Lecture 26

Constrained Nonlinear Problems

Necessary KKT Optimality Conditions

November 18, 2009

Outline

- Necessary Optimality Conditions for Constrained Problems
  - Karush-Kuhn-Tucker* (KKT) optimality conditions
    - □ Equality constrained problems
    - □ Inequality and equality constrained problems
  - Convex Inequality Constrained Problems
    - □ Sufficient optimality conditions

- The material is in Chapter 18 of the book
  - Section 18.1.1
  - Lagrangian Method in Section 18.2 (see 18.2.1 and 18.2.2)

*William Karush develop these conditions in 1939 as a part of his M.S. thesis at the University of Chicago; the same were developed independently later in 1951 by Harold W. Kuhn and Albert W. Tucker.

Constrained Optimization

Let $f : \mathbb{R}^n \to \mathbb{R}$, $g_j : \mathbb{R}^n \to \mathbb{R}$ for $j = 1, \ldots, m$, and $h_\ell : \mathbb{R}^n \to \mathbb{R}$ for $\ell = 1, \ldots, r$

Unconstrained NLP problem

$$\begin{array}{rl} \textbf{minimize} & f(x) \\ \textbf{no restrictions} & x \in \mathbb{R}^n \end{array}$$

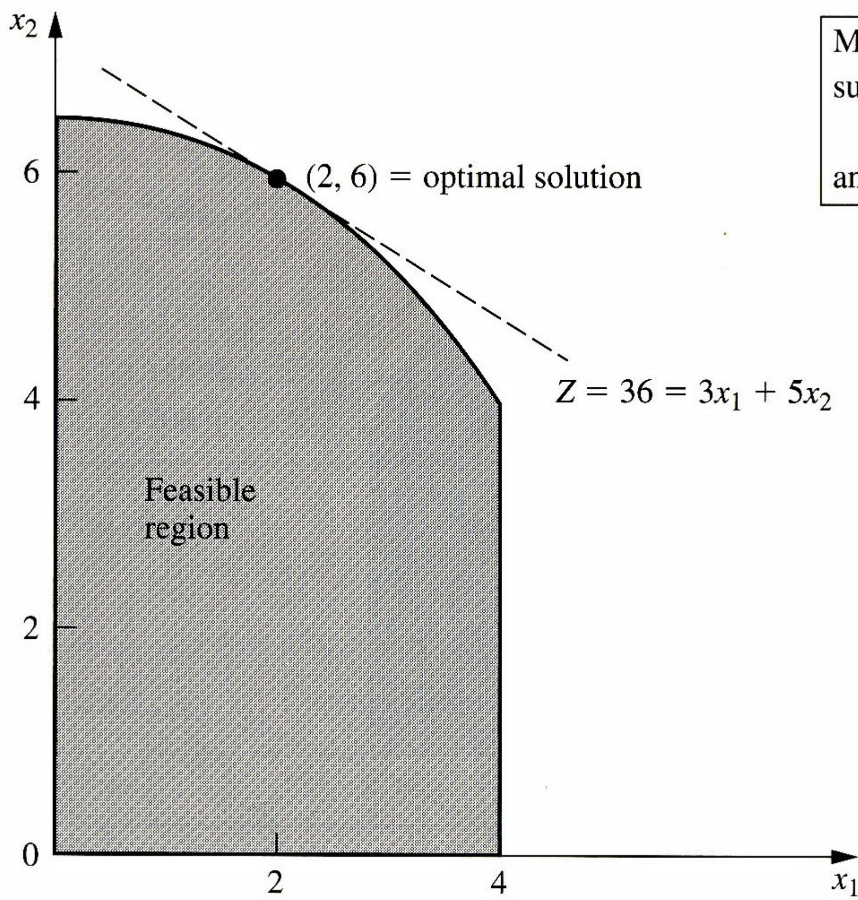
Constrained NLP problem

$$\begin{array}{rl} \textbf{minimize} & f(x) \\ \textbf{subject to} & g_j(x) \le 0, \quad j = 1, \ldots, m \\ & h_\ell(x) = 0, \quad \ell = 1, \ldots, r \end{array}$$

For constrained problem, we say that x **if feasible** if it satisfies all the constraints of the problem, i.e.,

