School of Computer Science and Applied Mathematics

APPM 3040A: Scientific Computing III

Chapter 1 : Part 1 - Introduction to Constrained Optimization

Lecturer: *Matthews Sejeso* ***Date:*** *February 2025*

By the completion of this lecture you should be able to:

1. Describe first order optimality conditions for constrained optimization.
2. Describe second order optimality conditions for constrained optimization.
3. Use first and second order optimality conditions to solve constrained optimization problems.

Reference:

- *Chapter 12, Jorge Nocedal and Stephen J. Wright, 'Numerical Optimization'.*

1.1 Introduction

In this chapter we consider *constrained* nonlinear problems - we minimize the objective function subject to constraints on the variables. A general formulation for these problems is

$$\min_{\mathbf{x}} f(\mathbf{x}) \quad \text{subject to} \begin{cases} c_i(\mathbf{x}) = 0 & i \in E, \\ c_i(\mathbf{x}) \geq 0 & i \in I, \end{cases} \tag{1.1}$$

where f and c_i are continuously differentiable, real-valued functions on a subset of $\mathbb{R}^n$, and E and I are finite sets of equality and inequality indices respectively. A point $\mathbf{x} \in \mathbb{R}^n$ is called a *feasible point* if it satisfy all constraints simultaneously. A collection of all feasible points form a *feasible set*, Ω, associated with problem (1.1):

$$\Omega = \{\mathbf{x} \in \mathbb{R}^n : c_i(\mathbf{x}) = 0 \text{ for } i \in E; \quad c_i(\mathbf{x}) \geq 0 \text{ for } i \in I\}. \tag{1.2}$$

The objective of constrained optimization is to find a feasible point $\mathbf{x}^* \in \Omega$ that attains the smallest value of the objective function $f(\mathbf{x})$. This point, $\mathbf{x}^*$, is referred to as the local minimizer.

The primary goal of this chapter is to derive the optimality conditions for constrained optimization problem (1.1). These conditions are of paramount importance as they underpin the development of algorithms designed for resolving such problems.

The derivation of optimality conditions for constrained optimization problems is more complicated than the derivation for unconstrained optimization problem. Recall the unconstrained optimization problem:

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}). \tag{1.3}$$

The optimality conditions for unconstrained optimization problems (1.3) are summarised as follows following:

- *Necessary conditions:* For unconstrained problem, if $\mathbf{x}^*$ is the local minimizer and f is continuously differentiable in the open neighbourhood of $\mathbf{x}^*$, then $\nabla f(\mathbf{x}^*) = \mathbf{0}$ and $\nabla^2 f(\mathbf{x}^*)$ is positive semidefinite.
- *Sufficient conditions:* Suppose that $\nabla^2 f$ is continuous in an open neighbourhood of $\mathbf{x}^*$ and that $\nabla f(\mathbf{x}^*) = \mathbf{0}$ and $\nabla^2 f(\mathbf{x}^*)$ is positive definite. Then $\mathbf{x}^*$ is a strict local minimizer of f.

1.2 First Order Optimality conditions

To introduce the basic principles behind the characterization of solutions of constrained optimization problems, we work through two simple examples. The ideas discussed in the examples will be made rigorous to define the first order optimality conditions for constrained optimization. We start by noting one item of terminology that recurs throughout this document: At a feasible point $\mathbf{x}$, the inequality constraint $i \in I$ is said to be active if $c_i(\mathbf{x}) = 0$ and inactive if the strict inequality $c_i(\mathbf{x}) > 0$ is satisfied.

1.2.1 Equality constrains

The first example is a two-variable problem with a single equality constraint:

$$\min_{x} x_1 + x_2, \qquad \text{subject to } x_1^2 + x_2^2 = 2. \tag{1.4}$$

In the language of (1.1), we have $f(\mathbf{x}) := x_1 + x_2$, and $c_1 = x_1^2 + x_2^2 - 2$, therefore $I = \emptyset$, and $E = \{1\}$. We can see by inspection that the feasible set for this problem is the circle of radius $\sqrt{2}$ centred at the origin$-$just the boundary of this circle, not its interior. Obviously, $\mathbf{x}^* = (-1, -1)$ is the optimal solution. From any other point on the circle, it is easy to find a way to move that stays feasible (i.e remains on the circle) while decreasing f. For instance from the point $\mathbf{x} = (\sqrt{2}, 0)$ any move in the clockwise direction around the circle has desired effect.

