the sufficient conditions of Theorem 15 hold. Then there exists an open neighborhood $\Omega \subset I\!R^n \times I\!R^m$ of the point $(x^\star, \lambda^\star)$ such that, if $(x_0, \lambda_0) \in \Omega$, the sequence $\{x_k, \lambda_k\}$ obtained solving the sequence of quadratic programming problems PQ_1^{k+1}, with $k = 0, 1, \cdots$, converges to $(x^\star, \lambda^\star)$.*
Moreover, the speed of convergence is superlinear, and, if f and g are three times differentiable, the speed of convergence is quadratic.

Remark. It is convenient to solve the sequence of quadratic programming problems PQ_1^{k+1}, instead of solving the equations (3.27) with Newton method, because, for the former it is possible to exclude converge to maxima or saddle points. $\diamond$

In the case of problem P_2, using considerations similar to the one above, it is easy to obtain the following sequence of quadratic programming problems

$$PQ_2^{k+1} \quad \begin{cases} \min\limits_{\delta} f(x_k) + \nabla f(x_k)'\delta + \dfrac{1}{2}\delta' \nabla^2_{xx} L(x_k, \lambda_k)\delta \\ \dfrac{\partial h(x_k)}{\partial x}\delta + h(x_k) \leq 0. \end{cases} \tag{3.28}$$

This sequence of problems has to be solved iteratively to generate a sequence $\{x_k, \lambda_k\}$ that, under hypotheses similar to those of Theorem 23, converges to a strict constrained local minimum of problem P_2.
The method described are the basis for a large class of iterative methods.
A first disadvantage of the proposed iterative schemes is that it is necessary to compute the second derivatives of the functions of the problem. This computation can be avoided, using the same philosophy of quasi-Newton methods.
A second disadvantage is in the fact that, being based on Newton algorithm, only local convergence can be guaranteed. However, it is possible to combine the method with global convergent methods: these are used to generate a pair $(\tilde{x}, \tilde{\lambda})$ sufficiently close to $(x^\star, \lambda^\star)$

and then recursive quadratic programming methods are used to obtain fast convergence to $(x^\star, \lambda^\star)$.
A third disadvantage is in the fact that there is no guarantee that the sequence of admissible sets generated by the algorithm is non-empty at each step.
Finally, it is worth noting that, contrary to most of the existing methods, quadratic programming methods do not rely on the use of a penalty term.

Remark. There are several alternative recursive quadratic programming methods which alleviate the drawbacks of the methods described. These are (in general) based on the use of quadratic approximation of penalty functions. For brevity, we do not discuss these methods. $\diamond$

3.6 Concluding remarks

In this section we briefly summarize advantages and disadvantages of the nonlinear programming methods discussed.
Sequential penalty functions methods are very simple to implement, but suffer from the ill conditioning associated to large penalties (*i.e.* to small values of ϵ). As a result, these methods can be used if approximate solutions are acceptable, or in the determination of initial estimates for more precise, but only locally convergent, methods. Note, in fact, that not only an approximation of the solution $x^\star$ can be obtained, but also an approximation of the corresponding multiplier $\lambda^\star$. For example, for problem P_1, a (approximate) solution $\bar{x}$ is such that

$$\nabla F_{\epsilon_k}(\bar{x}) = \nabla f(\bar{x}) + \frac{2}{\epsilon_k}\frac{\partial g(\bar{x})}{\partial x} g(\bar{x}) = 0,$$

hence, the term $\frac{2}{\epsilon_k} g(\bar{x})$ provides an approximation of $\lambda^\star$.
Sequential augmented Lagrangian functions do not suffer from ill conditioning, and yield faster speed of convergence then that attainable using sequential penalty functions.
The methods based on exact penalty functions do not require the solution of a sequence of problems. However, they require the inversion of a matrix of dimension equal to the number of constraints, hence their applicability is limited to problems with a small number of constraints.
Exact augmented Lagrangian functions can be built without inverting any matrix. However, the minimization has to be performed in an extended space.
Recursive quadratic programming methods require the solution, at each step, of a constrained quadratic programming problem. Their main problem is that there is no guarantee that the admissible set is non-empty at each step.
We conclude that it is not possible to decide which is the best method. Each method has its own advantages and disadvantages. Therefore, the selection of a nonlinear programming method has to be driven by the nature of the problem: and has to take into consideration the number of variables, the regularity of the involved functions, the required precision, the computational cost, ...

3.7 Exercises

Similarly to Section 2.10, this section contains a set of exercises related to the notions, concepts, algorithms and tools discussed in Chapter 3. Since nonlinear programming is a widely studied, and complex, area of optimization, there are several methods and algorithms which are introduced only via exercises and which are not covered in details in the text. It is my hope that the exercises could form the starting point for additional reading.

Exercise 28 Consider the minimization problem

$$\begin{cases} \min\limits_{x_1,x_2} 1 - x_1^2 - x_2^2, \\ x_1 \geq 0, \\ x_2 \geq 0, \\ x_1 + x_2 - 1 \leq 0. \end{cases}$$

a) Show that all points in the admissible set are regular points for the constraints.

b) State the first order necessary conditions of optimality for such a constrained optimization problem.

c) Using the conditions derived in part b), compute candidate optimal solutions.

d) Show that the admissible set is compact. Hence deduce the existence of a global minimizer for the optimization problem. Determine the global minimizer of the problem. Is this minimizer unique?

Solution 28

a) The admissible set is the shaded area in the figure below. The arrows denote the gradient of the constraints on the boundary of the admissible set. As can be seen, these vectors are always independent, therefore all points are regular points for the constraints.

b) Define the Lagrangian

$$L = 1 - x_1^2 - x_2^2 + \rho_1(-x_1) + \rho_2(-x_2) + \rho_3(x_1 + x_2 - 1).$$

The first order necessary conditions of optimality are

$$\begin{array}{lll} -2x_1 - \rho_1 + \rho_3 = 0, & -2x_2 - \rho_2 + \rho_3 = 0, & \\ -x_1 \leq 0, & -x_2 \leq 0, & x_1 + x_2 - 1 \leq 0, \\ \rho_1 \geq 0, & \rho_2 \geq 0, & \rho_3 \geq 0, \\ -\rho_1 x_1 = 0, & -\rho_2 x_2 = 0, & \rho_3(x_1 + x_2 - 1) = 0. \end{array}$$

c) To compute candidate optimal solutions, note that from the last line of the necessary conditions we have the following possibilities:

$$\begin{array}{ll} \text{P1: } \rho_1 = 0,\ \rho_2 = 0,\ \rho_3 = 0; & \text{P2: } \rho_1 = 0,\ \rho_2 = 0,\ \rho_3 > 0; \\ \text{P3: } \rho_1 = 0,\ \rho_2 > 0,\ \rho_3 = 0; & \text{P4: } \rho_1 = 0,\ \rho_2 > 0,\ \rho_3 > 0; \\ \text{P5: } \rho_1 > 0,\ \rho_2 = 0,\ \rho_3 = 0; & \text{P6: } \rho_1 > 0,\ \rho_2 = 0,\ \rho_3 > 0; \\ \text{P7: } \rho_1 > 0,\ \rho_2 > 0,\ \rho_3 = 0; & \text{P8: } \rho_1 > 0,\ \rho_2 > 0,\ \rho_3 > 0. \end{array}$$

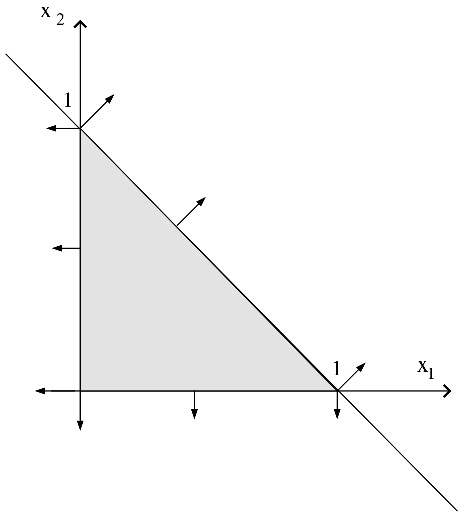
This yields the following candidate points

$$\begin{array}{cc} \text{P1: } x_1 = x_2 = 0 \text{ hence } f = 1; & \text{P2: } x_1 = x_2 = 1/2 \text{ hence } f = 1/2; \\ \text{P3: unfeasible;} & \text{P4: } x_1 = 1,\ x_2 = 0, \text{ hence } f = 0; \\ \text{P5: unfeasible;} & \text{P6: } x_1 = 0,\ x_2 = 1, \text{ hence } f = 0; \\ \text{P7: unfeasible;} & \text{P8: unfeasible.} \end{array}$$

As a result, we have only four candidate points:

$$P_1 = (0,0), \qquad P_2 = (1/2, 1/2), \qquad P_4 = (1,0), \qquad P_6 = (0,1).$$

d) The admissible set is closed (the constraints include the equality sign) and bounded (see the figure), hence compact. By Weierstrass theorem the function f has a global minimum in such a set. The function f attains its global minimum at P_4 and P_6, which are therefore both global minimizers. (This can be also shown noting that the problem is symmetric, i.e. changing x_1 into x_2 and x_2 into x_1 yields the same problem.)



Exercise 29 Consider the optimization problem

$$\begin{cases} \min\limits_{x_1,x_2} -x_1 - x_2 \\ x_1^2 + x_2^2 - 1 = 0 \end{cases}$$

a) Transform this minimization problem into an unconstrained minimization problem using the method of sequential penalty functions.

b) State the necessary conditions of optimality for the unconstrained problem of part a). Hence compute approximate candidate optimal solutions for the unconstrained optimization problem. Discuss the feasibility of these candidate optimal solutions.
(Hint: you may show that optimal points of the unconstrained problem are such that $x_1^\star = x_2^\star$. Moreover, use the fact that the solutions of $1 + 4x\dfrac{1-2x^2}{\epsilon} = 0$, for ϵ positive and small, are $\frac{\sqrt{2}}{2} + \frac{1}{8}\epsilon$, $-\frac{\sqrt{2}}{2} + \frac{1}{8}\epsilon$, $-\frac{1}{4}\epsilon$.)

c) Consider the stationary points of the sequential penalty function in part b). Consider the limit for $\epsilon \to 0$ of these stationary points and thus determine candidate optimal solutions for the original constrained optimization problem.

Solution 29

a) A sequential penalty function for the constrained problem is

$$F_\epsilon = -x_1 - x_2 + \frac{1}{\epsilon}(x_1^2 + x_2^2 - 1)^2.$$

b) The necessary conditions of optimality for F_ϵ are

$$0 = \nabla F_\epsilon = \begin{bmatrix} -1 + \frac{4x_1}{\epsilon}(x_1^2 + x_2^2 - 1) \\ \\ -1 + \frac{4x_2}{\epsilon}(x_1^2 + x_2^2 - 1) \end{bmatrix}.$$

As a result,

$$\frac{1}{x_1} = \frac{4}{\epsilon}(x_1^2 + x_2^2 - 1)$$
$$\frac{1}{x_2} = \frac{4}{\epsilon}(x_1^2 + x_2^2 - 1)$$

yielding $x_1 = x_2$. Let $x_1 = x_2 = x$. From the first equation we have

$$\frac{1}{x} = \frac{4}{\epsilon}(2x^2 - 1) \Rightarrow 1 + 4x\frac{1 - 2x^2}{\epsilon} = 0.$$

As stated, this equation has approximate solutions

$$\frac{\sqrt{2}}{2} + \frac{1}{8}\epsilon, \qquad -\frac{\sqrt{2}}{2} + \frac{1}{8}\epsilon, \qquad -\frac{1}{4}\epsilon.$$

As a result, F_ϵ has three stationary points:

$$P_1 \approx \left(\frac{\sqrt{2}}{2} + \frac{1}{8}\epsilon, \frac{\sqrt{2}}{2} + \frac{1}{8}\epsilon\right), \qquad P_2 \approx \left(\frac{-\sqrt{2}}{2} + \frac{1}{8}\epsilon, \frac{-\sqrt{2}}{2} + \frac{1}{8}\epsilon\right), \qquad P_3 \approx \left(-\frac{1}{4}\epsilon, -\frac{1}{4}\epsilon\right).$$

Note that none of the above points is feasible, for any $\epsilon > 0$.

c) The stationary points of F_ϵ are such that

$$\lim_{\epsilon\to 0} P_1 = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \qquad \lim_{\epsilon\to 0} P_2 = \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right), \qquad \lim_{\epsilon\to 0} P_3 = (0, 0).$$

Hence, P_1 and P_2 converge to the admissible set, and P_1 is a (local) solution of the optimization problem considered.

Exercise 30 Consider the optimization problem

$$\begin{cases} \min\limits_{x_1,x_2} x_1^2 - x_2^2, \\ \\ x_1 - x_2^2 = 0. \end{cases}$$

a) Sketch in the (x_1, x_2)-plane the level lines of the function to be minimized and the admissible set. (Hint: plot the level lines corresponding to $x_1^2 - x_2^2 = 0$ and $x_1^2 - x_2^2 = \pm 4$.)

b) Using first order necessary conditions, compute candidate optimal solutions. Use second order sufficient conditions to decide which of the candidate points is a local minimizer or a local maximizer.

c) Compute an exact penalty function for the minimization problem and verify that the candidate optimal solutions determined in part b) are stationary points of the exact penalty function.

Solution 30

a) The level sets and the admissible set are depicted in the figure below.

b) Let

$$L(x, \lambda) = x_1^2 - x_2^2 + \lambda(x_1 - x_2^2).$$

The first order necessary conditions are

$$2x_1 + \lambda = 0, \qquad -2x_2 - 2\lambda x_2 = 0, \qquad x_1 - x_2^2 = 0,$$

and these yield the candidate optimal points

$$P_1 = (0,0), \qquad P_2 = (1/2, \sqrt{2}/2), \qquad P_3 = (1/2, -\sqrt{2}/2),$$

with corresponding multipliers $\lambda_1 = 0, \lambda_2 = -1, \lambda_3 = -1$. The second order sufficient conditions are $s'\nabla^2_{xx}L(x^\star, \lambda^\star)s > 0$ for $s \neq 0$ such that $[1, -2x_2^\star]s = 0$. For P_1 one has $s = [0,1]'$ and $s'(\nabla^2_{xx}L)s < 0$, hence P_1 is a local maximizer. For P_2 one has $s = [\sqrt{2}, 1]'$ and $s'(\nabla^2_{xx}L)s = 4$, and for P_3 one has $s = [\sqrt{2}, -1]'$ and $s'(\nabla^2_{xx})Ls = 4$. Hence, P_2 and P_3 are local minimizers.

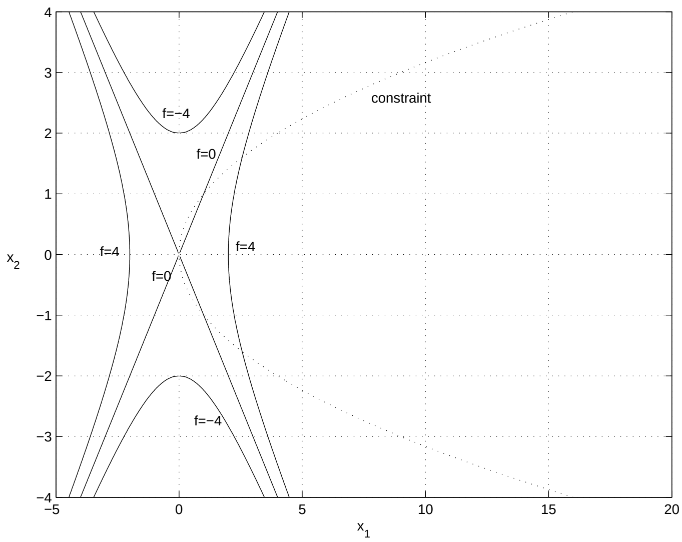
c) An exact penalty function for the considered problem is

$$G(x_1, x_2) = x_1^2 - x_2^2 - \frac{(2x_1 + 4x_2^2)(x_1 - x_2^2)}{1 + 4x_2^2} + \frac{(x_1 - x_2^2)^2}{\epsilon},$$

with $\epsilon > 0$. Its stationary points are the solutions of

$$0 = \nabla G(x_1, x_2) = \begin{bmatrix} 2\dfrac{-x_1\epsilon + 4x_1\epsilon x_2^2 - \epsilon x_2^2 + x_1 + 4x_1x_2^2 - x_2^2 - 4x_2^4}{(1+4x_2^2)\epsilon} \\ \\ 2x_2\dfrac{-\epsilon + 8\epsilon x_1^2 - 2x_1\epsilon - 2x_1 - 16x_1x_2^2 - 32x_1x_2^4 + 2x_2^2 + 16x_2^4 + 32x_2^6}{(1+4x_2^2)^2\epsilon} \end{bmatrix}.$$

By direct substitution we verify that, for any $\epsilon > 0$, the points P_1, P_2 and P_3 are stationary points of G.



