Chapter4: Nonlinear Constrained Optimization (Equality Constrained Optimization)

Mr. OL Sela

Github: https://github.com/OLSela12

Mail: selaol168@gmail.com

March 14, 2025

Outline

Outline

Problem

In this part we discuss methods for solving a class of nonlinear constrained optimization problems that can be formulated as

Problem

minimize

$$f(x)$$

subject to

$$\begin{aligned} h_i(x) &= 0, \quad i = 1, 2, ..., m, \\ g_j(x) &\leq 0, \quad j = 1, 2, ..., p, \end{aligned}$$

where $x \in \mathbb{R}^n, f : \mathbb{R}^n \to \mathbb{R}, h_i : \mathbb{R}^n \to \mathbb{R}, g_j : \mathbb{R}^n \to \mathbb{R}$ and $m + p \leq n$

Problem

In vector notation, the problem above can be represented in the following standard form:
minimize

$$f(x)$$

subject to

$$h(x) = 0,$$
$$g(x) \leq 0,$$

where $x \in \mathbb{R}^n, f : \mathbb{R}^n \to \mathbb{R}, h : \mathbb{R}^n \to \mathbb{R}^m, g : \mathbb{R}^n \to \mathbb{R}^p$ and $m + p \leq n$

Definition 1

Any point satisfying the constraints is called a feasible point. The set of all feasible points

$$\{x \in \mathbb{R}^n : h(x) = 0, g(x) \leq 0\}$$

is called a feasible set

Optimization problems of the above form are not new to us. Indeed, linear programming problems of the form

LPP

minimize

$$z = c^T x$$

subject to

$$\begin{aligned} Ax &= b, \\ x &\geq 0 \end{aligned}$$

which we studied in chapter $2, 3$.

Noted

$$\text{maximize } f(x) = \text{minimize } (-f(x)).$$

Example

Example 2

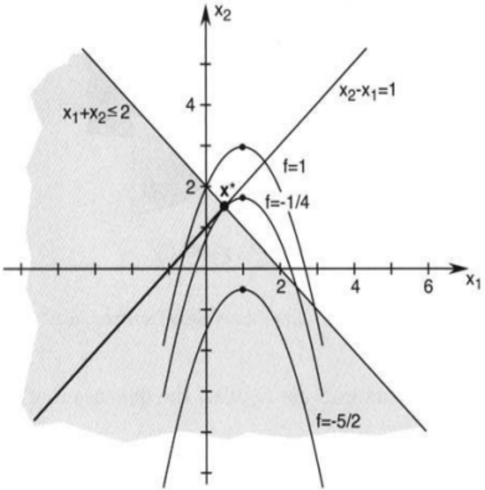
Consider the following optimization problem:
minimize

$$(x_1 - 1)^2 + x_2 - 2$$

subject to

$$x_2 - x_1 = 1,$$
$$x_1 + x_2 \leq 2.$$

In this case, the minimizer lies on the level set with $f = -1/4$. The minimizer of the objective function is $x^\star = [1/2, 3/2]^T$.


x2
x2-x1=1
4
x1+x2≤2
f=1
2
x*
f=-1/4
2
4
6
x1
f=-5/2

Outline

Equality Constrained Optimization

Problem

minimize

$$f(x)$$

subject to

$$h(x) = 0,$$

where $x \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$, $h_i : \mathbb{R}^n \to \mathbb{R}^m$, $h = [h_1, h_2, ..., h_m]^T$, and $m \leq n$

We assume that the function h is continuously differentiable, that is, $h \in \mathcal{C}^1$.

Example

Example 3

$$\text{minimize } 2x_1^2 + x_2^2$$
$$\text{subject to } x_1 + x_2 = 1$$

- Let us first consider the unconstrained case.
- Differentiate with respect to x_1 and x_2.

$$\frac{\partial f}{\partial x_1} = 4x_1$$

$$\frac{\partial f}{\partial x_2} = 2x_2$$

- These yield the solution $x_1 = x_2 = 0$
- Does not satisfy the constrain.