Notice, that at the optimal solution $\mathbf{x}^*$, the constrained normal $\nabla c_1(\mathbf{x}^*)$ is parallel to $\nabla f(\mathbf{x}^*)$. That is, there is a scalar λ_1^* such that

$$\nabla f(\mathbf{x}^*) = \lambda_1^* \nabla c_1(\mathbf{x}^*). \tag{1.5}$$

In particular, for this example $\lambda_1^* = -\frac{1}{2}$.

We now show that equation (1.5) is a necessary condition for optimality is the general case. Consider the Taylor series approximation of the objective and constraint functions. Unlike in unconstrained optimization problems, the direction $\mathbf{d}$ cannot be arbitrary. To retain feasibility with respect to the function $c_1(\mathbf{x}) = 0$, we require any small step $\mathbf{d}$ to satisfy $c_1(\mathbf{x} + \mathbf{d}) = \mathbf{0}$; that is

$$\mathbf{0} = c_1(\mathbf{x} + \mathbf{d}) \approx c_1(\mathbf{x}) + \nabla c_1(\mathbf{x})^T d = \nabla c_1(\mathbf{x})^T d. \tag{1.6}$$

Hence, the direction d retains feasibility with respect to equality constraint c_1, to first order, when it satisfies

$$\nabla c_1(\mathbf{x})^T \mathbf{d} = 0. \tag{1.7}$$

On the other hand, a direction of improvement must produce a decrease of the objective function f, so that

$$f(\mathbf{x} + \mathbf{d}) - f(\mathbf{x}) \approx \nabla f(\mathbf{x})^T \mathbf{d} < 0. \tag{1.8}$$

Hence, if there exists a direction $\mathbf{d}$ that satisfies both equation (1.7) and inequality (1.8), we conclude that improvement on the current point $\mathbf{x}$ is possible. It follows that a necessary condition for optimality for problem (1.4) is that *there exist no* direction $\mathbf{d}$ satisfying both equation (1.7) and inequality (1.8) simultaneously. The only way such a direction does not exist is when $\nabla f(\mathbf{x})$ and $\nabla c_1(\mathbf{x})$ are parallel. That is, if the condition $\nabla f(\mathbf{x}) = \lambda_1 \nabla c_1(\mathbf{x})$ holds at $\mathbf{x}$, for some scalar λ_1.

Hence, we introduce the *Lagrangian function*

$$L(\mathbf{x}, \lambda_1) = f(\mathbf{x}) - \lambda_1 c_1(\mathbf{x}). \tag{1.9}$$

Note that $\nabla_x L(\mathbf{x}, \lambda_1) = \nabla f(\mathbf{x}) - \lambda_1 \nabla c_1(\mathbf{x})$, we can state the optimality condition (1.5) equivalently as follows: at the optimal point $\mathbf{x}^*$, there is a scalar λ_1 such that

$$\nabla_x L(\mathbf{x}, \lambda_1) = \mathbf{0}. \tag{1.10}$$

This observation suggests that we can search for solutions of equality constrained problem (1.1) by searching for stationary points of the Lagrangian function. The scalar quantity λ_1 is called the Lagrange multiplier for the constraint $c_1(\mathbf{x}) = 0$.

Though the condition (1.5) (equivalently (1.10)) appears to be necessary for an optimal solution of problem (1.4), it is clearly not sufficient. For instance, in problem (1.4), the condition (1.5) is satisfied at the point $x = (1,1)$ with $\lambda_1 = \frac{1}{2}$. This point is obviously not a solution–in fact, it maximizes the objective function f on the circle. Moreover, in the case of equality constrained problems, we cannot turn condition (1.5) into sufficient condition simply by placing some restriction on the sign of λ_1. To see this, consider replacing the constraint $x_1^2 + x_2^2 - 2 = 0$ by its negative $2 - x_1^2 + x_2^2$ in problem (1.4). The solution of the problem is not affected, but the value of λ_1 that satisfies the condition (1.5) changes form $\lambda_1 = -\frac{1}{2}$ to $\lambda_1 = \frac{1}{2}$.