Exercise 31 Consider the minimization problem

$$\begin{cases} \min\limits_{x_1, x_2} -x_1x_2, \\ 0 \leq x_1 + x_2 \leq 2, \\ -2 \leq x_1 - x_2 \leq 2, \end{cases}$$

a) Sketch in the (x_1, x_2)-plane the level surfaces of the function to be minimized and the admissible set. Hence show that all points in the admissible set are regular points for the constraints.

b) Using only graphical considerations determine the global solution of the considered problem.

c) State first order necessary conditions of optimality for such a constrained optimization problem. Show that the point determined in part b) satisfies first order necessary conditions of optimality, for some selection of the multiplier ρ.

d) Show that the point determined in part b) satisfies second order sufficient conditions of optimality for such a constrained optimization problem.

Solution 31

a) The admissible set is the shaded area in the figure below. The arrows denote the gradient of the constraints on the boundary of the admissible set. As can be seen, these vectors are always independent, therefore all points are regular points for the constraints. The dashed lines represent level lines of the function f.

b) From the figure it can be seen that the minimum is achieved when the level line of the function f is tangent to the admissible set, *i.e.* at the point $p = (1, 1)$.

c) Consider the Lagrangian

$$L = -x_1x_2 + \rho_1(-x_1 - x_2) + \rho_2(x_1 + x_2 - 2) + \rho_3(-2 - x_1 + x_2) + \rho_4(x_1 - x_2 - 2).$$

The first order sufficient conditions of optimality are

$$0 = \nabla_x L = \begin{bmatrix} -x_2 - \rho_1 + \rho_2 - \rho_3 + \rho_4 \\ -x_1 - \rho_1 + \rho_2 + \rho_3 - \rho_4 \end{bmatrix}$$

$$-x_1 - x_2 \le 0, \qquad x_1 + x_2 - 2 \le 0, \qquad -x_1 - x_2 - 2 \le 0, \qquad x_1 - x_2 - 2 \le 0,$$

$$\rho_1 \ge 0, \qquad \rho_2 \ge 0, \qquad \rho_3 \ge 0, \qquad \rho_4 \ge 0,$$

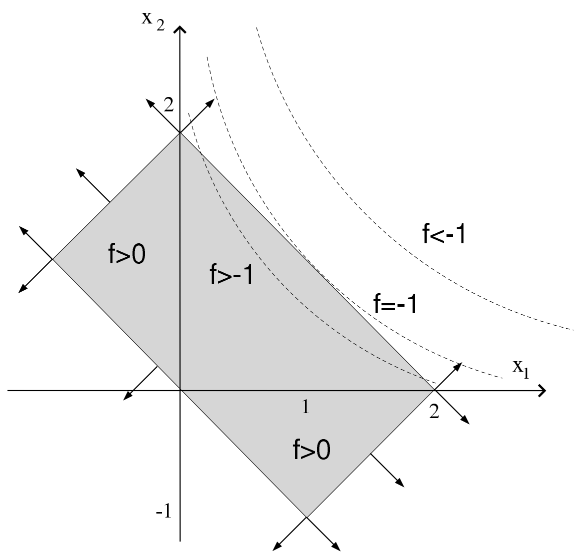
$$\rho_1(-x_1-x_2) = 0, \qquad \rho_2(x_1+x_2-2) = 0, \qquad \rho_3(-2-x_1+x_2) = 0, \qquad \rho_4(x_1-x_2-2) = 0.$$

Setting $(x_1, x_2) = (1, 1)$ and selecting $\rho_1 = 0$, $\rho_3 = 0$, $\rho_4 = 0$ satisfies all the above equations. Hence, the point $(x_1, x_2) = (1, 1)$, together with the given multipliers, satisfies first order necessary conditions of optimality.

d) To check second order sufficient conditions note that for $(x_1, x_2) = (1, 1)$ the only active constraint is $x_1 + x_2 - 2 \le 0$. Therefore we need to check positivity of $s'\nabla^2_{xx}Ls$ for $s = [s_1 \ s_2]'$ such that $[1 \ 1]s = 0$. This means $s_1 + s_2 = 0$, hence, solving for s_2, one has

$$s'\nabla^2_{xx}Ls = \begin{bmatrix} s_1 & -s_1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} s_1 \\ -s_1 \end{bmatrix} = 2s_1^2 > 0$$

for $s_1 \neq 0$. As a result, the point obtained from graphical considerations in part b) is indeed a local minimizer for the considered problem.



Exercise 32 Consider the optimization problem

$$\begin{cases} \min\limits_{x_1,x_2,x_3} Tx_1^2 + Tx_2^2 + x_3^2, \\ \\ x_1 + x_2 + x_3 - 1 = 0. \end{cases}$$

a) Transform this minimization problem into an unconstrained minimization problem by solving the constraint equation for x_3 and substituting the solution into the objective function.

b) Assume $T > 0$. Consider the unconstrained minimization problem determined in part a). Find (the unique) candidate optimal solution and show that this is indeed a local minimizer.

c) Assume $T > 0$. Exploit the results in parts a) and b) to determine the solution of the constrained optimization problem.

d) Assume $T > 0$. Consider the so-called l_1 penalty function

$$f_p = Tx_1^2 + Tx_2^2 + x_3^2 + \frac{|x_1 + x_2 + x_3 - 1|}{\epsilon},$$

with $\epsilon > 0$ and sufficiently small. Show that the unique stationary point of f_p coincides with the optimal solution determined in part c).
(Hint: recall that that $\frac{d|x|}{dx} = \text{sign}(x)$ and that $\text{sign}(0) \in [-1, 1]$. Moreover, use the fact that the stationary points of f_p do not depend upon the parameter ϵ.)

Solution 32

a) Solving the constrain equation for x_3 yields $x_3 = 1 - x_1 - x_2$. This is replaced in the function to minimize, hence resulting in the unconstrained minimization problem

$$\min_{x_1,x_2} \tilde{f}$$

with

$$\tilde{f} = Tx_1^2 + Tx_2^2 + (1 - x_1 - x_2)^2.$$

b) To determine candidate optimal solution consider the equations

$$0 = \nabla \tilde{f} = \begin{bmatrix} 2Tx_1 + 2x_1 + 2x_2 - 2 \\ 2Tx_2 + 2x_1 + 2x_2 - 2. \end{bmatrix}$$

These have the unique solution

$$x_1^\star = \frac{1}{T+2}, \qquad x_2^\star = \frac{1}{T+2}.$$

Note now that

$$\nabla^2 \tilde{f} = \begin{bmatrix} 2T+2 & 2 \\ 2 & 2T+2 \end{bmatrix}$$

and this is positive definite for $T > 0$. Hence, the obtained stationary point is a local minimizer for $\tilde{f}$.

c) To obtain a solution of the original problem it is enough to compute

$$x_3^\star = 1 - x_1^\star - x_2^\star = \frac{T}{T+2}.$$

d) To compute the stationary points of f_p consider the equations

$$0 = \nabla f_p = \begin{bmatrix} 2Tx_1 + \dfrac{\text{sign}(x_1 + x_2 + x_3 - 1)}{\epsilon} \\ 2Tx_2 + \dfrac{\text{sign}(x_1 + x_2 + x_3 - 1)}{\epsilon} \\ 2x_3 + \dfrac{\text{sign}(x_1 + x_2 + x_3 - 1)}{\epsilon} \end{bmatrix}.$$

These can be rewritten as

$$2Tx_1 = 2Tx_2 = 2x_3 = -\frac{\text{sign}(x_1 + x_2 + x_3 - 1)}{\epsilon}$$

yielding $x_1 = x_3/T$ and $x_2 = x_3/T$. Replacing this in the last equation yields

$$2x_3 = -\frac{\text{sign}(x_3/T + x_3/T + x_3 - 1)}{\epsilon}.$$

Note now that the solution of this equation may be independent of ϵ only if $x_3/T+x_3/T+x_3-1 = 0$, implying $x_3 = x_3^\star$. Finally, this implies that $x_1 = x_1^\star$ and $x_2 = x_2^\star$, *i.e.* the unique stationary point of f_p coincides with the optimal solution obtained in part c).

Exercise 33 Consider the optimization problem

$$\begin{cases} \min\limits_{x_1,x_2} x_1x_2, \\ x_1^2 + x_2^2 - 1 \leq 0. \end{cases}$$

a) State first order necessary conditions of optimality for such a constrained optimization problem.

b) Using the conditions derived in part a), compute candidate optimal solutions.

c) This constrained optimization problem can be transformed into an unconstrained optimization problem by defining the so-called barrier function

$$B_\epsilon(x) = x_1x_2 + \frac{\epsilon}{1 - x_1^2 - x_2^2},$$

with $\epsilon > 0$, and considering the unconstrained minimization of $B_\epsilon(x)$. Determine the stationary points x_ϵ of $B_\epsilon(x)$.
(Hint: show that all stationary points $\bar{x} = (\bar{x}_1, \bar{x}_2)$ are such that $\bar{x}_1 = -\bar{x}_2$, and then note that the solutions of the equation

$$x - \frac{2\epsilon x}{(2x^2 - 1)^2} = 0$$

are $x = 0$ and $x = \pm\frac{\sqrt{2\pm2\sqrt{2\epsilon}}}{2}$.)
Discuss the feasibility of the obtained stationary points x_ϵ. Compute $\lim\limits_{\epsilon\to0} x_\epsilon$ and compare this result with the results obtained in parts a) and b).

d) Discuss the advantages and disadvantages of the proposed barrier function method in comparison with the sequential penalty function method discussed in Section 3.4.1.

Solution 33

a) The Lagrangian of the problem is

$$L = x_1x_2 + \rho(x_1^2 + x_2^2 - 1).$$

The first order necessary conditions of optimality are

$$0 = \nabla_x L = \begin{bmatrix} x_2 + 2\rho x_1 \\ x_1 + 2\rho x_2 \end{bmatrix},$$

$$x_1^2 + x_2^2 - 1 \leq 0, \qquad \rho \geq 0, \qquad (x_1^2 + x_2^2 - 1)\rho = 0.$$

b) From the first two equations we have that if $\rho \neq 1/2$ then $x_1 = x_2 = 0$. If $\rho = 1/2$ then $x_1 + x_2 = 0$ and from the last equation $x_1^2 + x_2^2 - 1 = 0$. As a result $x_1 = \pm\frac{1}{\sqrt{2}}$ and $x_2 = \mp\frac{1}{\sqrt{2}}$. In conclusion we have three candidate solutions

$$P_1 = (0, 0), \qquad P_2 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \qquad P_3 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right).$$

c) To determine the stationary points of B_ϵ, consider the equation

$$0 = \nabla_x B_\epsilon = \begin{bmatrix} x_2 + 2\epsilon \dfrac{x_1}{(1 - x_1^2 - x_2^2)^2} \\ x_1 + 2\epsilon \dfrac{x_2}{(1 - x_1^2 - x_2^2)^2} \end{bmatrix}.$$

From these we obtain

$$\tfrac{x_2}{x_1} = -2\epsilon \frac{1}{(1 - x_1^2 - x_2^2)^2},$$
$$\tfrac{x_1}{x_2} = -2\epsilon \frac{1}{(1 - x_1^2 - x_2^2)^2},$$

hence $x_1/x_2 < 0$ and $x_2/x_1 = x_1/x_2$. As a result $x_1 = -x_2$. Replacing in the second equation yields

$$x_1 = 2\epsilon \frac{x_1}{(1 - 2x_1^2)^2}$$

hence we obtain five candidate solutions, namely

$$P_a = (0,0), \qquad P_b = \left(\frac{\sqrt{2 + 2\sqrt{2\epsilon}}}{2}, -\frac{\sqrt{2 + 2\sqrt{2\epsilon}}}{2} \right), \qquad P_c = \left(-\frac{\sqrt{2 + 2\sqrt{2\epsilon}}}{2}, \frac{\sqrt{2 + 2\sqrt{2\epsilon}}}{2} \right),$$

$$P_d = \left(\frac{\sqrt{2 - 2\sqrt{2\epsilon}}}{2}, -\frac{\sqrt{2 - 2\sqrt{2\epsilon}}}{2} \right), \qquad P_e = \left(-\frac{\sqrt{2 - 2\sqrt{2\epsilon}}}{2}, \frac{\sqrt{2 - 2\sqrt{2\epsilon}}}{2} \right).$$

Note that P_a, P_d and P_e are feasible, whereas P_b and P_c are not feasible. Finally $P_a = P_1$,

$$\lim_{\epsilon \to 0} P_b = \lim_{\epsilon \to 0} P_d = P_2$$

and

$$\lim_{\epsilon \to 0} P_c = \lim_{\epsilon \to 0} P_e = P_3.$$

d) The proposed method is preferable to the sequential penalty function method because it provides feasible solutions also for $\epsilon > 0$. However, the function B_ϵ is not defined on all $I\!R^2$, hence it may be difficult to perform a numerical minimization.

Exercise 34 Let $Q \in I\!R^{n\times n}$ with $Q = Q' > 0$, $x \in I\!R^n$, $A \in I\!R^{m\times n}$ $b \in I\!R^m$ and $y \in I\!R^m$. Consider the minimization problem

$$P : \begin{cases} \min\limits_x \dfrac{1}{2} x'Qx, \\ Ax - b \le 0, \end{cases}$$

and the so-called dual problem

$$D : \begin{cases} \min\limits_y \dfrac{1}{2} y'AQ^{-1}A'y + b'y, \\ -y \le 0. \end{cases}$$

a) Write first order necessary conditions of optimality for the problem P. (Denote the multiplier with ρ.)

b) Write first order necessary conditions of optimality for the problem D. (Denote the multiplier with σ.)

c) Let $y_\star$ and $\sigma_\star$ be such that the optimality conditions in part b) hold. Show that

$$x_\star = -Q^{-1}A'y_\star, \qquad \rho_\star = y_\star,$$

are such that the optimality conditions in part a) hold.

d) Consider the minimization problem

$$P_1 : \begin{cases} \min\limits_x \dfrac{1}{2}x'x, \\ x_1 + 1 \leq 0, \end{cases}$$

with $x \in I\!R^n$ and $x = [x_1, x_2, \cdots, x_n]'$. Exploiting the results above solve this problem. (Hint: write the dual D_1 of problem P_1, solve problem D_1, and then obtain a solution to problem P_1 exploiting the results in part c).)

Solution 34

a) Let $L_P = \dfrac{1}{2}x'Qx + \rho'(Ax - b)$ be the Lagrangian for problem P. The first order necessary conditions of optimality for problem P are

$$Qx_\star + A'\rho_\star = 0, \qquad Ax_\star - b \leq 0, \qquad \rho_\star \geq 0, \qquad \rho_\star'(Ax_\star - b) = 0.$$

b) Let $L_D = \dfrac{1}{2}y'AQ^{-1}A'y + b'y + \sigma'(-y)$ be the Lagrangian for problem D. The first order necessary conditions of optimality for problem D are

$$AQ^{-1}A'y_\star + b - \sigma_\star = 0, \qquad -y_\star \leq 0, \qquad \sigma_\star \geq 0, \qquad \sigma_\star'(-y_\star) = 0.$$

c) Replacing $x_\star = -Q^{-1}A'y_\star$ and $\rho_\star = y_\star$ in the equations in part a) yields

$$\begin{aligned} Q(-Q^{-1}A'y_\star) + A'y_\star &= 0, \\ A(-Q^{-1}A'y_\star) - b &\leq 0, \\ y_\star &\geq 0, \\ y_\star'(A(-Q^{-1}A'y_\star) - b) &= 0. \end{aligned}$$

The first of the above equations holds trivially. For the second one note that

$$A(-Q^{-1}A'y_\star) - b = -\sigma_\star \leq 0,$$

by the third of the equations in b). The third equation holds by the second of the equations in b). The fourth equation holds exploiting the first and the fourth of the equations in b), hence we conclude the claim.

d) Problem P_1 is of the form of problem P with $Q = I$, $A = [1, 0, \cdots, 0]$ and $b = -1$. Hence, the dual D_1 is

$$D_1 : \begin{cases} \min\limits_y \dfrac{1}{2}y^2 - y, \\ -y \leq 0, \end{cases}$$

with $y \in I\!R$. The problem D_1 has the solution $y_\star = 1$ and $\sigma_\star = 0$. Hence, the solution to problem P_1 is

$$x_\star = -[1, 0, \cdots, 0]', \qquad \rho_\star = 1.$$

Exercise 35 Consider the optimization problem

$$\begin{cases} \min\limits_{x_1,\ x_2} x_1^2 + x_2, \\ x_1^2 + (x_2 - 1)^2 = 4. \end{cases}$$

a) Sketch in the (x_1, x_2)-plane the level lines of the function to be minimized and the admissible set. Hence show that all points in the admissible set are regular points for the constraints.

b) Using only graphical considerations determine the solution of the considered problem.

c) Show that the considered problem can be solved by eliminating the variable x_1 and obtaining the optimization problem

$$\begin{cases} \min\limits_{x_2} 4 - (x_2 - 1)^2 + x_2 \\ -1 \leq x_2 \leq 3. \end{cases}$$

d) Solve the optimization problem in part c) and hence obtain a solution for the considered optimization problem.

e) Suppose that one wants to solve the considered optimization problem using recursive quadratic programming methods. Write the quadratic programming problem associated with the considered optimization problem.