Definition 4

A point $x^\star$ satisfying the constraints

$$h_1(x^\star) = 0, \dots, h_m(x^\star) = 0$$

is said to be a regular point of the constraints if the gradient vectors

$$\nabla h_1(x^\star), \dots, \nabla h_m(x^\star)$$

are linearly independent.

Let $Dh(x^\star)$ be the Jacobian matrix of $h = [h_1, ...h_m]^T$ at $x^\star$, given by

$$Dh\left(x^\star\right) = \begin{bmatrix} Dh_1\left(x^\star\right) \\ \vdots \\ Dh_m\left(x^\star\right) \end{bmatrix} = \begin{bmatrix} \nabla h_1(x^\star)^T \\ \vdots \\ \nabla h_m(x^\star)^T \end{bmatrix}$$

Then, $x^\star$ is regular if and only if rank $Dh(x^\star) = m$ (i.e., the Jacobian matrix is of full rank)

Example

Example 5

For example, let

$$h_1(x) = x_1, h_2(x) = x_2 - x_3^2$$

In this case,

$$\nabla h_1(x) = [1, 0, 0]^T, \nabla h_2(x) = [0, 1, -2x_3]^T$$

Question , $\nabla h_1(x), \nabla h_2(x)$ are linearly independent??

Outline

Lagrange Condition

We now generalize Lagrange's theorem.

Lagrange's Theorem.

Let $x^\star$ be a local minimizer (maximizer) of

$$f : \mathbb{R}^n \to \mathbb{R}$$

,subject to $h(x) = 0$, where $h : \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$.
Assume that $x^\star$ is a regular point. Then,there exists $\lambda^\star \in \mathbb{R}^m$ such that

$$Df(x^\star) + {\lambda^\star}^T Dh(x^\star) = 0^T \tag{1}$$

Proof: see in book

- It is convenient to introduce the Lagrangian function $L : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, given by,

$$L(x, \lambda) = f(x) + \lambda^T h(x) \tag{2}$$

- The Lagrange condition for a local minimizer $x^\star$ can be represented using the Lagrangian function as

$$Dh(x^\star, \lambda^\star) = 0^T \tag{3}$$

for some $\lambda^\star$, where the derivative operation D is with respect to the entire argument $[x^T, \lambda^T]^T$

In other words, the necessary condition in Lagrange's theorem is equivalent to the first-order necessary condition for unconstrained optimization applied to the Lagrangian function

- To see the above, denote the derivative of L with respect to x as $D_x L$ and the derivative of L with respect to λ as $D_\lambda L$. Then

$$DL(x,\lambda) = [D_x L(x,\lambda), D_\lambda L(x,\lambda)]$$

  Note that $DL(x,\lambda) = Df(x) + \lambda^T Dh(x)$ and $D_\lambda L(x,\lambda) = h(x)^T$.

- Therefore, Lagrange's theorem for a local minimizer $x^\star$ can be stated as

$$D_x L(x^\star, \lambda^\star) = 0^T,$$
$$D_\lambda L(x^\star, \lambda^\star) = 0^T$$

- The Lagrange condition is used to find possible extremizers.This entails solving the equations

$$D_x L(x, \lambda) = 0^T,$$
$$D_\lambda L(x, \lambda) = 0^T$$

Outline

Example

Example 6

Find the minimum value of

$$f(x, y, z) = 2x^2 + y^2 + z^2$$

subject to the constraint

$$2x - 3y - 4z = 49$$

.