1.2.2 Inequality constraint

The second example is a two-variable problem with single inequality constraint: This is a slight modification of problem (1.4), in which the equality constraint is replaced by an inequality. We consider the problem,

$$\min x_1 + x_2, \quad \text{subject to } 2 - x_1^2 + x_2^2 \leq 0, \tag{1.11}$$

for which the feasible region is the *circle and its interior*. Here the objective function $f(\mathbf{x}) = x_1 + x_2$ and $c_1(x) = 2 - x_1^2 + x_2^2$, therefore $E = \emptyset$ and $I = \{1\}$.

Note that the constraint normal $\nabla c_1(x)$ points toward the interior of the feasible region at each point on the boundary of the circle. By inspection, we see that the optimal solution is still $\mathbf{x}^* = (-1,-1)$ and the condition (1.5) holds with Lagrange multiplier $\lambda_1^* = \frac{1}{2}$. However, this inequality constrained problem differs form the equality constrained problem (1.4) in that the sign of the Lagrange multiplier plays a significant role, as we now argue.

As before, we conjecture that a given feasible point $\mathbf{x}$ is not optimal if we can find a step $\mathbf{d}$ that retains feasibility and decreases the objective function f to first order. The main difference between equality constrained problem (1.4) and inequality constrained problem (1.11) comes in the handling of the feasibility condition. As in (1.8), the direction $\mathbf{d}$ improves the objective function, to first order, if $\nabla f(\mathbf{x})^T \mathbf{d} < 0$. Meanwhile, for the inequality constraint, the direction $\mathbf{d}$ retains feasibility if

$$0 \leq c_1(\mathbf{x}) + \nabla c_1(\mathbf{x})^T \mathbf{d},$$

so to first order, feasibility is retained if the following condition holds

$$c_1(\mathbf{x}) + \nabla c_1(\mathbf{x})^T \mathbf{d} \geq 0. \tag{1.12}$$

In determining whether a direction $\mathbf{d}$ exists that satisfies both conditions (1.8) and (1.12), we consider the following two cases.

- ***Case I:*** Consider the first case in which x lines strictly inside the circle, the strict inequality $c_1(\mathbf{x}) > 0$ holds. In this vector any vector $\mathbf{d}$ satisfies the condition (1.12), provided that its length is sufficiently small. In particular, the problem is just like unconstrained optimization problem, whenever $\nabla f(\mathbf{x}^*) \neq 0$ we can obtain a direction d that satisfies both (1.8) and (1.12) simultaneously by setting

$$\mathbf{d} = -c_1(\mathbf{x}) \frac{\nabla f(\mathbf{x})}{\|\nabla f(\mathbf{x})\|}.$$

The only situation in which such direction fails to exist is when

$$\nabla f(\mathbf{x}) = 0. \tag{1.13}$$

- ***Case II:*** Now, consider the case in which x lies on the boundary of the circle, so that $c_1(x) = 0$. The conditions (1.8) and (1.12) therefore become

  $$\nabla f(\mathbf{x})^T \mathbf{d} < 0, \quad \text{and } \nabla c_1(\mathbf{x})^T \mathbf{d} \geq 0.$$

  The first case of this conditions defines an open half-space, while the second defines a closed half-space. We can easily see by inspection that the two regions fail to intersect only when $\nabla f(\mathbf{x})$ and $\nabla c_1(\mathbf{x})$ points in the same direction, that is when

  $$\nabla f(\mathbf{x}) = \lambda_1 \nabla c_1(\mathbf{x}), \quad \text{for some } \lambda_1 \geq 0. \tag{1.14}$$

  Note that the sign of the multiplier is significant here. If condition (1.5) were satisfied with a negative value of λ_1, then $\nabla f(\mathbf{x})$ and ∇c_1 would point in opposite directions. We see that the set of directions that satisfy both (1.8) and (1.12) would make up an entire open half-plane.

Similar to equality constraint, the optimality conditions for both *case I* and *case II* can be summarized neatly with reference to the Lagrangian function. When no first-order feasibility descent direction exists at some point x^* , we have that

$$\nabla_x L(\mathbf{x}^*, \lambda_1^*) = 0, \quad \text{for some } \lambda_1^* \geq 0, \tag{1.15}$$

where we also require that

$$\lambda_1^* c_1(\mathbf{x}) = 0 \tag{1.16}$$

Condition (1.16) is know as a *complementarity condition*; it implies that the Lagrange multiplier λ_1 can be strictly positive only when the corresponding constraint $c_1(\mathbf{x})$ is active. Conditions of this type play a central role in constrained optimization. Hence, (1.15) reduces to $\nabla f(\mathbf{x}^*) = 0$, as required by case I (1.13). In case II, condition (1.16) allows λ_1^* to take on a nonnegative value, so (1.15) becomes equivalent to (1.14).