Solution 35

a) The level lines and the admissible set are depicted in the figure below. Note that the constraint is always active, and that the gradient of the constraint is never zero, hence all points are regular points.

b) The solution of the problem is obtained when the level set of f is tangent to the admissible set in its lower point. As a result, the optimal point is $(x_1, x_2) = (0, -1)$.

c) We can solve the constraint yielding

$$x_1^2 = 4 - (x_2 - 1)^2.$$

Replacing in f we obtain the function to minimize

$$\tilde{f} = 4 - (x_2 - 1)^2 + x_2.$$

Note that x_2 is not *free*. In fact, from the constraint

$$(x_2 - 1)^2 = 4 - x_1^2 \leq 4$$

we obtain

$$-1 \leq x_2 \leq 3.$$

This shows that eliminating the variable x_1 yields the constrained scalar problem given in part b).

d) Note that a solution to the minimization problem in part c) is obtained at a stationary point of $\tilde{f}$ or at the boundary of the admissible set. The function $\tilde{f}$ has a stationary point (a local maximizer) for $x_2 = 3/2$. Note now that

$$\tilde{f}(-1) = -1, \qquad \tilde{f}(3/2) = 21/4, \qquad \tilde{f}(3) = 3.$$

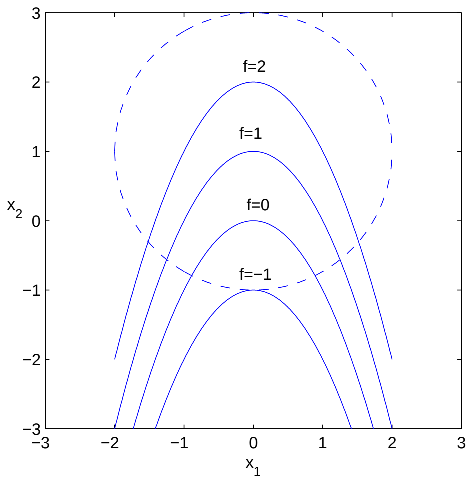
Therefore the function $\tilde{f}$ attains its minimum for $x_2 = -1$. Replacing this into the constraint yields $x_1 = 0$ and this coincides with the optimal solution obtained in part b).

e) Consider the optimization problem $\min\limits_{x} f(x)$ subject to $g(x) = 0$. Using recursive quadratic programming methods to solve this problem one obtains the quadratic programming problem

$$PQ_1^{k+1} \begin{cases} \min\limits_{\delta} f(x_k) + \nabla f(x_k)'\delta + \dfrac{1}{2}\delta' \nabla^2_{xx} L(x_k, \lambda_k)\delta, \\ \dfrac{\partial g(x_k)}{\partial x}\delta = 0, \end{cases}$$

where $L = f + \lambda' g$, $\delta = x - x_k$, and x_k and λ_k are the current estimates of the solution and of the multiplier. For the specific example one has to replace the functions f and g in the above expression yielding

$$PQ_1^{k+1} \begin{cases} \min\limits_{\delta_1, \delta_2} x_{1,k}^2 + x_{2,k} + 2x_{1,k}\delta_1 + \delta_2 + \delta_1^2, \\ 2x_{1,k}\delta_1 + 2(x_{2,k} - 1)\delta_2 = 0. \end{cases}$$



Exercise 36 Consider the optimization problem

$$\begin{cases} \max\limits_{x_1,\ x_2,\ x_3} & x_1^\alpha\ x_2^\alpha\ x_3^\alpha \\ x_1 + x_2 + x_3 - 1 = 0 \\ x_1 \geq 0 \\ x_2 \geq 0 \\ x_3 \geq 0, \end{cases}$$

with $\alpha > 1$.

a) State first order necessary conditions of optimality for this constrained optimization problem.

b) Using the conditions derived in part a), compute candidate optimal solutions. Show that there is only one candidate solution such that $x_1 \neq 0$, $x_2 \neq 0$ and $x_3 \neq 0$.

c) Consider the candidate optimal solution with $x_1 \neq 0$, $x_2 \neq 0$ and $x_3 \neq 0$ determined in part b). Show using second order sufficient conditions that such a candidate optimal point is a local maximizer.

Solution 36

a) Let (note the change of sign in the objective function to transform the problem into a minimization problem)

$$L = -x_1^\alpha x_2^\alpha x_3^\alpha + \lambda(x_1 + x_2 + x_3 - 1) + \mu_1(-x_1) + \mu_2(-x_2) + \mu_3(-x_3).$$

The first order necessary conditions of optimality for the problem are

$$-\alpha x_1^{\alpha-1} x_2^\alpha x_3^\alpha + \lambda - \mu_1 = 0, \qquad -\alpha x_1^\alpha x_2^{\alpha-1} x_3^\alpha + \lambda - \mu_2 = 0, \qquad -\alpha x_1^\alpha x_1^\alpha x_3^{\alpha-1} + \lambda - \mu_3 = 0,$$

$$x_1 + x_2 + x_3 - 1 = 0, \qquad -x_1 \leq 0, \qquad -x_2 \leq 0, \qquad -x_3 \leq 0,$$

$$\mu_1 \geq 0, \qquad \mu_2 \geq 0, \qquad \mu_3 \geq 0, \qquad x_1\mu_1 = 0, \qquad x_2\mu_2 = 0, \qquad x_3\mu_3 = 0.$$

b) Consider the condition $x_1\mu_1 = 0$. This implies $\mu_1 = 0$ or $x_1 = 0$. If $x_1 = 0$ then, since $\alpha > 1$,

$$\lambda = \mu_1 = \mu_2 = \mu_3 = \kappa \geq 0,$$

for some constant κ. If $\kappa > 0$ then $x_2 = 0$ and $x_3 = 0$ which is not feasible. If $\kappa = 0$ then any x_2 and x_3 such that

$$x_2 + x_3 - 1 = 0, \qquad x_2 \geq 0, \qquad x_3 \geq 0,$$

satisfy necessary conditions of optimality. We obtain similar conclusions from the conditions $x_2\mu_2 = 0$ and $x_3\mu_3 = 0$. To obtain other candidate solutions we have to consider the case $x_1 \neq 0$, $x_2 \neq 0$ and $x_3 \neq 0$. In this case, $\mu_1 = \mu_2 = \mu_3 = 0$, and

$$\begin{aligned} \alpha x_1^{\alpha-1}x_2^\alpha x_3^\alpha - \lambda &= 0, \\ \alpha x_1^\alpha x_2^{\alpha-1}x_3^\alpha - \lambda &= 0, \\ \alpha x_1^\alpha x_1^\alpha x_3^{\alpha-1} - \lambda &= 0. \end{aligned}$$

The above equations imply $x_1 = x_2 = x_3$ which, together with the constraint $x_1 + x_2 + x_3 - 1 = 0$, yields the candidate optimal solution $(x_1, x_2, x_3) = (1/3, 1/3, 1/3)$. In summary, all candidate optimal solutions are

$$\begin{aligned} &x_1 = x_2 = x_3 = 1/3 \\ &x_1 = 0,\ x_2 + x_3 = 1,\ x_2 \geq 0,\ x_3 \geq 0, \\ &x_2 = 0,\ x_1 + x_3 = 1,\ x_1 \geq 0,\ x_3 \geq 0, \\ &x_3 = 0,\ x_1 + x_2 = 1,\ x_1 \geq 0,\ x_2 \geq 0. \end{aligned}$$

c) At the point $(x_1, x_2, x_3) = (1/3, 1/3, 1/3)$ the only active constraint is the equality constraint. Hence, the second order sufficient conditions of optimality are $s'\nabla^2 Ls > 0$, for all $s \neq 0$ such that $[1,\ 1,\ 1]s = 0$. Note now that at the considered point

$$\nabla^2 L = -\left(\frac{1}{3}\right)^{3\alpha-2} \left[\begin{array}{ccc} \alpha(\alpha-1) & \alpha^2 & \alpha^2 \\ \alpha^2 & \alpha(\alpha-1) & \alpha^2 \\ \alpha^2 & \alpha^2 & \alpha(\alpha-1) \end{array}\right]$$

and the admissible s can be parameterized as

$$s_a = [\beta, 0, -\beta]' \qquad s_b = [\gamma, -\gamma, 0]'.$$

As a result

$$s_a'\nabla^2 Ls_a = 2\left(\frac{1}{3}\right)^{3\alpha-2}\beta^2\alpha > 0, \qquad s_b'\nabla^2 Ls_b = 2\left(\frac{1}{3}\right)^{3\alpha-2}\gamma^2\alpha > 0,$$

which show that the considered point is a local minimizer (hence a maximizer for the original problem).

Exercise 37 Consider the problem of approximating a matrix $Q \in I\!R^{n\times n}$ with a matrix of the form $A = \rho I$, with I the identity matrix of dimension $n \times n$ and $\rho \geq 0$.
As a measure of the distance between the two matrices we could use either the square of the Frobenius norm or the infinity norm. The Frobenius norm of a matrix $L \in I\!R^{n\times n}$ is defined as

$$\|L\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n L_{ij}^2},$$

where the L_{ij}'s denote the entry of the matrix L. The infinity norm of a matrix $L \in I\!R^{n\times n}$ is defined as

$$\|L\|_\infty = \max_i \sum_{j=1}^n |L_{ij}|.$$

The optimal approximation problem is thus the problem of determining the nonnegative constant ρ which minimizes

$$\|Q - \rho I\|_F^2$$

or

$$\|Q - \rho I\|_\infty.$$

a) Show that the considered optimal approximation problems can be written as constrained minimization problems with one inequality constraint.

b) Consider the Frobenius norm. Solve the problem derived in part a). Show that if $\mathrm{trace}(Q) > 0$ then the optimal ρ is positive, and if $\mathrm{trace}(Q) \leq 0$ then the optimal ρ is zero.

c) Consider the infinity norm and assume that $n = 2$, hence

$$Q = \left[\begin{array}{cc} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{array} \right],$$

that $0 < Q_{11} < Q_{22}$ and that $|Q_{12}| = |Q_{21}|$.

i) Sketch the graph of the function to be minimized.

ii) Argue that the optimal solution $\rho_\star$ is such that

$$0 < Q_{11} < \rho_\star < Q_{22}.$$

iii) Compute the optimal solution $\rho_\star$.

Solution 37

a) The optimal approximation problems can be written as

$$P_f : \left\{ \begin{array}{l} \min\limits_{\rho} \|Q - \rho I\|_F^2, \\ \\ \rho \geq 0, \end{array} \right. \qquad \text{or} \qquad P_\infty : \left\{ \begin{array}{l} \min\limits_{\rho} \|Q - \rho I\|_\infty, \\ \\ \rho \geq 0. \end{array} \right.$$

b) Note that

$$\begin{array}{rcl} \|Q - \rho I\|_F^2 & = & (Q_{11} - \rho)^2 + Q_{12}^2 + \cdots + Q_{1n}^2 + \\ & & Q_{21}^2 + (Q_{22} - \rho)^2 + Q_{23}^2 + \cdots + Q_{2n}^2 + \cdots + \\ & & Q_{n1}^2 + \cdots + Q_{2,n-1}^2 + (Q_{nn} - \rho)^2 \end{array}$$

hence

$$\|Q - \rho I\|_F^2 = n\rho^2 - 2\rho \overbrace{(Q_{11} + Q_{22} + \cdots + Q_{nn})}^{\mathrm{trace}(Q)} + \text{constant terms}.$$

If $\mathrm{trace}(Q) > 0$ the function $\|Q - \rho I\|_F^2$, which is convex, has a global minimizer for $\rho = \dfrac{\mathrm{trace}(Q)}{n}$. If $\mathrm{trace}(Q) \leq 0$ the function $\|Q - \rho I\|_F^2$ is monotonically increasing for $\rho \geq 0$, hence it achieves its minimum, in the set $\rho \geq 0$, for $\rho = 0$.

c) The optimal approximation problem is now

$$\tilde{P}_\infty : \left\{ \begin{array}{l} \min\limits_{\rho} \left(\max \left(|Q_{11} - \rho| + |Q_{12}|, |Q_{21}| + |Q_{22} - \rho| \right) \right), \\ \\ \rho \geq 0. \end{array} \right.$$

A sketch of the function to be minimized is in the figure below. From this it is clear that $0 < Q_{11} < \rho_\star < Q_{22}$. Note that $\rho_\star$ is such that

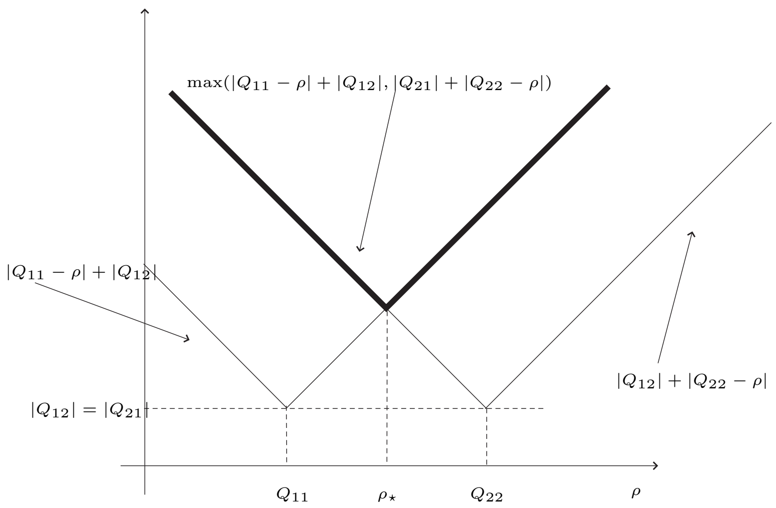
$$|Q_{11} - \rho_\star| + |Q_{12}| = |Q_{21}| + |Q_{22} - \rho_\star|.$$

However, because $0 < Q_{11} < \rho_\star < Q_{22}$ this can be rewritten as

$$\rho_\star - |Q_{11}| + |Q_{12}| = |Q_{21}| + |Q_{22}| - \rho_\star.$$

As a result (recall that $Q_{11} > 0$, $Q_{22} > 0$ and $|Q_{12}| = |Q_{21}|$)

$$\rho_\star = \frac{Q_{11} + Q_{22}}{2}.$$



Exercise 38 Consider the optimization problem

$$\begin{cases} \min\limits_{x_1,x_2} x_1^2 + x_2, \\ x_1^2 + (x_2-1)^2 \geq 1, \\ x_1^2 + (x_2-2)^2 \leq 4. \end{cases}$$

a) Sketch in the (x_1, x_2)-plane the admissible set and show that there is a point which is not a regular point for the constraints.

b) State first order necessary conditions of optimality for such a constrained optimization problem.

c) Find candidate optimal solutions for the considered problem.

d) Prove that the non-regular point for the constraints is the global minimizer for the considered problem.

Solution 38

a) The admissible set is the set outside a circle of radius one and centered at $(0, 1)$ and inside a circle of radius two and centered at $(0, 2)$, which is the shaded region in the figure below. The point $(0, 0)$ is not a regular point for the constraints because at this point both constraints are active and their gradients, namely

$$\begin{bmatrix} 2x_1 \\ 2(x_2-1) \end{bmatrix}, \qquad \begin{bmatrix} 2x_1 \\ 2(x_2-2) \end{bmatrix},$$

evaluated at the point, are linearly dependent.

b) To write necessary conditions of optimality rewrite first the constraints as

$$1 - x_1^2 - (x_2-1)^2 \leq 0 \qquad x_1^2 + (x_2-2)^2 - 4 \leq 0$$

and define the Lagrangian function

$$L(x_1, x_2, \mu_1, \mu_2) = x_1^2 + x_2 + \mu_1(1 - x_1^2 - (x_2-1)^2) + \mu_2(x_1^2 + (x_2-2)^2 - 4).$$

The necessary conditions of optimality are

$$\begin{aligned} \frac{dL}{dx_1} = 2x_1 - 2\mu_1 x_1 + 2\mu_2 x_1 = 0, &\qquad \frac{dL}{dx_2} = 1 - 2\mu_1(x_2-1) + 2\mu_2(x_2-2) = 0, \\ 1 - x_1^2 - (x_2-1)^2 \leq 0, &\qquad x_1^2 + (x_2-2)^2 - 4 \leq 0, \\ \mu_1 \geq 0, &\qquad \mu_2 \geq 0, \\ \mu_1(1 - x_1^2 - (x_2-1)^2) = 0, &\qquad \mu_2(x_1^2 + (x_2-2)^2 - 4) = 0. \end{aligned}$$

c) To find candidate optimal solutions we exploit the complementarity conditions, hence we have four possibilities.