Example 7

Find the maximum value of

$$f(x, y) = 4xy$$

,where $x > 0, y > 0$ subject to the constraint

$$\frac{x^2}{9} + \frac{y^2}{16} = 1$$

Example

Example 8

Given a fixed area of cardboard, we wish to construct a closed cardboard box with maximum volume. We can formulate and solve this problem using the Lagrange condition. Denote the dimensions of the box with maximum volume by x_1, x_2, and x_3, and let the given fixed area of cardboard be A. The problem can then be formulated as

$$\begin{aligned} \text{maximize } & x_1 x_2 x_3 \\ \text{subject to } & x_1 x_2 + x_2 x_3 + x_3 x_1 = \frac{A}{2} \end{aligned}$$

Example

Example 9

Consider the problem of extremizing the objective function

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

on the ellipse

$$\{[x_1, x_2]^T : h(x) = x_1^2 + 2x_2^2 - 1 = 0\}$$

Example 10

A consumer's preferences are represented by the utility function

$$u(x_1, x_2) = 2lnx_1 + lnx_2 \tag{4}$$

If the budget constraint is $p_1x_1 + p_2x_2 = M$,
Determine the demand functions, that is, the optimal values $x_1^\star$ and $x_2^\star$; in terms ofp_1, p_2 and M.

Outline

Second-Order Conditions

We assume that $f : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^m$ are twice continuously differentiable: $f, h \in \mathcal{C}^2$.

- Let
$$L(x, \lambda) = f(x) + \lambda^T h(x) = f(x) + \lambda_1 h_1(x) + ... + \lambda_m h_m(x).$$
be the Lagrangian function
- Let
$$L'(x, \lambda) = F(x) + \lambda_1 H_1(x) + ... + \lambda_m H_m(x),$$
where $F(x)$ is the Hessian matrix of $L(x, \lambda)$ at x and $H_k(x)$ is the Hessian matrix of h_k at $x, k = 1, 2, ..., m$, given by
$$H_k(x) = \begin{bmatrix} \frac{\partial^2 h_k}{\partial x_1^2}(x) & \cdots & \frac{\partial^2 h_k}{\partial x_n \partial x_1}(x) \\ \vdots & & \\ \frac{\partial^2 h_k}{\partial x_1 \partial x_n}(x) & \cdots & \frac{\partial^2 h_k}{\partial x_n^2}(x) \end{bmatrix}$$

We introduce the notation $[\lambda H(x)]$;

$$[\lambda H(x)] = \lambda_1 H_1(x) + \ldots + \lambda_m H_m$$

Using the notation above, we can write

$$L'(x, \lambda) = F(x) + [\lambda H(x)]$$

Second-Order Necessary Conditions.

Theorem 11

Let $x^\star$ be a local minimizer (maximizer) of

$$f : \mathbb{R}^n \to \mathbb{R}$$

,subject to $h(x) = 0$, where $h : \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$ and $f, g \in \mathcal{C}^2$. Assume that $x^\star$ is a regular point. Then, there exists $\lambda^\star \in \mathbb{R}^m$ such that

- $Df(x^\star) + {\lambda^\star}^T Dh(x^\star) = 0^T$
- *For all $y \in T(x^\star)$, we have $y^T L'(x^\star, \lambda^\star) y \geq 0$.*

Noted: $T(x)$ is a Tangent Space.

Second-Order Sufficient Conditions.

Theorem 12

Suppose that $f, h \in \mathcal{C}^2$, *and there exists a point* $x^\star \in \mathbb{R}^n$ *and* $\lambda^\star \in \mathbb{R}^m$ *such that*

- $Df(x^\star) + \lambda^{\star T} Dh(x^\star) = 0^T$
- *For all* $y \in T(x^\star)$,$y \neq 0$, *we have* $y^T L'(x^\star, \lambda^\star) y > 0$.

Then, $x^\star$ *is a strict local minimizer of* f *subject to* $h(x) = 0$.

Example

Example 13

Consider

min:

$$f(x) = \frac{1}{2}x^T Ax$$

subject to :

$$Ax = b$$

where $Q > 0$ (Q is positive definite on $\mathbb{R}^n$)

Example

Outline

Thank for your attention.