1.2.3 First order Optimality conditions

The above discussion suggests that a number of conditions are important in characterization of solutions for the constrained problem (1.1). These include the relation $\nabla_x L(\mathbf{x}, \lambda) = 0$, the nonnegative of λ_i for all inequality constraints $c_i(\mathbf{x})$, and the complementarity condition $\lambda_i c_i(\mathbf{x}) = 0$ that is required for all the inequality constraints. We now generalize the observations made above to state the first-order conditions in a rigorous fashion.

In general, the Lagrangian for the constrained optimization problem (1.1) is defined as

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \sum_{i \in E \cup I} \lambda_i c_i(\mathbf{x}). \tag{1.17}$$

Note that *every constraint c_i has its associated Lagrange multiplier λ_i*. We need the following definitions to state the first-order optimality conditions.

Definition 1.1 (Active set)**.** *The active set $A(x)$ of any feasible point $\mathbf{x}$ is the union of the equality set E with the indices of active inequality constraints; that is*

$$A(\mathbf{x}) = E \cup \{i \in I | c_i(\mathbf{x}) = 0\}. \tag{1.18}$$

Next, we need to pay more attention to properties of the constraints gradients. The vector $\nabla c_i(\mathbf{x})$ is often called the normal to the constraint $c_i(\mathbf{x})$ at the point $\mathbf{x}$, because it is usually a vector that is perpendicular to the contours of the constraint c_i at $\mathbf{x}$. In the case of inequality constraint, it points toward the feasible side of the constraint. It is possible, however, that $\nabla c_i(\mathbf{x})$ vanishes due to the algebraic representation of c_i. So that the term $\lambda_i \nabla c_i(\mathbf{x})$ vanishes for all values of λ_i and does not play a role in the Lagrangian gradient $\nabla_x L(\mathbf{x}, \lambda)$.

For instance, if we replace the constraint in problem (1.4) by the equivalent condition

$$c_1(\mathbf{x}) = \left(x_1^2 + x_2^2 - 2\right)^2 = 0.$$

Then we have the gradient

$$\nabla c_1(\mathbf{x}) = \begin{bmatrix} 4\left(x_1^2 + x_2^2 - 2\right) x_1 \\ 4\left(x_1^2 + x_2^2 - 2\right) x_2 \end{bmatrix}.$$

It is obvious that $\nabla c_1(\mathbf{x}) = 0$ for all feasible points. In particular the condition $\nabla f(\mathbf{x}) = \lambda_1 \nabla c_1(\mathbf{x})$ does no longer holds at the optimal point $\mathbf{x} = (-1, -1)$. We usually make an assumption called *constrained qualification* to ensure such degenerate behaviour does not occur at the value of $\mathbf{x}$ in question. One such constrained qualification is defined here.

Definition 1.2 (LICQ)**.** *Given the point $\mathbf{x}^*$ and the active set $A(\mathbf{x}^*)$ defined by (1.18), we say that the linear independence constraint qualification (LICQ) hold if the set of active constraint gradients $\nabla c_i(\mathbf{x}^*)$, $i \in A(\mathbf{x}^*)$ is linearly independent.*

Note that if *LICQ* holds, none of the active constraint gradients can be zero. This condition allow us to state the following optimality conditions for a general nonlinear programming problem (1.1). These conditions provide the foundation for many of the algorithms designed for problem (1.1). They are called first-order conditions because they concern themselves with properties of the gradients of the objective and constraint functions.

Theorem 1.1 (First Order Necessary Conditions under LICQ)**.** *Suppose that $\mathbf{x}^*$ is a local solution of (1.1) and the LICQ holds at $\mathbf{x}^*$. Then there is a Lagrange multiplier λ^* with components λ_i^*, $i \in E \cup I$, such that the following conditions are satisfied at $(\mathbf{x}^*, \lambda^*)$*

$$\begin{aligned}
\nabla_x L(\mathbf{x}^*, \lambda^*) &= 0, && (1.19)\\
c_i(\mathbf{x}^*) &= 0, \quad \textit{for all } i \in E, && (1.20)\\
c_i(\mathbf{x}^*) &\geq 0, \quad \textit{for all } i \in I, && (1.21)\\
\lambda_i^* &\geq 0, \quad \textit{for all } i \in I, && (1.22)\\
\lambda_i^* c_i(\mathbf{x}^*)&, \quad \textit{for all } i \in E \cup I. && (1.23)
\end{aligned}$$

Here, $L(\mathbf{x}, \lambda)$ is the Lagrangian function of an constraint problem, defined as $L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \sum_i \lambda_i c_i(\mathbf{x})$.