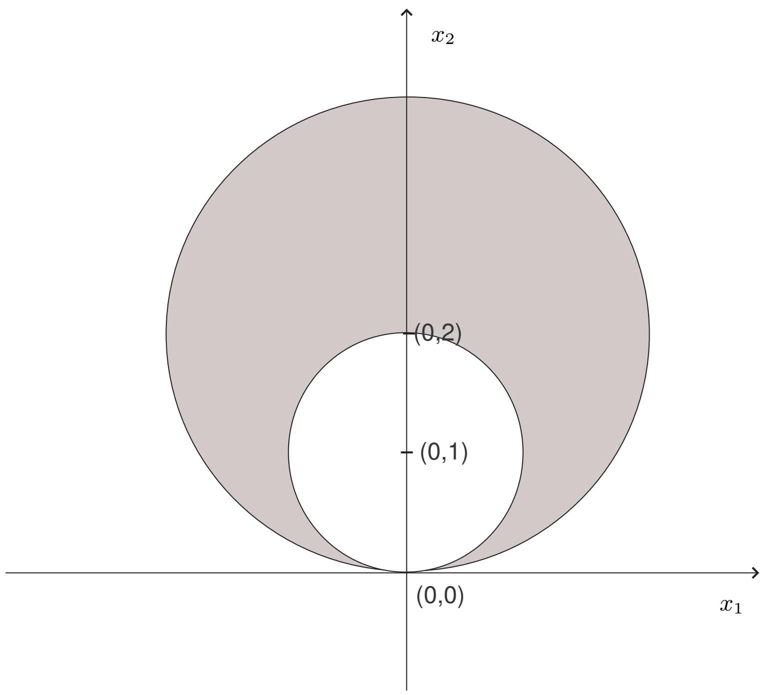
- $\mu_1 = 0$ and $\mu_2 = 0$.
  This selection yields $0 = \frac{dL}{dx_2} = 1$, hence no candidate optimal solution.
- $\mu_1 = 0$ and $x_1^2 + (x_2 - 2)^2 - 4 = 0$.
  This selection yields, from $0 = \frac{dL}{dx_1}$, either $x_1 = 0$ or $\mu_2 = -1$. The first option yields $x_2 = 0$ or $x_2 = 4$, whereas the second option violates the positivity of μ_2. Moreover, the selection $x_1 = 0$ and $x_2 = 4$ yields, from $0 = \frac{dL}{dx_2}$, $\mu_2 < 0$, hence it is not a candidate solution.
- $1 - x_1^2 - (x_2 - 1)^2 = 0$ and $\mu_2 = 0$.
  This selection yields, from $0 = \frac{dL}{dx_1}$, $x_1 = 0$ or $\mu_1 = 1$. The first option yields $x_2 = 0$ or $x_2 = 2$. The second option yields, from $0 = \frac{dL}{dx_2}$, $x_2 = 3/2$, hence, from $1 - x_1^2 - (x_2 - 1)^2 = 0$, $x_1 = \pm\frac{\sqrt{3}}{2}$.
- $1 - x_1^2 - (x_2 - 1)^2 = 0$ and $x_1^2 + (x_2 - 2)^2 - 4 = 0$.
  The only point consistent with these conditions is $(0, 0)$.

In summary the candidate solutions obtained so far are as follows.

- $(0, 0)$.
- $(0, 2)$.
- $\left(\pm\frac{\sqrt{3}}{2}, \frac{3}{2}\right)$.

Hence there are four candidate optimal solutions.

d) The nonregular point $(0, 0)$ is such that $x_1^2 + x_2 = 0$. Note now that the function $x_1^2 + x_2$ is always nonnegative in the admissible set and it is zero, in the admissible set, if and only if $x_1 = x_2 = 0$. Hence the nonregular point is a global minimizer for the considered problem. Note that it is not possible to associate, in a unique way, a pair of optimal multipliers to this optimal point.



Exercise 39 Consider the optimization problem

$$\begin{cases} \min\limits_{x_1,\ x_2} x_1^2 + x_2^2, \\ -x_1 \le 0, \\ x_2 - x_1 - 1 = 0. \end{cases}$$

a) Sketch in the (x_1, x_2)-plane the level surfaces of the function to be minimized and the admissible set. Hence show that all points in the admissible set are regular points for the constraints.

b) Using only graphical considerations determine the solution of the considered problem.

c) This constrained optimization problem can be transformed into an unconstrained optimization problem by defining the so-called mixed penalty-barrier function

$$F_\epsilon(x_1, x_2) = x_1^2 + x_2^2 + \frac{1}{\epsilon}(x_2 - x_1 - 1)^2 + \frac{\epsilon}{x_1},$$

with $\epsilon > 0$ and considering the unconstrained minimization of $F_\epsilon(x_1, x_2)$. Determine the stationary points of $F_\epsilon(x_1, x_2)$.
(Hint: solve $\nabla_{x_2} F_\epsilon(x_1, x_2) = 0$ for x_2, and replace the obtained solution in the equation $\nabla_{x_1} F_\epsilon(x_1, x_2) = 0$. Solve this last equation assuming that $x_1 = \alpha\epsilon^{1/2}$, for some $\alpha > 0$ to be determined, and neglecting all terms ϵ^k, for $k \ge 1/2$.)

d) Show that the stationary point of $F_\epsilon(x_1, x_2)$ computed in part c) tends, as ϵ tends to zero, to the optimal solution determined in part b).

Solution 39

a) The admissible set, and the level surfaces of the function to be minimized are as in the figure below. There are two constraints active at the point $(0, 1)$ and their gradients, at this point, are independent. At any other admissible point there is only one active constraint, the equality constraint, and its gradient is always nonzero (it is a constant vector). Thus all points are regular points for the constraints.

b) The optimal solution is obtained considering the smallest circle centered at the origin intersecting the admissible set. Hence, the optimal solution is the point $(0, 1)$.

c) The stationary points of the mixed penalty-barrier function are the solutions of

$$0 = \nabla F_\epsilon = \begin{bmatrix} 2x_1 - \dfrac{2}{\epsilon}(x_2 - x_1 - 1) - \dfrac{\epsilon}{x_1^2} \\ 2x_2 + \dfrac{2}{\epsilon}(x_2 - x_1 - 1) \end{bmatrix}.$$

Solving the second equation yields

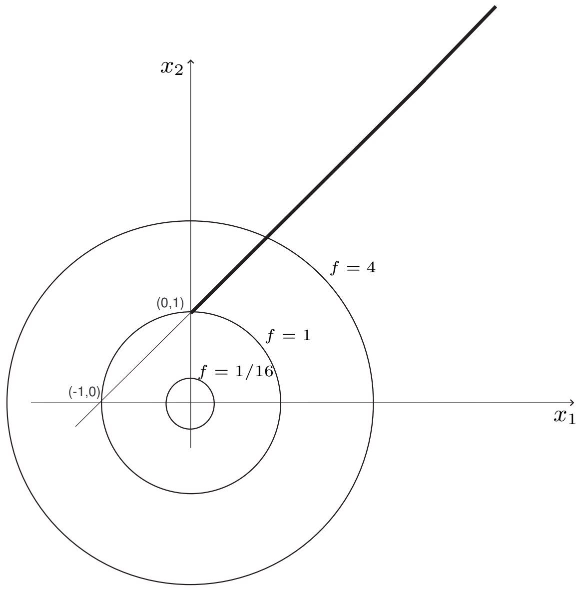
$$x_2 = \frac{x_1 + 1}{\epsilon + 1},$$

and replacing this in the first equation yields

$$0 = \frac{x_1^3(2\epsilon + 4) + 2x_1^2 - \epsilon(1 + \epsilon)}{(\epsilon + 1)x_1^2}.$$

Setting $x_1 = \alpha\sqrt{\epsilon}$ and neglecting all ϵ^k terms, with $k \ge 1/2$, yields $0 = (2\alpha^2 - 1)$, hence (recall that $\alpha > 0$) $x_1 = \sqrt{\epsilon/2}$, and $x_2 = \frac{\sqrt{\epsilon/2}+1}{\epsilon+1}$.

d) As $\epsilon \to 0$, the stationary point of the mixed penalty-barrier function tends to $(0, 1)$, which is the optimal solution of the considered problem.



Exercise 40 Consider the optimization problem

$$\begin{cases} \min\limits_{x_1,\ x_2} x_1 x_2, \\ \dfrac{1}{2}x_1^2 + 2x_2^2 = 1. \end{cases}$$

a) State first order necessary conditions of optimality for such a constrained optimization problem.

b) Using the conditions in part a) determine candidate optimal solutions for the considered problem.

c) Transform the minimization problem into an unconstrained minimization problem using the method of the exact augmented Lagrangian functions and write explicitly the exact augmented Lagrangian function for the considered problem.

d) Show that the candidate optimal solutions determined in part b) are stationary points of the exact augmented Lagrangian function.

e) Find the global minimizer for the considered problem. Is the global minimizer unique?

Solution 40

a) Define the Lagrangian

$$L(x_1, x_2, \lambda) = x_1 x_2 + \lambda(\frac{1}{2}x_1^2 + 2x_2^2 - 1).$$

The first order necessary conditions of optimality are

$$0 = \frac{dL}{dx_1} = x_2 + \lambda x_1, \qquad 0 = \frac{dL}{dx_2} = x_1 + 4\lambda x_2, \qquad \frac{1}{2}x_1^2 + 2x_2^2 - 1 = 0.$$

b) The conditions $\frac{dL}{dx_1} = \frac{dL}{dx_2} = 0$ can be rewritten as

$$\begin{bmatrix} \lambda & 1 \\ 1 & 4\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0.$$

If $4\lambda^2 - 1 \neq 0$ the above equation implies $x_1 = x_2 = 0$, which is not an admissible point. If $4\lambda^2 - 1 = 0$, or $\lambda = \pm\frac{1}{2}$. then $x_2 = \mp\frac{1}{2}x_1$, and replacing in the constrains yields the candidate solutions with the corresponding multipliers, namely

$$(x_1, x_2, \lambda) = \left(1, -\frac{1}{2}, \frac{1}{2}\right), \qquad (x_1, x_2, \lambda) = \left(-1, \frac{1}{2}, \frac{1}{2}\right),$$

$$(x_1, x_2, \lambda) = \left(1, \frac{1}{2}, -\frac{1}{2}\right), \qquad (x_1, x_2, \lambda) = \left(-1, -\frac{1}{2}, -\frac{1}{2}\right).$$

c) The exact augmented Lagrangian function for a constraint optimization problem with equality constraints is

$$S(x, \lambda) = f(x) + \lambda' g(x) + \frac{1}{\epsilon}\|g(x)\|^2 + \eta\|\frac{\partial g(x)}{\partial x}\nabla_x L(x, \lambda)\|^2,$$

with $\epsilon > 0$ and $\eta > 0$. For the considered problem, we have

$$S(x_1, x_2, \lambda) = x_1 x_2 + \lambda(\frac{1}{2}x_1^2 + 2x_2^2 - 1) + \frac{1}{\epsilon}(\frac{1}{2}x_1^2 + 2x_2^2 - 1)^2 + \eta\left(\begin{bmatrix} x_1 & 4x_2 \end{bmatrix}\begin{bmatrix} x_2 + \lambda x_1 \\ x_1 + 4\lambda x_2 \end{bmatrix}\right)^2.$$

d) The stationary points of the function $S(x_1, x_2, \lambda)$ are the solutions of the equations

$$0 = \frac{dS}{dx_1} = x_2 + \lambda x_1 + \frac{2x_1}{\epsilon}(\frac{1}{2}x_1^2 + 2x_2^2 - 1) + 2\eta(5x_1 x_2 + \lambda x_1^2 + 16\lambda x_2^2)(5x_2 + 2\lambda x_1),$$

$$0 = \frac{dS}{dx_2} = x_1 + 4\lambda x_2 + \frac{8x_2}{\epsilon}(\frac{1}{2}x_1^2 + 2x_2^2 - 1) + 2\eta(5x_1 x_2 + \lambda x_1^2 + 16\lambda x_2^2)(5x_1 + 32\lambda x_2),$$

$$0 = \frac{dS}{d\lambda} = \frac{1}{2}x_1^2 + 2x_2^2 - 1 + 2\eta(5x_1 x_2 + \lambda x_1^2 + 16\lambda x_2^2)(x_1^2 + 16x_2^2).$$

Replacing the candidate points obtained in part b) shows that indeed they are stationary points for the augmented Lagrangian function. (Note that this is true for any ϵ and η.)

e) To find the global minimum we evaluate the function to be minimized at the candidate optimal solutions:

$$(x_1 x_2)_{x_1=1, x_2=-1/2} = -\frac{1}{2}, \qquad (x_1 x_2)_{x_1=-1, x_2=1/2} = -\frac{1}{2},$$

$$(x_1 x_2)_{x_1=1, x_2=1/2} = \frac{1}{2}, \qquad (x_1 x_2)_{x_1=-1, x_2=-1/2} = \frac{1}{2}.$$

Hence, the points $(1, -1/2)$ and $(-1, 1/2)$ are both global minimizers. (Note that the points $(1, 1/2)$ and $(-1, -1/2)$ are both global maximizers.)

Exercise 41 Consider the optimization problems

$$P_{min} \left\{ \begin{array}{l} \min\limits_{x_1,\ x_2} |x_1| + |x_2|, \\ x_1^2 + x_2^2 = 1, \end{array} \right.$$

and

$$P_{max} \left\{ \begin{array}{l} \max\limits_{x_1,\ x_2} |x_1| + |x_2|, \\ x_1^2 + x_2^2 = 1. \end{array} \right.$$

a) Sketch in the (x_1, x_2)-plane the admissible set and the level lines of the function $|x_1| + |x_2|$.

b) Using only graphical considerations determine the solutions of the considered problems.

c) State first order necessary conditions of optimality for these constrained optimization problems. Show that the optimal solutions determined in part b) satisfy the necessary conditions of optimality. (Hint: use the fact that $\text{sign}(0) = 0$.)

d) Write a penalty function F_ϵ for problem P_{max}. Show that, for $\epsilon > 0$ and sufficiently small, the stationary points of F_ϵ approach the optimal solutions determined in part b). (Do not compute explicitly the stationary points of F_ϵ.)
(Hint: for ϵ sufficiently small, the stationary points of F_ϵ are such that $x_1 \neq 0$ and $x_2 \neq 0$.)

Solution 41

a) The admissible set is the circle of radius one and with center at $(0, 0)$. The level lines of the function $|x_1| + |x_2|$ are squares with their vertices on the x_1- and x_2- axes, as indicated in the figure below.

b) The solution to problem P_{min} is obtained considering the smallest square level set intersecting the admissible set. Hence there are four optimal solutions, namely the points $(0, \pm 1)$ and $(\pm 1, 0)$.
The solution to problem P_{max} is obtained considering the largest square level set intersecting the admissible set. Hence there are four optimal solutions, namely the points $\left(\pm\frac{\sqrt{2}}{2}, \pm\frac{\sqrt{2}}{2}\right)$.

c) Define the Lagrangian

$$L(x_1, x_2, \lambda) = \pm(|x_1| + |x_2|) + \lambda(x_1^2 + x_2^2 - 1),$$

where the $+$ sign has to be used for P_{min} and the $-$ sign has to be used for P_{max}. The first order necessary conditions of optimality are

$$0 = \frac{dL}{dx_1} = \text{sign}(x_1) + 2\lambda x_1, \qquad 0 = \frac{dL}{dx_2} = \text{sign}(x_2) + 2\lambda x_2, \qquad x_1^2 + x_2^2 - 1 = 0,$$

and a direct substitution shows that the solutions determined in part b) satisfy the necessary conditions of optimality.

d) A penalty function for problem P_{max} is

$$F_\epsilon(x_1, x_2) = -(|x_1| + |x_2|) + \frac{1}{\epsilon}(x_1^2 + x_2^2 - 1)^2.$$

The stationary points of F_ϵ are the solutions of the equations

$$0 = -\text{sign}(x_1) + \frac{4}{\epsilon}x_1(x_1^2 + x_2^2 - 1), \qquad 0 = -\text{sign}(x_2) + \frac{4}{\epsilon}x_2(x_1^2 + x_2^2 - 1).$$

If we assume that the stationary points of F_ϵ, for ϵ sufficiently small, are away from $x_1 = 0$ and from $x_2 = 0$, then the stationary points are such that

$$\frac{\text{sign}(x_1)}{x_1} = \frac{\text{sign}(x_2)}{x_2},$$

which implies $x_2 = \pm x_1$. Replacing this in the first of the equations above yields

$$0 = -\text{sign}(x_1) + \frac{4}{\epsilon}x_1(2x_1^2 - 1),$$

or equivalently

$$\frac{\epsilon}{4}\text{sign}(x_1) = x_1(2x_1^2 - 1).$$

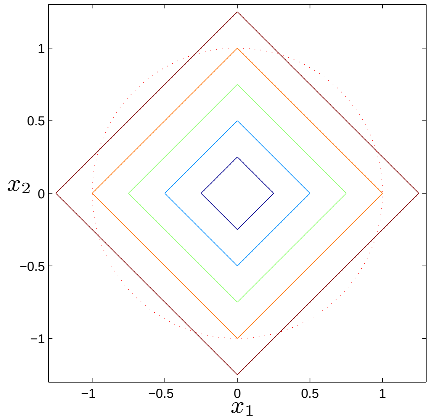
For ϵ sufficiently small the solutions of this equation are of the form

$$x_1 = \pm\frac{\sqrt{2}}{2} + o(\epsilon).$$

As a result, the stationary points of F_ϵ are of the form

$$\left(\pm\left(\frac{\sqrt{2}}{2} + o(\epsilon)\right), \pm\left(\frac{\sqrt{2}}{2} + o(\epsilon)\right)\right),$$

i.e. they are close to the optimal solutions of the problem P_{max} for ϵ sufficiently small.