The conditions (1.20)-(1.23) are often known as the *Karush-Kuhn-Tucker* conditions, or simply *KKT* conditions for short. They are necessary conditions for the optimum of a constrained problem. For the given problem (1.1) and solution point x^*, there may be many vectors λ_* for which the conditions (1.20)-(1.23) are satisfied. When the *LIQC* holds, however, the optimal λ^* is unique.

1.3 Second Order Optimality condition

As in the unconstrained case, the first order conditions are not sufficient to guarantee a local minimum. For this, we turn to the second order sufficient conditions. The first order conditions– the KKT conditions–tell us how the first derivatives of f and the active constraints c_i are related at the optimal solution $\mathbf{x}^*$. For the second order conditions, we are concerned with the behaviour of the Hessian of the Lagrangian, denoted $\nabla^2_{xx}L(\mathbf{x}, \lambda)$, at locations there the KKT conditions hold. In particular, we look for positive definiteness in a subspace defined by the linearised constraints. We define this subspace by F:

Definition 1.3 (Feasible space)**.** *Given a point x^* and the active constraint set $A(x^*)$, the set F is defined by*

$$w \in F(\lambda^*) \implies \begin{cases} \nabla c_i(x^*)^T w = 0, & \text{for all } i \in E \\ \nabla c_i(x^*)^T w = 0, & \text{for all } i \in A(x^*) \cap I \text{ with } \lambda_i > 0, \\ \nabla c_i(x^*)^T w \geq 0, & \text{for all } i \in A(x^*) \cap I \text{ with } \lambda_i = 0. \end{cases} \tag{1.24}$$

The subset $F(\lambda^*)$ contains the direction $\mathbf{w}$ that tend to adhere to the active inequality constraints for which the Lagrange multiplier component λ_i^* is positive, as well as to the equality constraints. The first theorem defines a necessary condition involving the second derivative: If the point $\mathbf{x}^*$ is local solution, then the curvature of the Lagrangian along directions in the feasible space $F(\lambda^*)$ must be nonnegative.

Theorem 1.2 (Second Order Necessary Conditions under LICQ)**.** *Suppose that $\mathbf{x}^*$ is a local solution of problem (1.1) and that the LICQ condition is satisfied. Let λ^* be a Lagrange multiplier vector such that the KKT conditions (1.20)-(1.23) are satisfied, and let $F(\lambda^*)$ be define as above. Then*

$$\mathbf{w}^T \nabla^2_{xx} L(\mathbf{x}^*, \lambda^*)\mathbf{w} \geq 0, \qquad \text{for all } \mathbf{w} \in F(\lambda^*). \tag{1.25}$$

The second theorem defines a sufficient condition involving the second derivative: If the point $\mathbf{x}^*$ is local solution, then the curvature of the Lagrangian along directions in $F(\lambda^*)$ must be strictly positive. Note that the constraint qualification is not required.

Theorem 1.3 (Second Order Sufficient Conditions)**.** *Suppose that for some feasible point $\mathbf{x}^*$ there is a Lagrange multiplier vector λ^* such that the KKT conditions (1.20)-(1.23) are satisfied. Suppose also that*

$$\mathbf{w}^T \nabla^2_{xx} L(\mathbf{x}^*, \lambda^*)\mathbf{w} > 0, \qquad \text{for all } \mathbf{w} \in F(\lambda^*), \quad \mathbf{w} \neq \mathbf{0}. \tag{1.26}$$

Then $\mathbf{x}^*$ is a strict local solution for problem (1.1).

If the problem is simple, the KKT conditions (1.20)-(1.23) can be used to identify the KKT points (x^*, λ^*). The points can then be classified using the second order sufficient condition (1.26).

Example 1.1. *Consider the following equality constrained problem:*

$$\min_{\mathbf{x}} f(\mathbf{x}) := x_1 + x_2 \qquad \text{subject to } c_1(x) = x_1^2 + x_2^2 - 2 = 0$$

The Lagrangian is

$$L(\mathbf{x}, \lambda) = x_1 + x_2 - \lambda_1(x_1^2 + x_2^2 - 2), \tag{1.27}$$

and the optimality conditions are

$$\nabla_x L(\mathbf{x}, \lambda) = \begin{bmatrix} 1 - 2\lambda_1 x_1 \\ 1 - 2\lambda_1 x_2 \end{bmatrix} = 0 \implies \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2\lambda_1} \\ \frac{1}{2\lambda_1} \end{bmatrix}$$
$$\nabla_{\lambda_1} L(\mathbf{x}, \lambda) = x_1^2 + x_2^2 - 2 = 0 \implies \lambda_1 = \pm\frac{1}{2}.$$

Note there is no restriction on the sign of λ_1, since we are dealing with equality constraint. Hence, we get the KKT points as $\mathbf{x} = (1,1)^T$ for $\lambda = -\frac{1}{2}$ and $\mathbf{x} = (-1,-1)^T$ for $\lambda_1 = -\frac{1}{2}$. To establish which point is a minimizer as opposed to other types of stationary points, we need to look at the second order conditions. We need to find the F set for each λ_1 value.

- *For $\lambda_1 = \frac{1}{2}$, we are only interested in the first condition of (1.24) since we are dealing with equality constraint. Thus, let $\mathbf{w} = (w_1, w_2)^T$, then*

$$\nabla c_1(\mathbf{x}) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}_{\mathbf{x}=(1,1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \implies \nabla c(\mathbf{x}^*)^T \mathbf{w} = w_1 + w_2 = 0 \implies w_2 = -w_1.$$

  Hence the set F is simply defined by

$$F(\lambda_1 = \frac{1}{2}) = \{\mathbf{w} | w_2 = -w_1\}$$

  The Hessian of the Lagrangian in the set defined by $F(\frac{1}{2})$ is

$$\mathbf{w}^T \nabla^2_{xx} L(x^*, \lambda^*) \mathbf{w} = \begin{bmatrix} w_1 & -w_1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} w_1 \\ -w_1 \end{bmatrix} = -2w_1^2 < 0.$$

  Hence the Hessian is negative definite in the set $F(\lambda_1 = \frac{1}{2})$. Therefore the KKT point $\mathbf{x} = (1, 1)$ with $\lambda_1 = \frac{1}{2}$ is not a minimizer of the problem.

- *Similarly, for $\lambda_1 = -\frac{1}{2}$, with the corresponding point $\mathbf{x} = (-1, -1)^T$, we have*

$$\nabla c_1(\mathbf{x}) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}_{\mathbf{x}=(-1,-1)} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \implies \nabla c(\mathbf{x}^*)^T \mathbf{w} = -w_1 - w_2 = 0 \implies w_2 = -w_1.$$

  Hence the set F is defined by

$$F(\lambda_1 = -\frac{1}{2}) = \{\mathbf{w} | w_2 = -w_1\}$$

  The Hessian of the Lagrangian in the set defined by $F(\lambda_1 = -\frac{1}{2})$ is

$$\mathbf{w}^T \nabla^2_{xx} L(x^*, \lambda^*) \mathbf{w} = \begin{bmatrix} w_1 & -w_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ -w_1 \end{bmatrix} = 2w_1^2 > 0.$$

  Hence the Hessian is positive definite in the set $F(-\frac{1}{2})$. Therefore the KKT point $\mathbf{x} = (-1, -1)$ with $\lambda_1 = -\frac{1}{2}$ is a strict minimizer of the problem.

Example 1.2. *Consider the constraint optimization problem*

$$\min -0.1(x_1 - 4)^2 + x_2^2 \quad \text{subject to } x_1^2 + x_2^2 - 1 \geq 0, \tag{1.28}$$

in which we seek to minimise a non-convex function over the exterior of the unit circle. The objective function is not bounded below in the feasible region. Therefore, no global solution exists, but it may still be possible to identify a strict local solution on the boundary of the constraint. We search for such solution by using the KKT conditions (1.20)-(1.23) and the second order conditions of Theorem 1.3.