Exercise 42 Consider the optimization problem

$$\begin{cases} \min\limits_{x_1,\ x_2} x_1^3 - x_1^2 x_2 + 2x_2^2, \\ x_1 \geq 0, \\ x_2 \geq 0. \end{cases}$$

a) State first order necessary conditions of optimality for this constrained optimization problem.

b) Using the conditions derived in part a) compute candidate optimal solutions. Show that there is one candidate solution on the boundary of the admissible set and one in the interior of the admissible set.

c) Using second order sufficient conditions of optimality show that the candidate solution inside the admissible set is not a local minimizer.

d) Show that the candidate optimal solution on the boundary of the admissible set is a local minimizer. (Hint: show that the function to be minimized is zero at the candidate optimal solution, and it is strictly positive in all admissible points in a neighborhood of the candidate optimal solution).

e) Show that the function to be minimized is not bounded from below in the admissible set. Hence, argue that the problem does not have a global solution.
(Hint: consider the function to be minimized along the line $x_2 = 2x_1$, and study its behaviour for $x_1 > 0$ and large.)

Solution 42

a) Define the Lagrangian

$$L(x_1, x_2, \rho_1, \rho_2) = x_1^3 - x_1^2 x_2 + 2x_2^2 + \rho_1(-x_1) + \rho_2(-x_2).$$

The first order necessary conditions of optimality are

$$0 = \frac{dL}{dx_1} = 3x_1^2 - 2x_1 x_2 - \rho_1, \qquad 0 = \frac{dL}{dx_2} = -x_1^2 + 4x_2 - \rho_2,$$

$$-x_1 \leq 0, \qquad -x_2 \leq 0, \qquad \rho_1 > 0, \qquad \rho_2 > 0,$$

$$-x_1\rho_1 = 0, \qquad -x_2\rho_2 = 0.$$

b) Using the complementarity conditions, *i.e.* the last two conditions, we have four possibilities.

- $\rho_1 = 0$ and $\rho_2 = 0$. This yields the candidate optimal solutions $(x_1, x_2) = (0, 0)$ and $(x_1, x_2) = (6, 9)$.

- $\rho_1 = 0$ and $x_2 = 0$. This yields the candidate optimal solution $(x_1, x_2) = (0, 0)$.
- $x_1 = 0$ and $\rho_2 = 0$. This yields the candidate optimal solution $(x_1, x_2) = (0, 0)$.
- $x_1 = 0$ and $x_2 = 0$.

In summary there are two candidate optimal solutions: the point $(0, 0)$, on the boundary of the admissible set, and the point $(6, 9)$ in the interior of the admissible set.

c) The second order sufficient condition of optimality for the candidate point in the interior of the admissible set is $\nabla^2 L(6, 9) > 0$. Note that

$$\nabla^2 L(6,9) = 2 \begin{bmatrix} 9 & -6 \\ -6 & 2 \end{bmatrix},$$

and that $\det \nabla^2 L(6, 9) < 0$, which implies that $\nabla^2 L(6, 9)$ is not positive definite. Hence the candidate optimal point in the interior of the admissible set is not a local minimizer (tt is a saddle point).

d) To show that the point $(0, 0)$ is a local minimizer note that the function f to be minimized is such that $f(0, 0) = 0$, $f(x_1, 0) > 0$ for $x_1 > 0$, and $f(0, x_2) > 0$ for $x_2 > 0$. Consider now straight lines described by $x_2 = \alpha x_1$, with $\alpha > 0$. Then

$$f(x_1, \alpha x_1) = \alpha^2 \left(\frac{1-\alpha}{\alpha^2} x_1^3 + 2x_1^2 \right),$$

which is positive for all $\alpha > 0$ and all $x_1 > 0$ and sufficiently small. Since the function f is zero at the candidate optimal solution $(0, 0)$ and strictly positive in all admissible point in a neighborhood of this point, then the point is a local minimizer.

e) The function f along the line $x_2 = 2x_1$ is given by $f(x_1, 2x_1) = -x_1^3 + 4x_1^2$ and this function is not bounded from below, *i.e.* $\lim_{x_1 \to \infty} f(x_1, 2x_1) = -\infty$. This implies that the considered optimization problem does not have a global solution.

Exercise 43 Consider the optimization problem

$$\begin{cases} \max_{x_1,\, x_2,\, x_3} (x_1 x_2 + x_2 x_3 + x_1 x_3), \\ x_1 + x_2 + x_3 = 3. \end{cases}$$

a) State first order necessary conditions of optimality for this constrained optimization problem and show that there exists only one candidate optimal solution.

b) Using second order sufficient conditions of optimality show that the candidate solution is a local maximizer.

c) Consider the use of an exact penalty function for the solution of the problem.

   i) Write an exact penalty function G for the problem.

   ii) Show that the function is well-defined for every (x_1, x_2, x_3).

   iii) Show that the exact penalty function has only one stationary point and this coincides with the optimal solution of the problem determined in part b).

Solution 43

a) Define the Lagrangian (note the change in sign due to the transformation of the maximization problem into a minimization problem)

$$L(x_1, x_2, x_3, \lambda) = -(x_1 x_2 + x_2 x_3 + x_1 x_3) + \lambda(x_1 + x_2 + x_3 - 3).$$

The first order necessary conditions of optimality are

$$0 = \frac{dL}{dx_1} = -x_2 - x_3 + \lambda, \qquad 0 = \frac{dL}{dx_2} = -x_1 - x_3 + \lambda,$$

$$0 = \frac{dL}{dx_3} = -x_2 - x_1 + \lambda, \qquad 0 = x_1 + x_2 + x_3 - 3.$$

This is system a linear equations with the unique solution $(x_1, x_2, x_3, \lambda) = (1, 1, 1, 2)$. Hence the problem has only one candidate optimal solution.

b) Note that

$$\nabla^2 L = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$$

and

$$\frac{\partial g}{\partial x} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

The candidate optimal solution is a minimizer if $s'\nabla^2 L s > 0$ for all $s \neq 0$ such that $s' \frac{\partial g}{\partial x} = 0$. The set of such s's can be described by linear combinations of the vectors

$$s_1' = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix} \qquad s_2' = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}.$$

Note that

$$[s_1, s_2]'\nabla^2 L[s_1, s_2] = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} > 0,$$

hence the candidate optimal solution is a local minimizer.

c) The exact penalty function for a constraint optimization problem with equality constraints is

$$G(x) = f(x) - g'(x)\left(\frac{\partial g}{\partial x}\frac{\partial g'}{\partial x}\right)^{-1}\frac{\partial g}{\partial x}\nabla f + \frac{1}{\epsilon}\|g(x)\|^2,$$

with $\epsilon > 0$.

i) For the considered problem we have

$$G(x_1, x_2, x_3) = -(x_1x_2+x_2x_3+x_1x_3)+\frac{2}{3}(x_1+x_2+x_3-3)(x_1+x_2+x3)+\frac{1}{\epsilon}(x_1+x_2+x_3-3)^2.$$

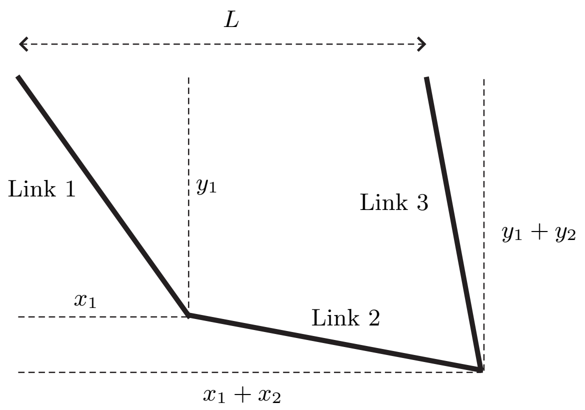
ii) The function is well-defined for all (x_1, x_2, x_3) since $\dfrac{\partial g}{\partial x}\dfrac{\partial g'}{\partial x}$ is a full rank matrix (it is a nonzero constant).

iii) The stationary points of the function $G(x_1, x_2, x_3)$ are the solutions of the equations

$$0 = \nabla G = \begin{bmatrix} \frac{1}{3}(4x_1 + x_2 + x_3) - 2 + \frac{2}{\epsilon}(x_1 + x_2 + x_3 - 3) \\ \frac{1}{3}(x_1 + 4x_2 + x_3) - 2 + \frac{2}{\epsilon}(x_1 + x_2 + x_3 - 3) \\ \frac{1}{3}(x_1 + x_2 + 4x_3) - 2 + \frac{2}{\epsilon}(x_1 + x_2 + x_3 - 3) \end{bmatrix}.$$

These equations have a unique solution $(x_1, x_2, x_3) = (1, 1, 1)$ which does not depend upon ϵ and coincides with the optimal solution determined in part b).

Exercise 44 A chain with three links, each of length one, hangs between two points at the same height, a distance $L > 1$ apart (see the figure below). To find the form in which the chain hangs we minimize the potential energy.
Let (x_i, y_i) be the displacement of the right end of the ith link , from the right end of the $(i-1)$th link.



The potential energy is therefore

$$V(y_1, y_2, y_3) = \frac{1}{2}y_1 + (y_1 + \frac{1}{2}y_2) + (y_1 + y_2 + \frac{1}{2}y_3).$$

a) The condition that the hanging points are a distance L apart can be translated in the constraint

$$x_1 + x_2 + x_3 = L.$$

Express this constraint in terms of the variables y_i.
(Hint: use Pythagoras' Theorem!)

b) Show that the condition that the height of the hanging points is the same can be expressed with the constraint

$$y_1 + y_2 + y_3 = 0.$$

c) Consider the problem of minimizing the potential energy $V(y_1, y_2, y_3)$ subject to the constraints determined in parts b) and c).

i) Write necessary conditions of optimality for the considered optimization problem.

ii) Using physical considerations it may be noted that candidate optimal solutions should be such that the chain has a _/ shape or a /‾\\ shape. Show that these two shapes yield values for y_1, y_2, y_3 and for the Lagrangian multipliers such that the necessary conditions of optimality are met.
(Hint: note that for both shapes $y_2 = 0$.)

iii) By evaluating the potential energy at the candidate optimal solutions determined in part c.ii) determine the shape that minimizes the potential energy.

Solution 44

a) From the figure above we obtain, for $i = 1, 2, 3$, $x_i^2 + y_i^2 = 1$, hence $x_i = \sqrt{1 - y_i^2}$, yielding the constrain

$$\sqrt{1 - y_1^2} + \sqrt{1 - y_2^2} + \sqrt{1 - y_3^2} = L.$$

b) Since the height of the left hanging point is at zero, and the y-coordinate of the last link is $y_1+y_2+y_3$ then the condition that both hanging points are at the same height is given by $y_1 + y_2 + y_3 = 0$.

c) The optimization problem to solve is thus

$$\begin{cases} \min\limits_{y_1, y_1, y_3} \frac{5}{2}y_1 + \frac{3}{2}y_2 + \frac{1}{2}y_3, \\ \\ \sqrt{1 - y_1^2} + \sqrt{1 - y_2^2} + \sqrt{1 - y_3^2} - L = 0, \\ \\ y_1 + y_2 + y_3 = 0. \end{cases}$$

i) Define the Lagrangian

$$L(y_1, y_2, y_3, \lambda_1, \lambda_2) = \frac{5}{2}y_1 + \frac{3}{2}y_2 + \frac{1}{2}y_3 + \lambda_1(\sqrt{1-y_1^2} + \sqrt{1-y_2^2} + \sqrt{1-y_3^2} - L) + \lambda_2(y_1 + y_2 + y_3).$$

The first order necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial y_1} = \frac{5}{2} - \lambda_1 \frac{y_1}{\sqrt{1-y_1^2}} + \lambda_2, \qquad 0 = \frac{\partial L}{\partial y_2} = \frac{3}{2} - \lambda_1 \frac{y_2}{\sqrt{1-y_2^2}} + \lambda_2,$$

$$0 = \frac{\partial L}{\partial y_3} = \frac{1}{2} - \lambda_1 \frac{y_3}{\sqrt{1-y_3^2}} + \lambda_2,$$

$$\sqrt{1-y_1^2} + \sqrt{1-y_2^2} + \sqrt{1-y_3^2} - L = 0, \qquad y_1 + y_2 + y_3 = 0.$$

ii) The indicated shapes are such that $y_2 = 0$ and$y_1 = -y_3$. Replacing these conditions in the necessary conditions of optimality yields

$$0 = \frac{\partial L}{\partial y_1} = \frac{5}{2} - \lambda_1 \frac{y_1}{\sqrt{1-y_1^2}} + \lambda_2, \qquad 0 = \frac{\partial L}{\partial y_2} = \frac{3}{2} - \lambda_2,$$

$$0 = \frac{\partial L}{\partial y_3} = \frac{1}{2} - \lambda_1 \frac{y_1}{\sqrt{1-y_1^2}} + \lambda_2,$$

$$2\sqrt{1-y_1^2} + 1 - L = 0, \qquad 0 = 0.$$

These equations have the two solutions

$$y_1 = \pm\frac{1}{2}\sqrt{3 + 2L - L^2}, \qquad \lambda_1 = \pm\frac{L-1}{\sqrt{3+2L-L^2}}, \qquad \lambda_2 = \frac{3}{2}.$$

The one with positive y_1 corresponds to the /‾\ shape, the one with negative y_1 corresponds to the _/ shape. Note that all square roots are well-defined since $L > 1$.

iii) The potential energy for the above shapes is $V(y_1, y_2, y_3) = 2y_1$. Hence the candidate optimal solution with negative y_1 yields a local minimizer.

Exercise 45 The economy class luggage policy of an airline on a transatlantic flight reads:

Each passenger is allowed one piece of luggage. The three linear dimensions, when added together, must not exceed 150 *cm.*

The problem of maximizing the volume of the luggage can be posed and solved with the following steps.

a) Let $x_1 > 0$, $x_2 > 0$ and $x_3 > 0$ be the three linear dimensions (in cm) of a piece of luggage. Write the considered optimization problem as a minimization problem subject to one inequality constraint. (Do not include the constraints $x_1 > 0$, $x_2 > 0$ and $x_3 > 0$ in the formulation of the problem.)

b) State first order necessary conditions of optimality for this constrained optimization problem.

c) Using the conditions derived in part b) compute candidate optimal solutions.

d) Using second order sufficient conditions of optimality determine which of the candidate optimal solutions determined in part c) is a local maximizer.

e) Which is the geometric shape of the 'optimal luggage'?

Solution 45

a) The considered optimization problem can be written as

$$\begin{cases} \max\limits_{x_1,x_2,x_3} \; x_1x_2x_3, \\ x_1 + x_2 + x_3 \le 150. \end{cases}$$

b) Define the Lagrangian (note the change in sign of the objective function)

$$L(x_1, x_2, x_3, \rho) = -x_1 x_2 x_3 + \rho(x_1 + x_2 + x_3 - 150).$$

The first order necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x_1} = -x_2 x_3 + \rho, \qquad 0 = \frac{\partial L}{\partial x_2} = -x_1 x_3 + \rho, \qquad 0 = \frac{\partial L}{\partial x_3} = -x_2 x_1 + \rho,$$

$$\rho \geq 0, \qquad x_1 + x_2 + x_3 - 150 \leq 0, \qquad \rho(x_1 + x_2 + x_3 - 150) = 0.$$

c) Using the complementarity condition, *i.e.* the last condition, we have two cases.
Case 1: $\rho = 0$. This implies $x_1 x_2 = x_2 x_3 = x_1 x_3 = 0$, yielding the sets of candidate solutions

$$x_1 = x_2 = \rho = 0, x_3 \leq 150, \qquad x_1 = x_3 = \rho = 0, x_2 \leq 150, \qquad x_2 = x_3 = \rho = 0, x_1 \leq 150.$$

Case 2: $x_1 + x_2 + x_3 = 150$. This yields the candidate solutions

$$x_1 = x_2 = \rho = 0, x_3 = 150, \qquad x_1 = x_3 = \rho = 0, x_2 = 150, \qquad x_2 = x_3 = \rho = 0, x_1 = 150,$$

and

$$x_1 = x_2 = x_3 = 50, \rho = 50^2.$$

d) Note that

$$\nabla^2 L(x_1, x_2, x_3) = -\begin{bmatrix} 0 & x_3 & x_2 \\ x_3 & 0 & x_1 \\ x_2 & x_1 & 0 \end{bmatrix}.$$

All candidate solutions obtained in Case 1, for which no constrain is active, are such that $\nabla^2 L$ has a positive, a negative and a zero eigenvalue. As a result, all solutions obtained in Case 1, are saddle points. Consider now the candidate solutions obtained in Case 2, and such that $\rho = 0$. For such solutions the condition of strict complementarity does not hold, hence it is not possible to use second order sufficient conditions to classify these points. Finally, consider the candidate optimal solution

$$x_1 = x_2 = x_3 = 50, \qquad \rho = 50^2.$$

The second order sufficient condition require $s'\nabla^2 L(50, 50, 50)s > 0$ for all non-zero s such that

$$\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} s = 0.$$

Such s can be parameterized as

$$s = \begin{bmatrix} s_1 & s_2 & -s_1 - s_2 \end{bmatrix},$$

yielding

$$s'\nabla^2 L(50, 50, 50)s = 100(s_1^2 + s_2^2 + s_1 s_2),$$

which is positive for all non-zero s_1 and s_2. As a result, this candidate optimal solution is a local minimizer. (It is a local maximizer for the original problem).

e) The optimal luggage is a cube!

Exercise 46 Consider the optimization problem

$$\begin{cases} \min\limits_{x_1, x_2} \ x_1 x_2^2, \\ x_1^2 + x_2^2 \leq 2. \end{cases}$$

a) State first order necessary conditions of optimality for this constrained optimization problem.

b) Using the conditions derived in part a) compute candidate optimal solutions.

c) Evaluating the objective function at the candidate optimal solutions determined in part b) derive the solution of the considered optimization problem.

d) The considered constrained optimization problem can be solved minimizing the so-called logarithmic penalty function given by

$$P_l(x_1, x_2) = x_1 x_2^2 - \epsilon \log(2 - x_1^2 - x_2^2),$$

with $\epsilon > 0$.

  i) State first order necessary condition of optimality for P_l.

  ii) Show that the stationary points of P_l are such that

$$x_2^2 = 2x_1^2.$$

  iii) Using the results in part d.ii) show that the stationary points of P_l are such that

$$x_1(3x_1^3 - 2x_1 - \epsilon) = 0.$$

  Hence argue that, as ϵ approaches zero the stationary points of P_l approach candidate optimal solutions for the considered problem.

Solution 46

a) Define the Lagrangian

$$L(x_1, x_2, \rho) = x_1 x_2^2 + \rho(x_1^2 + x_2^2 - 2).$$

The first order necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x_1} = x_2^2 + 2\rho x_1, \qquad 0 = \frac{\partial L}{\partial x_2} = 2x_1 x_2 + 2\rho x_2,$$

$$x_1^2 + x_2^2 - 2 \le 0, \qquad \rho > 0, \qquad \rho(x_1^2 + x_2^2 - 2) = 0.$$

b) Using the complementarity conditions, *i.e.* the last condition, we have two possibilities.

- $\rho = 0$. This yields the candidate optimal solutions

$$P_1 : \ (x_1, x_2) = (\alpha, 0)$$

  with $|\alpha| \le \sqrt{2}$. Note that at $(x_1, x_2) = (\pm\sqrt{2}, 0)$ the condition of strict complementarity does not hold.

- $x_1^2 + x_2^2 - 2 = 0$. This yields the candidate optimal solutions

$$P_2 : \ (x_1, x_2) = (\pm\sqrt{2}, 0),$$

  with $\rho \ge 0$, and

$$P_3 : \ (x_1, x_2) = (-\frac{\sqrt{6}}{3}, \pm\frac{\sqrt{12}}{3}),$$

  with $\rho = \frac{\sqrt{6}}{3}$.

In summary there are infinitely many candidate optimal solutions, some of which such that second order sufficient conditions cannot be used.

c) The values of the objective function at candidate optimal points are

$$f(P_1) = 0, \qquad f(P_2) = 0, \qquad f(P_3) = -\frac{4}{9}\sqrt{6}.$$

Hence P_3 is the solution of the considered problem.

d) i) The first order necessary condition of optimality for P_l are

$$0 = \frac{\partial P_l}{\partial x_1} = x_2^2 + 2\epsilon \frac{x_1}{2 - x_1^2 - x_2^2}, \qquad 0 = \frac{\partial P_l}{\partial x_2} = 2x_1 x_2 + 2\epsilon \frac{x_2}{2 - x_1^2 - x_2^2}.$$

ii) The equations defining the stationary points of P_l yield, for nonzero x_1 and x_2,

$$-2\epsilon\frac{1}{2-x_1^2-x_2^2} = \frac{x_2^2}{x_1} = 2x_1,$$

hence stationary points are such that

$$x_2^2 = 2x_1^2.$$

If $x_1 = 0$ then the necessary conditions yield $x_2 = 0$, and similarly for $x_2 = 0$. Hence, the above relation holds for any x_1 and x_2.

iii) Replacing the above relation in the equation

$$0 = \frac{\partial P_l}{\partial x_1}$$

yields

$$2x_1\frac{3x_1^3 - 2x_1 - \epsilon}{3x_1^2 - 2} = 0$$

from which we infer that, as $\epsilon \to 0$, $x_1 \to 0$ or $x_1 \to \pm\frac{\sqrt{6}}{3}$. As a result, as ϵ goes to zero the stationary points of P_l *approach* the candidate optimal solutions of the problem.

Exercise 47 Consider the optimization problem

$$\begin{cases} \max\limits_{x_1,x_2,x_3} \; x_1 + 2x_2 + x_3, \\ x_1^2 + x_2^2 + x_3^2 \leq 1. \end{cases}$$

a) State first order necessary condition of optimality for this constrained optimization problem.

b) Using the conditions derived in part a) compute candidate optimal solutions.

c) Using second order sufficient conditions of optimality determine the solution of the optimization problem.

d) Consider the change of variables

$$x_1 = r\cos\theta\sin\phi, \qquad x_2 = r\sin\theta\sin\phi, \qquad x_3 = r\cos\phi,$$

with $r \geq 0$, $\theta \in [0, 2\pi)$, and $\phi = [0, 2\pi)$.

i) Rewrite the considered optimization problem in the new variables and show that the resulting problem can be written in the form

$$\begin{cases} \max\limits_{r,\theta,\phi} \; r\Psi(\theta,\phi), \\ r \leq 1, \\ \theta \in [0, 2\pi), \\ \phi \in [0, 2\pi), \end{cases}$$

Determine the function $\Psi(\theta, \phi)$.

ii) Argue that the problem is equivalent to the unconstrained optimization problem

$$\max_{\theta,\phi} \Psi(\theta, \phi).$$

iii) Find candidate solutions of the unconstrained optimization problem in part d.ii), and show that one of the candidate solutions coincides with the optimal solution determined in part c).

Solution 47

a) Define the Lagrangian (note the sign change due to the transformation of the maximization problem into a minimization problem)

$$L(x_1, x_2, x_3, \rho) = -x_1 - 2x_2 - x_3 + \rho(x_1^2 + x_2^2 + x_3^2 - 1).$$

The first order necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x_1} = -1 + 2\rho x_1, \qquad 0 = \frac{\partial L}{\partial x_2} = -1 + 2\rho x_2, \qquad 0 = \frac{\partial L}{\partial x_2} = -1 + 2\rho x_3,$$

$$\rho \geq 0, \qquad x_1^2 + x_2^2 + x_3^2 - 1 \leq 0, \qquad \rho(x_1^2 + x_2^2 + x_3^2 - 1) = 0.$$

b) Using the complementarity conditions, *i.e.* the last condition, we have two possibilities.

- $\rho = 0$. This does not yield any candidate optimal solution.
- $x_1^2 + x_2^2 + x_3^2 - 1 = 0$, $\rho > 0$. This yields the candidate optimal solution

$$P: \ (x_1, x_2, x_3) = \frac{1}{2\rho}(1, 2, 1)$$

with $\rho \geq 0$ such that $\frac{3}{2\rho^2} = 1$.

In summary there is only one candidate optimal solution given by

$$P: \ (x_1, x_2, x3) = \frac{\sqrt{6}}{6}(1, 2, 1).$$

c) The Hessian of the Lagrangian is $\nabla^2 L = 2\rho I$, with I the identity matrix. Hence, the Hessian is positive definite at the candidate optimal solution which is therefore a (local) minimizer for the problem (note that we have changed the sign of the objective function to transform the maximization problem into a minimization one).

d) i) Applying the change of variable to the objective function yields the transformed objective function

$$r(\cos\theta \sin\phi + 2\sin\theta \sin\phi + \cos\phi),$$

whereas the constraint is transformed into $r^2 \leq 1$, which is equivalent to $r \leq 1$ since $r \geq 0$. As a result, the function Ψ is given by

$$\Psi(\theta, \phi) = (\cos\theta \sin\phi + 2\sin\theta \sin\phi + \cos\phi).$$

ii) The objective function in the transformed variables is separable, *i.e.* it is the product of two functions of different variables, namely r and Ψ. As a result, the maximization is achieved maximizing Ψ and r. The latter is maximized for $r = 1$. The former has to be maximized for $\theta \in [0, 2\pi)$ and $\phi \in [0, 2\pi)$. However, since Ψ is periodic in θ and ϕ it can be maximized disregarding the constraints.

iii) The stationary points of Ψ are the solutions of

$$\sin\phi(2\cos\theta - \sin\theta) = 0 \qquad \cos\theta\cos\phi + 2\sin\theta\cos\phi + \sin\phi = 0.$$

The first equation yields

- $\phi = 0$ or $\phi = \pi$, which replaced in the second equation yield $\theta = -\arctan 2$;
- $\theta = \arctan 2$, yielding $\phi = \arctan\sqrt{5}$.

In the original coordinates the first candidate solutions yield $(x_1, x_2, x_3) = (0, 0, \pm 1)$, whereas the second candidate solution give the optimal solution determined in part c).

Exercise 48 The methods of optimization can be used to solve simple geometric problems. Consider the following list of problems. For each of them, formulate the problem as an optimisation problem, defining the decision variables, the cost to be optimised, and the admissible set, and provide explicit solutions.

a) Show that of all rectangles with a fixed positive area the one with the smallest perimeter is a square.

b) Show that of all rectangles with a fixed positive perimeter the one with the largest area is a square.

c) Find the rectangle of largest area that has its base on the x-axis and its other two vertices above the x-axis and on the parabola $y = 8 - x^2$.

d) A piece of wire 10 meters long is cut into two pieces. One piece is bent to form a square, the other piece is bent to form an equilateral triangle. How should the wire be cut so that the total area of the square and of the triangle is a maximum or a minimum?

Solution 48

a) Let x and y be the height and length of the rectangle, then we want to minimize $P = 2x + 2y$ subject to $xy = A$, $x > 0$ and $y > 0$. Using the constraint on the area we have

$$y = \frac{A}{x},$$

hence we need to minimize

$$P = 2x + 2\frac{A}{x}$$

subject to $x > 0$. Ignoring the positivity constraint on x, the stationary points of P are the solutions of

$$0 = \frac{dP}{dx} = 2 - 2\frac{A}{x^2},$$

i.e. $x = \pm\sqrt{A}$. The only feasible solution is $x = \sqrt{A}$, which is a global minimizer, since P is convex for $x > 0$, yielding $y = \sqrt{A}$, hence the rectangle with minimum perimeter is a square with perimeter $P = 4\sqrt{A}$.

b) Let x and y be the height and length of the rectangle, then we want to maximize $A = xy$ subject to $P = 2x + 2y$, $x > 0$ and $y > 0$. Using the constraint on the perimeter we have

$$A = x\frac{P - 2x}{2} = \frac{1}{2}Px - x^2,$$

subject to $x > 0$. Ignoring the positivity constraint on x, the stationary points of A are the solutions of $0 = \frac{dA}{dx} = \frac{1}{2}P - 2x$, *i.e.* $x = \frac{1}{4}P$. This solution is positive, hence feasible, and it is a global maximizer since the function is concave, hence the rectangle with maximum area is a square with area $A = \frac{P^2}{16}$.

c) The area of the rectangle is (note that to have two vertices on the given parabola, the two vertices on the x-axis should be symmetric with respect to the y-axis)

$$A = 2xy = 2x(8 - x^2) = 16x - 2x^3,$$

with $0 \le x \le \sqrt{8}$. The function A is continuous in the interval $[0, \sqrt{8}]$, hence its global maximum is either a stationary point or an extreme of the interval. Stationary points are the solutions of $0 = \frac{dA}{dx} = 16 - 6x^2$, *i.e.* $x = \pm\sqrt{8/3}$. Note now that

$$x = 0 \Rightarrow A = 0, \qquad x = \sqrt{8/3} \Rightarrow A = \frac{64}{3}\frac{\sqrt{2}}{\sqrt{3}}, \qquad x = \sqrt{8} \Rightarrow A = 0,$$

hence the optimal solution is $x = \sqrt{8/3}$.

d) Let x be the length of the wire used for the square, and $10 - x$ the length used for the triangle. Each side of the square is $x/4$, and its area is $x^2/16$. Each side of the triangle is $(10 - x)/3$, and its area is $\frac{\sqrt{3}}{4}\frac{(10-x)^2}{3^2}$. The total area enclosed is

$$A = \frac{1}{16}x^2 + \frac{\sqrt{3}}{36}(10 - x)^2,$$

with $x \in [0, 10]$. Extrema, *i.e.* minimizers and maximizers are either stationary points in $[0, 10]$ or the extremes of the intervals. The function A has only a stationary point (since it is a quadratic function), namely $x = \frac{80\sqrt{3}}{18+8\sqrt{3}} \approx 4.35$. Note now that

$$x = 0 \Rightarrow A = \frac{100\sqrt{3}}{36} \approx 4.81, \qquad x \approx 4.35 \Rightarrow A \approx 2.72, \qquad x = 10 \Rightarrow A = 6.25.$$

Hence, to have the minimum area, use 4.35m of wire for the square and the rest for the triangle, to have maximum area use all the wire for the square!

Exercise 49 Consider the optimisation problem

$$\begin{cases} \min\limits_{x_1,x_2} \ -x_1 + x_2, \\ 0 \le x_1 \le 1, \\ x_2 \ge x_1^2. \end{cases}$$

a) State first order necessary conditions of optimality for this constrained optimisation problem.

b) Using the conditions in part a) determine a candidate optimal solution $x^\star$ for the considered optimisation problem.

c) Transform the optimization problem into an optimization problem with equality constrains by adding an auxiliary variable and disregarding, for simplicity, the constraints $0 \le x_1 \le 1$.

State first order necessary conditions of optimality for this transformed problem. Determine a candidate optimal solution and show that it coincides with the solution determined in part b).

Solution 49

a) Define the Lagrangian

$$L(x_1, x_2, \rho_1, \rho_2, \rho_3) = -x_1 + x_2 + \rho_1(-x_1) + \rho_2(x_1 - 1) + \rho_3(x_1^2 - x_2).$$

The first order necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x_1} = -1 - \rho_1 + \rho_2 + 2\rho_3 x_1, \qquad 0 = \frac{\partial L}{\partial x_2} = 1 - \rho_3,$$

$$-x_1 \le 0, \qquad x_1 - 1 \le 0, \qquad x_1^2 - x_2 \le 0, \qquad \rho_1 \ge 0, \qquad \rho_2 \ge 0, \qquad \rho_3 \ge 0,$$

$$\rho_1 x_1 = 0, \qquad \rho_2(x_1 - 1) = 0, \qquad \rho_3(x_1^2 - x_2) = 0.$$

b) To begin with note that ρ_3 has to be equal to one, hence $x_1^2 = x_2$. Consider now the following four possibilities.

- $\rho_1 = 0$, $\rho_2 = 0$. This yields $x_1 = 1/2$ and $x_2 = 1/4$.
- $\rho_1 = 0$, $\rho_2 > 0$. This yields $x_1 = 1$, $x_2 = 1$, $\rho_2 = -1$, which is not admissible.
- $\rho_1 > 0$, $\rho_2 = 0$. This yields $x_1 = 0$, $x_2 = 0$, $\rho_1 = -1$, which is not admissible.
- $\rho_1 > 0$, $\rho_2 > 0$. This yields $x_1 = 0$ and $x_1 = 1$, which is meaningless.