By defining the Lagrangian for (1.28) in the usual way, it is easy to verify that

$$\nabla_x L(\mathbf{x}, \lambda) = \begin{bmatrix} -0.2(x_1 - 4) - 2\lambda_1 x_1 \\ 2x_2 - 2\lambda_1 x_2 \end{bmatrix} \tag{1.29}$$

$$\nabla^2_{xx} L(\mathbf{x}, \lambda) = \begin{bmatrix} -0.2 - 2\lambda_1 & 0 \\ 0 & 2 - 2\lambda_1 \end{bmatrix} \tag{1.30}$$

The point $\mathbf{x}^* = (1, 0)$ satisfies the KKT conditions with $\lambda_1 = 0.3$ and the active set $A(\mathbf{x}^*) = \{1\}$. To check that the second order sufficient conditions are satisfied at this point, we note that

$$\nabla c_1(\mathbf{x}) = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

so the space F defined by (1.24) is simply

$$F(\lambda^*) = \{\mathbf{w}|w_1 = 0\} = \left\{(0, w_2)^T | w_2 \in \mathbb{R}\right\}.$$

Now, by substituting $\mathbf{x}^*$ and λ^* into (1.30), we have for any $\mathbf{w} \in F$ with $\mathbf{w} \neq \mathbf{0}$ that

$$\mathbf{w}^T \nabla^2_{xx} L(\mathbf{x}^*, \lambda^*)\mathbf{w} = \begin{bmatrix} 0 & w_2 \end{bmatrix} \begin{bmatrix} -0.4 & 0 \\ 0 & 1.4 \end{bmatrix} \begin{bmatrix} 0 \\ w_2 \end{bmatrix} = 1.4 w_2^2 > 0.$$

Hence, the second order sufficient conditions are satisfied, and we conclude from Theorem 1.3 that $\mathbf{x}^* = (1, 0)$ is a strict local minimizer for problem (1.28).

PROBLEM SET I

1. Consider the following inequality constrained optimization problem

$$\begin{aligned} \min_x \; & \left(x_1 - \frac{3}{2}\right)^2 + \left(x_2 - \frac{1}{2}\right)^4 \\ \text{subject to: } & 1 - x_1 - x_2 \geq 0 \\ & x_1 - x_2 - 1 \leq 0 \\ & x_2 - x_1 - 1 \leq 0 \\ & 1 + x_1 + x_2 \geq 0 \end{aligned}$$

   Answer the following questions:

   (a) Graphically, find the optimal solution of the problem.

   (b) Verify whether the Linear Independent Constrained qualification is satisfied.

   (c) Find λ^* such that the Karush-Kuhn-Tucker conditions are satisfied.

2. Consider the feasible region defined by

$$\begin{aligned} 4x_2^2 + x_1^2 &\leq 4 \\ (x_1 - 2)^2 + x_2^2 &\leq 5 \\ x_1 \geq 0, \quad x_2 &\geq 0 \end{aligned}$$

   Show that the LICQ is not satisfied at the the feasible point $\mathbf{x}^* = [0, 1]^T$.

3. Use the First and Second order optimality conditions to solve the following constrained optimization problems.

   (a)

$$\begin{aligned} \min \quad & x_1 \\ \text{s.t: } & x_2 \geq 0, \\ & (x_1 - 1)^2 + x_2^2 - 1 \leq 0. \end{aligned}$$

   (b)

$$\begin{aligned} \min \quad & -0.1(x_1 - 4)^2 + x_2^2 \\ \text{s.t: } & 1 - x_2^2 - x_2^2 \leq 0. \end{aligned}$$

4. Consider the problem of finding a point on the parabola $y = \frac{1}{5}(x - 1)^2$ that is closest to the point $(x, y) = (1, 2)$ in the Euclidean norm sense.

(a) Formulate this problem as constrained optimization problem.

(b) Find the KKT points for this problem. Is the LICQ satisfied?

(c) Which points are solutions?

(d) By directly substituting the constraint into the objective function and eliminating the variable x, we obtain an unconstrained optimization problem. Show that the solution of this problem cannot be the solution of the original problem.

5. Find the maxima of $f(x) = x_1 x_2$ over a unit disk defined by the inequality constraint $x_1^2 + x_2^2 \leq 1$.

6. Solve the following constrained optimization problem

$$\min x_1 + x_2 \quad \text{such that} \quad 2 - x_1^2 - x_2^2 \geq 0, \quad x_2 \geq 0.$$