In summary the only candidate solution is $x_1 = 1/2$, $x_2 = 1/4$, $\rho_1 = 0$, $\rho_2 = 0$, $\rho_3 = 1$.

c) The inequality constraint $x_1^2 - x_2 \le 0$ can be rewritten as $x_1^2 - x_2 + y^2 = 0$, where y is an auxiliary variable. The problem is thus transformed into the problem

$$\begin{cases} \min\limits_{x_1,x_2,y} \ -x_1 + x_2, \\ x_1^2 - x_2 + y^2 = 0. \end{cases}$$

The Lagrangian for this problem is $L = -x_1 + x_2 + \lambda(x_1^2 - x_2 + y^2)$, and the necessary conditions of optimality are $0 = \frac{dL}{dx_1} = -1 + 2\lambda x_1$, $\quad 0 = \frac{dL}{dx_2} = 1 - \lambda$, $\quad 0 = \frac{dL}{dy} = 2\lambda y$, $\quad x_1^2 - x_2 + y^2 = 0. The only candidate solution is \lambda = 1$, $y = 0$, $x_1 = 1/2$, $x_2 = 1/4$, which coincides with the one determined in part b).

Exercise 50 Consider the optimisation problem

$$\begin{cases} \min_{y_1, y_2} \; y_1 y_2, \\ y_1^2 + y_2^2 \leq 1. \end{cases}$$

a) Sketch in the (y_1, y_2)-plane the admissible set and the level lines of the the function $y_1 y_2$. Hence, using only graphical considerations determine the optimal solutions of the considered problem.

b) State first order necessary conditions of optimality for this constrained optimisation problem.

c) Using the conditions derived in part b) compute candidate optimal solutions. Show that the optimal solutions derived graphically in part a) satisfy the necessary conditions of optimality.

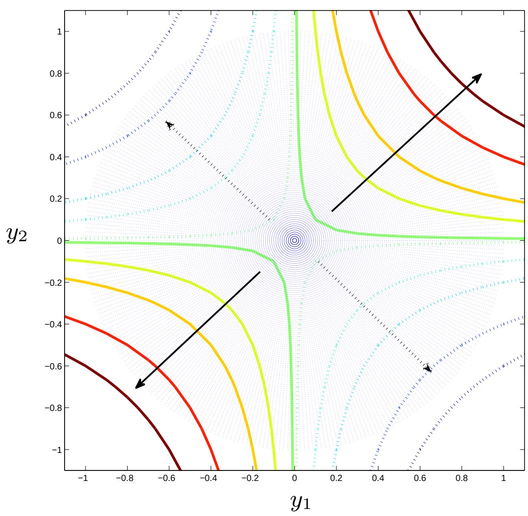
d) The considered problem can be transformed into a linear programming problem using the change of variable $x_1 = (y_1 - y_2)^2, x_2 = (y_1 + y_2)^2$.

   i) Write the equations describing the transformed problem.
   (Hint: note that the transformed problem has three inequality constraints.)

   ii) Sketch in the (x_1, x_2)-plane the admissible set and the level lines of the cost function. Hence determine the optimal solution of the transformed problem.

   iii) Show how the optimal solution of the transformed problem can be used to determine the optimal solutions of the original problem.

Solution 50



a) The admissible set is the shaded area in the figure above. The level lines are the solid (positive values of f) and dotted (negative values of f) lines. The value of the function f increases in the direction of the solid arrows, and decreases in the direction of the dotted arrows. The solution of the problem is obtained for negative values of f at a point in which the level lines are tangent to the circle $y_1^2 + y_2^2 = 1$. At such points $y_1 = -y_2$, hence the (global) minimizers are the point

$$P_1 = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right), \qquad P_2 = \left(\frac{-\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right).$$

The value of the function at the optimal points is $f(P_1) = f(P_2) = -\frac{1}{2}$.

b) The Lagrangian of the problem is

$$L(y_1, y_2, \rho) = y_1 y_2 + \rho(y_1^2 + y_2^2 - 1).$$

The first order necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial y_1} = y_2 + 2\rho y_1, \qquad 0 = \frac{\partial L}{\partial y_2} = y_1 + 2\rho y_2,$$

$$y_1^2 + y_2^2 - 1 \leq 0, \qquad \rho \geq 0, \qquad \rho(y_1^2 + y_2^2 - 1) = 0.$$

c) Consider the two cases.

- $\rho = 0$. In this case $y_1 = y_2 = 0$.
- $\rho > 0$. Consider the equations
$$0 = \frac{\partial L}{\partial y_1} = \frac{\partial L}{\partial y_2}.$$
For any $\rho > 0$, $y_1 = y_2 = 0$ is a solution. In addition, if $\rho = 1/2$ there are infinitely many solutions of the form $(y_1, y_2) = (\alpha, -\alpha)$, where α is any real number. Note now that $\rho > 0$ implies, by the complementarity condition, $y_1^2 + y_2^2 - 1 = 0$. Hence, $2\alpha^2 = 1$, yielding $\alpha = \pm\frac{\sqrt{2}}{2}$.

In summary the candidate optimal solutions are

- $(y_1, y_2) = (0, 0)$, with $\rho \geq 0$.
- P_1 and P_2 with $\rho = 1/2$.

d) i) Note that x_1 and x_2 are non-negative by definition and that $x_1 = y_1^2 - 2y_1 y_2 + y_2^2$, and $x_2 = y_1^2 + 2y_1 y_2 + y_2^2$. Hence

$$\frac{x_2 - x_1}{4} = y_1 y_2, \qquad \frac{x_1 + x_2}{2} = y_1^2 + y_2^2.$$

As a result, in the variables x_1 and x_2 the problem is

$$\begin{cases} \displaystyle\min_{x_1, x_2} \frac{x_2 - x_1}{4}, \\ x_1 \geq 0, \\ x_2 \geq 0, \\ \displaystyle\frac{x_1 + x_2}{2} \leq 1. \end{cases}$$

ii) The admissible set is the shaded area in the figure below. The level lines are the solid (positive values of f) and dotted (negative values of f) lines. The value of the function f increases in the direction of the solid arrow. The optimal solution is the point

$$P = (2, 0).$$

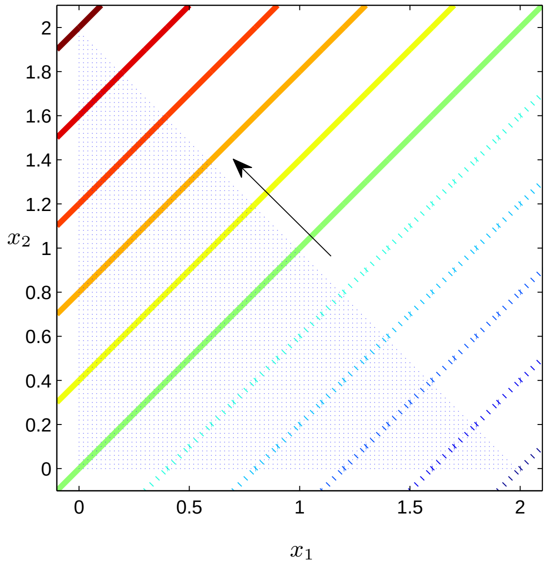
iii) The point P in the (x_1, x_2)-variables is transformed into points in the (y_1, y_2)-variables solving the equations

$$(y_1 - y_2)^2 = 2 \qquad (y_1 + y_2)^2 = 0.$$

These equations have the solutions

$$(y_1, y_2) = \left(\pm\frac{\sqrt{2}}{2}, \mp\frac{\sqrt{2}}{2}\right),$$

which coincide with the points P_1 and P_2 determined in part c).


2
1.8
1.6
1.4
1.2
x_2
1
0.8
0.6
0.4
0.2
0
0
0.5
1
1.5
2
x_1

Chapter 4

Global optimization

4.1 Introduction

Given a function $f : \mathbb{R}^n \to \mathbb{R}$, global optimization methods aim at finding the global minimum of f, *i.e.* a point $x^\star$ such that

$$f(x^\star) \le f(x)$$

for all $x \in \mathbb{R}^n$. Among these methods it is possible to distinguish between deterministic methods and probabilistic methods. In the following sections we provide a very brief introductions to global minimization methods. It is worth noting that this is an active area of research.

4.2 Deterministic methods

4.2.1 Methods for Lipschitz functions

Consider a function $f : \mathbb{R}^n \to \mathbb{R}$ and suppose it is Lipschitz with constant $L > 0$, *i.e.*

$$|f(x_1) - f(x_2)| \le L\|x_1 - x_2\|, \tag{4.1}$$

for all $x_1 \in \mathbb{R}^n$ and $x_2 \in \mathbb{R}^n$. Note that equation (4.1) implies that

$$f(x) \ge f(x_0) - L\|x - x_0\| \tag{4.2}$$

and that

$$f(x) \le f(x_0) + L\|x - x_0\|, \tag{4.3}$$

for all $x \in \mathbb{R}^n$ and $x_0 \in \mathbb{R}^n$, see Figure 4.1 for a geometrical interpretation.

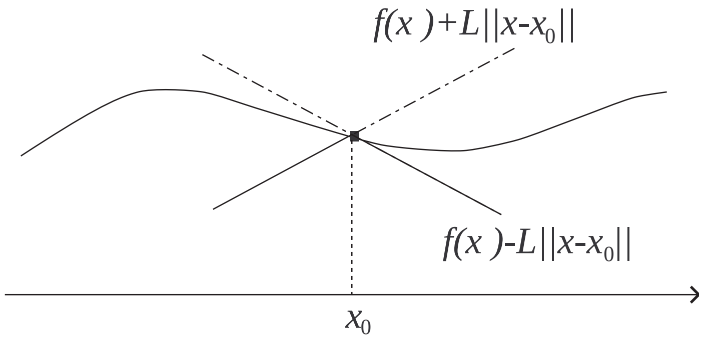


Figure 4.1: Geometrical interpretation of the Lipschitz conditions (4.2) and (4.3).

Methods for Lipschitz functions are suitable to find a global solution of the problem

$$\min_x f(x),$$

with

$$x \in I_n = \{x \in \mathbb{R}^n \mid A_i \le x_i \le B_i\},$$

and $A_i < B_i$ given, under the assumptions that the set I_n contains a global minimizer of f, the function f is Lipschitz in I_n and the Lipschitz constant L of f in I_n is known. Under these assumptions it is possible to construct a very simple global minimization algorithm, known as Schubert-Mladineo algorithm, as follows.

Step 0. Given $x_0 \in I_n$ and $\tilde{L} > L$.

Step 1. Set $k = 0$.

Step 2. Let

$$F_k(x) = \max_{j=0,\cdots,k} \{f(x_j) - \tilde{L}\|x - x_j\|\}$$

and compute x_{k+1} such that

$$F_k(x_{k+1}) = \min_{x \in I_n} F_k(x).$$

Step 4. Set $k = k + 1$ and go to **Step 2**.

Remark. The functions F_k in **Step 2** of the algorithm have a very special form. This can be exploited to construct special algorithms solving the problem

$$\min_{x \in I_n} F_k(x)$$

in a finite number of iterations. $\diamond$

For Schubert-Mladineo algorithm it is possible to prove the following statement.

Theorem 24 *Let $f^\star$ be the minimum value of f in I_n, let $x^\star$ be such that $f(x^\star) = f^\star$ and let $F_k^\star$ be the minima of the functions F_k in I_n. Let*

$$\Phi = \{x \in I_n \mid f(x) = f^\star\}$$

and let $\{x_k\}$ be the sequence generated by the algorithm. Then

- $\lim_{k \to \infty} f(x_k) = f^\star$*;*
- *the sequence $\{F_k^\star\}$ is non-decreasing and $\lim_{k \to \infty} F_k^\star = f^\star$;*
- $\lim_{k \to \infty} \inf_{x \in \Phi} \|x - x_k\| = 0$*;*
- $f(x_k) \geq f^\star \geq F_{k-1}(x_k)$*.*

Schubert-Mladineo algorithm can be given, if $x \in I_1 \subset I\!R$, a simple geometrical interpretation, as shown in Figure 4.2.
The main advantage of Schubert-Mladineo algorithm is that it does not require the computation of derivatives, hence it is also applicable to functions which are not everywhere

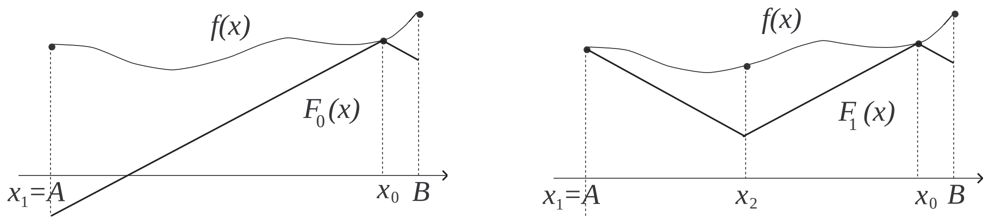


Figure 4.2: Geometrical interpretation of Schubert-Mladineo algorithm.

differentiable. Moreover, unlike other global minimization algorithms, it is possible to prove the convergence of the sequence $\{x_k\}$ to the global minimizer. Finally, it is possible to define a simple *stopping* condition. For, note that if $\{x_k\}$ and $\{F_k^\star\}$ are the sequences generated by the algorithm, then

$$f(x_k) \geq f^\star \geq F_k^\star$$

and

$$f(x_k) \geq f^\star \geq f(x_k) + r_k,$$

where $r_k = F_k^\star - f(x_k)$ and $\lim_{k\to\infty} r_k = 0$. As a result, if $|r_k| < \epsilon$, for some $\epsilon > 0$, the point x_k gives a good approximation of the minimizer of f.
The main disadvantage of the algorithm is in the assumption that the set I_n contains a global minimizer of f in $I\!R^n$. Moreover, it may be difficult to compute the Lipschitz constant L.

4.2.2 Methods of the trajectories

The basic idea of the global optimization methods known as methods of the trajectories is to construct trajectories which go through all local minimizers. Once all local minimizers are determined, the global minimizer can be easily isolated. These methods have been originally proposed in the 70's, but only recently, because of increased computer power and of a reformulation using tools from differential geometry, they have proved to be effective.
The simplest and first method of the trajectories is the so-called Branin method. Consider the function f and assume ∇f is continuous. Fix x_0 and consider the differential equations

$$\frac{d}{dt}\nabla f(x(t)) = \pm\nabla f(x(t)) \qquad x(0) = x_0. \tag{4.4}$$

The solutions $x(t)$ of such differential equations are such that

$$\nabla f(x(t)) = \nabla f(x_0)e^{\pm t},$$

i.e. $\nabla f(x(t))$ is parallel to $\nabla f(x_0)$ for all t. Using these facts it is possible to describe Branin algorithm.

Step 0. Given x_0.

Step 1. Compute the solution $x(t)$ of the differential equation

$$\frac{d}{dt}\nabla f(x(t)) = -\nabla f(x(t))$$

with $x(0) = x_0$.

Step 2. The point $x^\star = \lim_{t\to\infty} x(t)$ is a stationary point of f, in fact $\lim_{t\to\infty} \nabla f(x(t)) = 0$.

Step 3. Consider a perturbation of the point $x^\star$, *i.e.* the point $\tilde{x} = x^\star + \epsilon$ and compute the solution $x(t)$ of the differential equation

$$\frac{d}{dt}\nabla f(x(t)) = \nabla f(x(t)).$$

Along this trajectory the gradient $\nabla f(x(t))$ increases, hence the trajectory escapes from the *region of attraction* of x_0.

Step 4. Fix $\bar{t} > 0$ and assume that $x(\bar{t})$ is sufficiently away from x_0. Set $x_0 = x(\bar{t})$ and go to **Step 1**.

Note that, if the perturbation ϵ and the time $\bar{t}$ are properly selected, at each iteration the algorithm generates a new stationary point of the function f.

Remark. If $\nabla^2 f$ is continuous then the differential equations (4.4) can be written as

$$\dot{x}(t) = \pm\left[\nabla^2 f(x(t))\right]^{-1} \nabla f(x(t)).$$

Therefore Branin method is a continuous equivalent of Newton method. Note however that, as $\nabla^2 f(x(t))$ may become singular, the above equation may be meaningless. In such a case it is possible to modify Branin method using ideas borrowed from quasi-Newton algorithms. $\diamond$

Branin method is very simple to implement. However, it has several disadvantages.

- It is not possible to prove convergence to the global minimizer.
- Even if the method yields the global minimizer, it is not possible to know how many iterations are needed to reach such a global minimizer, *i.e.* there is no stopping criterion.
- The trajectories $x(t)$ are attracted by all stationary points of f, *i.e.* both minimizers and maximizers.
- There is not a systematic way to select ϵ and $\bar{t}$.

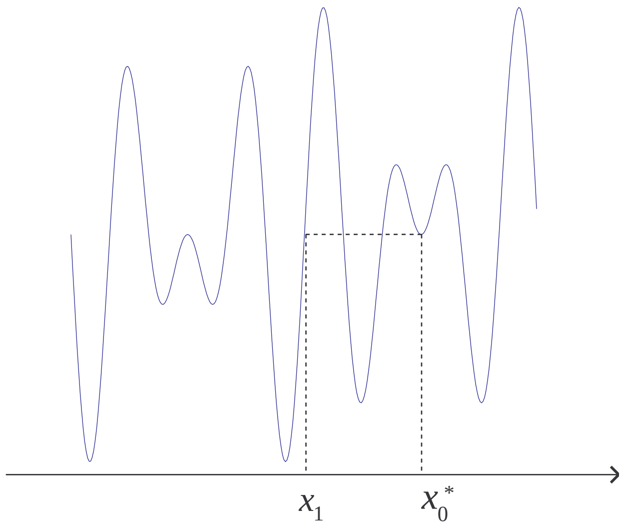


Figure 4.3: Interpretation of the tunneling phase.

4.2.3 Tunneling methods

Tunneling methods have been proposed to find, in an efficient way, the global minimizer of a function with several (possibly thousands) of local minimizers.
Tunneling algorithms are composed of a sequence of cycles, each having two phases. The first phase is the minimization phase, *i.e.* a local minimizer is computed. The second phase is the tunneling phase, *i.e.* a new starting point for the minimization phase is computed.

Minimization phase

Given a point x_0, a local minimization, using any unconstrained optimization algorithm, is performed. This minimization yields a local minimizer $x_0^\star$.

Tunneling phase

A point $x_1 \neq x_0^\star$ such that

$$f(x_1) = f(x_0^\star)$$

is determined. See Figure 4.3 for a geometrical interpretation.

In theory, tunneling methods generate a sequence $\{x_k^\star\}$ such that

$$f(x_{k+1}^\star) \leq f(x_k^\star)$$

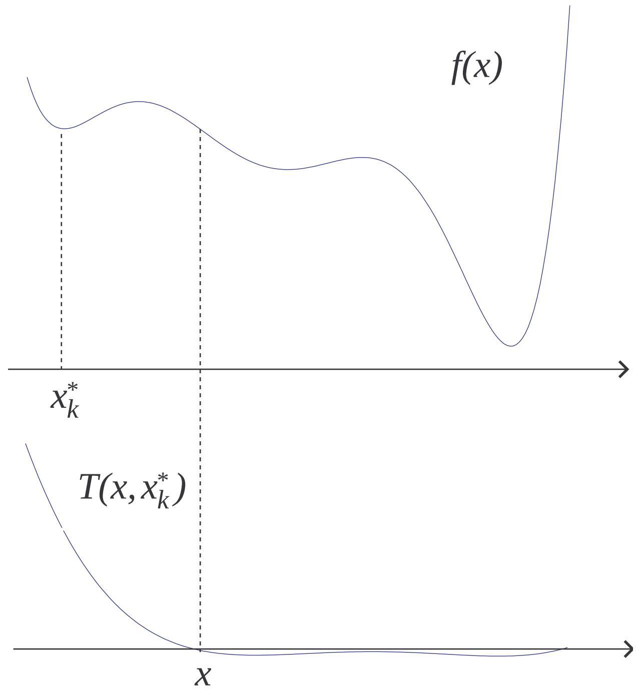


Figure 4.4: The functions $f(x)$ and $T(x, x_k^\star)$.

and the sequence $\{x_k^\star\}$ converges to the global minimizer without *passing* through all local minimizers. This is the most important advantage of tunneling methods. The main disadvantage is the difficulty in performing the tunneling phase. In general, given a point $x_k^\star$ a point x such that $f(x) = f(x_k^\star)$ is constructed searching for a zero of the function (see Figure 4.4)

$$T(x, x_k^\star) = \frac{f(x) - f(x_k^\star)}{\|x - x_k^\star\|^{2\lambda}},$$

where the parameter $\lambda > 0$ has to be selected such that $T(x_k^\star, x_k^\star) > 0$.
Finally, it is worth noting that tunneling methods do not have a stopping criterion, *i.e.* the algorithm attempts to perform the tunneling phase even if the point $x_k^\star$ is a global minimizer.

4.3 Probabilistic methods

4.3.1 Methods using random directions

In this class of algorithms at each iteration a randomly selected direction, having unity norm, is selected. The theoretical justification of such an algorithm rests on Gaviano theorem. This states that the sequence $\{x_k\}$ generated using the iteration

$$x_{k+1} = x_k + \alpha_k d_k,$$

where d_k is randomly selected on a unity norm sphere and α_k is such that

$$f(x_k + \alpha_k d_k) = \min_{\alpha} f(x_k + \alpha d_k),$$

is such that for any $\epsilon > 0$ the probability that

$$f(x_k) - f^\star < \epsilon,$$

where $f^\star$ is a global minimum of f, tends to one as $k \to \infty$.

4.3.2 Multistart methods

Multistart methods are based on the fact that for given sets D and A, with measures $m(D)$ and $m(A)$, and such that

$$1 \geq \frac{m(A)}{m(D)} = \alpha \geq 0,$$

the probability that, selecting N random points in D, one of these points is in A is

$$P(A, N) = 1 - (1 - \alpha)^N.$$

As a result

$$\lim_{N \to \infty} P(A, N) = 1.$$

Therefore, if A is a neighborhood of a global minimizer of f in D, we conclude that, selecting a sufficiently large number of random points in D, one of these will (almost surely) be close to the global minimizer. Using these considerations it is possible to construct a whole class of algorithms, with similar properties, as detailed hereafter.

Step 0. Set $f^\star = \infty$.

Step 1. Select a random point $x_0 \in I\!R^n$.

Step 2. If $f(x_0) > f^\star$ go to **Step 1**.

Step 3. Perform a local minimization starting from x_0 and yielding a point $x_0^\star$. Set $f^\star = f(x_0^\star)$.

Step 4. Check if $x_0^\star$ satisfies a stopping criterion. If not, go to **Step 1**.

4.3.3 Stopping criteria

The main disadvantage of probabilistic algorithms is the lack of a theoretically sound stopping criterion. The most promising and used stopping criterion is based on the construction of a *probabilistic approximation* $\tilde{P}(w)$ of the function

$$P(w) = \frac{m(\{x \in D \mid f(x) \leq w\})}{m(D)}.$$

Once the function $\tilde{P}(w)$ is known, a point $x^\star$ is regarded as a good approximation of the global minimizer of f if

$$\tilde{P}(f(x^\star)) \leq \epsilon \ll 1.$$

4.4 Exercises

Similarly to Sections 2.10 and 3.7, this section contains some exercises to illustrate how global minimization methods can be used.

Exercise 51 Consider the discrete time system

$$x_{k+1} = ax_k$$

with $x_k \in \mathbb{R}$, and output $y_k = x_k$. Consider also the auxiliary discrete time system

$$\xi_{k+1} = \alpha\xi_k$$

with $\xi_k \in \mathbb{R}$, output $\eta_k = \xi_k$, and such that $\xi_0 = x_0 \neq 0$.
Consider now the problem of determining the constant α such that the cost

$$J(\alpha) = \frac{1}{2}\left(e_1^2 + e_2^2 + \cdots e_N^2\right)$$

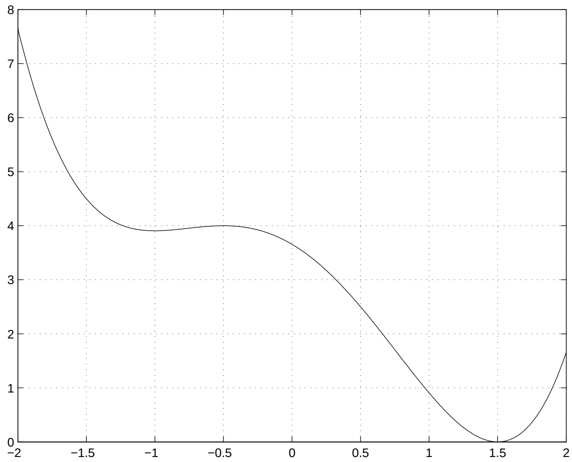
is minimized, where $e_i = y_i - \eta_i$ and $N \geq 1$.

a) Pose the above problem as an unconstrained optimization problem in the decision variable α, parametrized by a and x_0.

b) Assume $N = 1$. Show that $J(a) = 0$ and $J(\alpha) > 0$ for all $\alpha \neq a$. Hence show that the function $J(\alpha)$ has a unique local minimizer which is also a global minimizer.

c) Suppose $N = 2$. Compute the stationary points of $J(\alpha)$. Note that the number of stationary points is a function of the value of a. Hence, determine the local minimizers and the local maximizers of the function $J(\alpha)$.

d) For $N = 2$ and $a = 3/2$, the function $J(\alpha)$ is as shown in the figure below. Let $L = 12$ be the Lipschitz constant of $J(\alpha)$ for $\alpha \in [-2, 2]$. Apply four steps of the Schubert-Mladineo algorithm for the minimization of the function $J(\alpha)$ assuming that a global minimizer is in the set $I_1 = \{\alpha \in \mathbb{R} \mid -2 \leq \alpha \leq 2\}$ and that the starting point of the algorithm is selected to be $\alpha = 2$.



Solution 51

a) The problem can be formulated as

$$\min_{\alpha\in\mathbb{R}} J(\alpha) = \min_{\alpha\in\mathbb{R}} \frac{1}{2}(a-\alpha)^2x_0^2 + (a^2-\alpha^2)^2x_0^2 + \cdots + (a^N-\alpha^N)^2x_0^2,$$

i.e. as an unconstrained optimization problem in the decision variable α and parameterized by a and x_0.

b) For $N = 1$ one has $J(\alpha) = \frac{1}{2}(a-\alpha)^2 x_0^2$. Hence, $J(a) = 0$ and $J(\alpha) > 0$ for all $\alpha \neq a$. This shows (using the very definition of global minimum) that $\alpha = a$ is a global minimum.

c) If $N = 2$ one has $J(\alpha) = \frac{1}{2}\left((a-\alpha)^2 + (a^2-\alpha^2)^2\right) x_0^2$. Hence $\frac{dJ(\alpha)}{d\alpha} = -(a-\alpha)(2\alpha^2 + 2\alpha a + 1)x_0^2$. Therefore, the stationary points are

$$P_1 = a, \qquad P_2 = -\frac{a}{2} + \frac{\sqrt{a^2-2}}{2}, \qquad P_3 = -\frac{a}{2} - \frac{\sqrt{a^2-2}}{2}.$$

We conclude that, if $|a| < \sqrt{2}$ there is only one stationary point, whereas if $|a| \geq \sqrt{2}$ there are three stationary points. Computing second derivatives we have that P_1 is always a local minimizer, and, for $|a| \geq \sqrt{2}$, P_2 is a local maximizer and P_3 is a local minimizer.

d) A sketch of the application of the Schubert-Mladineo algorithm is shown at the end of the chapter. Note that x_4 is very close to the global minimizer.

Exercise 52 Consider the function

$$f = x_1^4 - x_1 x_2 + x_2^4$$

and the problem of finding its global minimizer.

a) Write the formulae for the so-called Branin system, that is the system

$$\dot{x} = -[\nabla^2 f]^{-1} \nabla f,$$

for the considered function f.

b) Compute the equilibria of the Branin system determined in part a). Show that these equilibria coincide with the stationary points of the function f. Show that f is radially unbounded. Hence determine the global minimizer of f.

c) Consider the linearization of the Branin system, computed in part a), around its equilibrium at $x = 0$. Show that this linearized system has two eigenvalues equal to -1, hence deduce that the point $x = 0$ is locally attractive.

d) Write now the formulae for the modified Branin system

$$\dot{x} = -\det(\nabla^2 f)[\nabla^2 f]^{-1} \nabla f,$$

for the function f above. Consider the linearization of the modified Branin system at $x = 0$ and show that this equilibrium point is unstable.

e) Give reasons for the modified Branin method being preferable to the Branin method when determining a global minimizer for the considered function f.

Solution 52

a) The Branin system is

$$\dot{x} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \frac{1}{144 x_1^2 x_2^2 - 1} \begin{bmatrix} -48 x_2^2 x_1^3 - 8 x_2^3 - x_1 \\ 8 x_1^3 + x_2 - 48 x_1^2 x_2^3 \end{bmatrix}.$$

b) The equilibria of the Branin system are $P_1 = (0,0)$, $P_2 = (1/2, 1/2)$ and $P_3 = (-1/2, -1/2)$. Note now that these are also such that $\nabla f(P_i) = 0$, for $i = 1, 2, 3$. Hence, the equilibria of Branin system coincide with the stationary points of f. Note now that $\lim_{\|x\| \to \infty} f(x) = +\infty$, hence f is radially unbounded. Moreover $f(P_1) = 0$ and $f(P_2) = f(P_3) = -1/8$. Hence the global minimum of f is $-1/8$ and there are two global minimizers, P_2 and P_3.

c) The linearization of the Branin system around the point P_1 is described by

$$\dot{x} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} x.$$

The linearized system has two eigenvalues equal to -1 and this shows that the point P_1 is locally attractive.

d) The modified Branin system is

$$\dot{x} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -48x_2^2x_1^3 - 8x_2^3 - x_1 \\ 8x_1^3 + x_2 - 48x_1^2x_2^3 \end{bmatrix}.$$

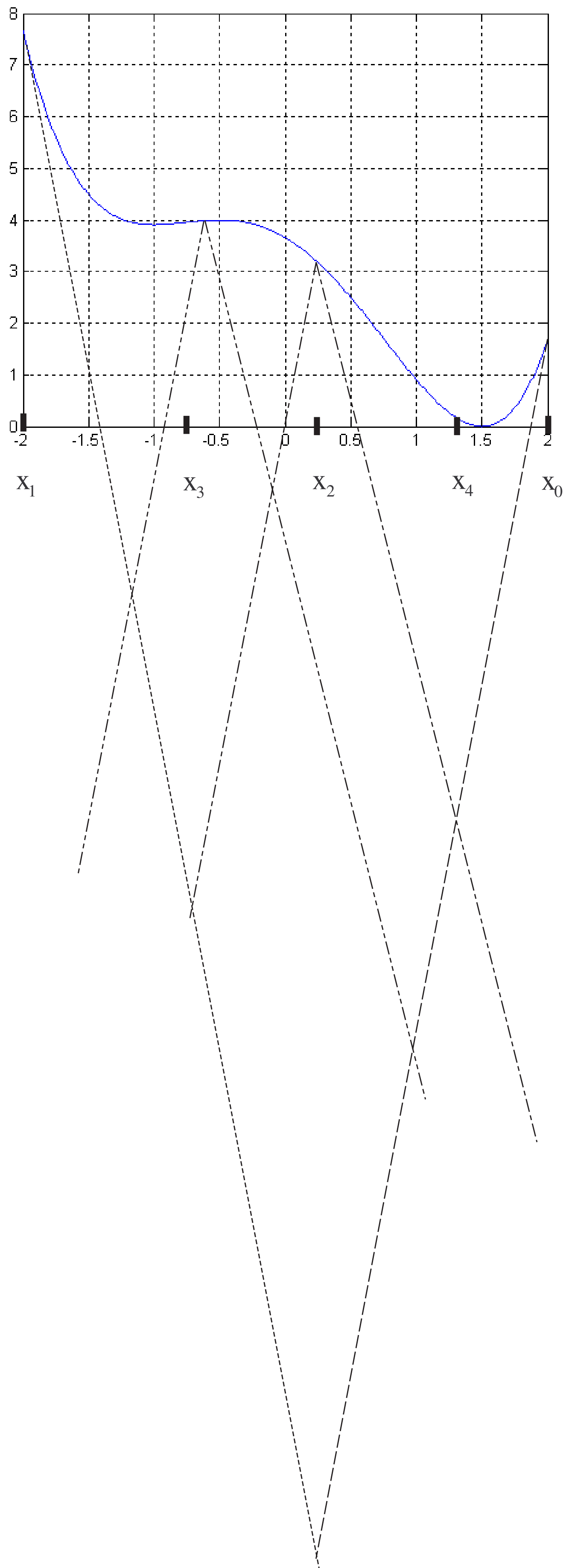
Its linearization around P_1 is described by

$$\dot{x} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x,$$

and this shows that the point P_1 is an unstable equilibrium of the modified Branin system.

e) The modified Branin system has the following advantages:

- the differential equations are defined for all x;
- the point P_1, which is a local maximizer, is unstable therefore almost all trajectories of the system are not be *attracted* by P_1.



8
7
6
5
4
3
2
1
0
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
x_1
x_3
x_2
x_4
x_0