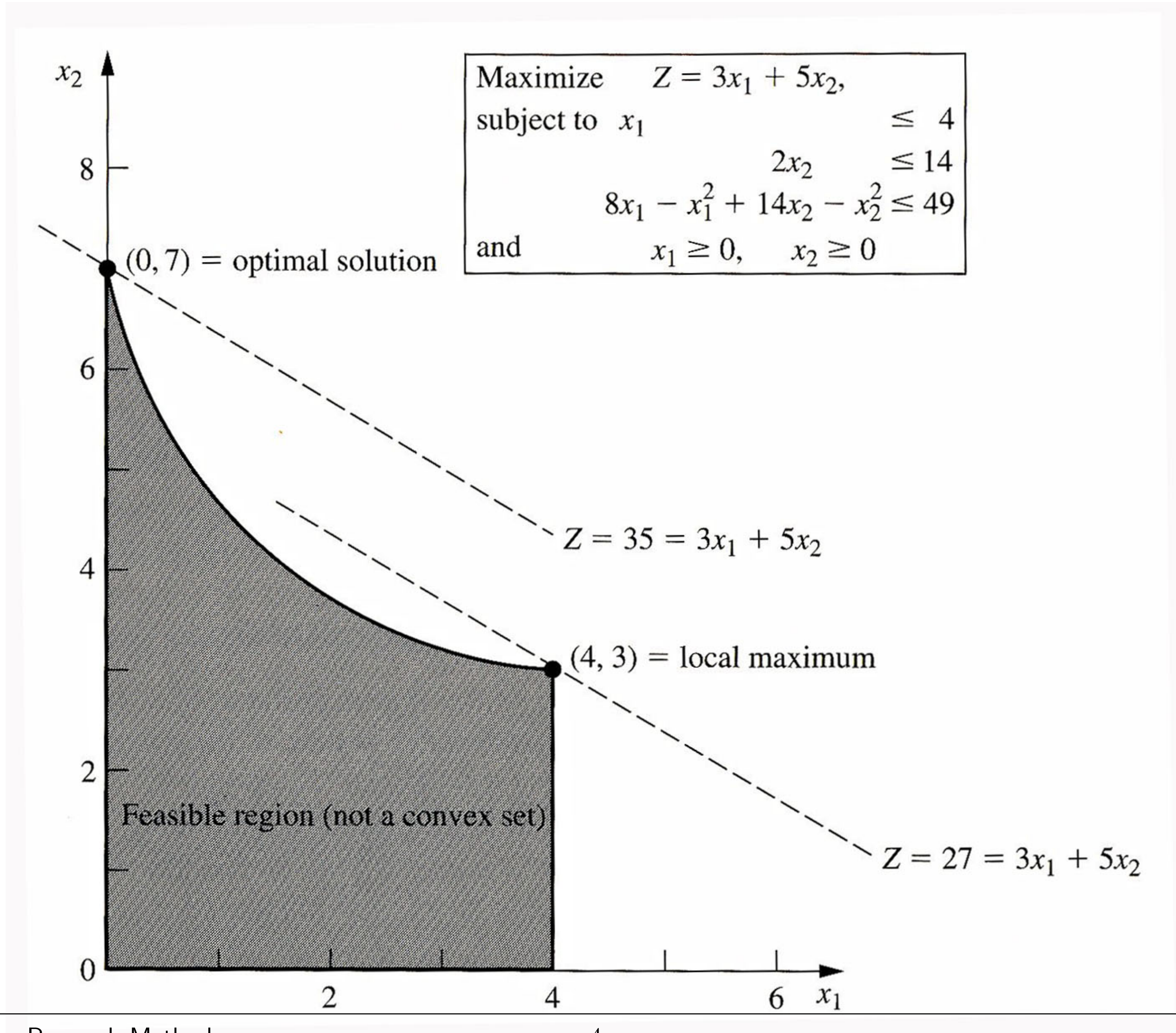
$$g_j(x) \le 0, \quad j = 1, \ldots, m, \qquad h_\ell(x) = 0, \quad \ell = 1, \ldots, r$$

Graphical Illustrations



Maximize	$Z = 3x_1 + 5x_2,$
subject to	$x_1 \leq 4$
	$9x_1^2 + 5x_2^2 \leq 216$
and	$x_1 \geq 0, \quad x_2 \geq 0$

The set of all feasible points constitute the feasible region of the problem


Maximize $Z = 3x_1 + 5x_2$,
subject to $x_1 \leq 4$
$2x_2 \leq 14$
$8x_1 - x_1^2 + 14x_2 - x_2^2 \leq 49$
and $x_1 \geq 0, \quad x_2 \geq 0$
(0, 7) = optimal solution
$Z = 35 = 3x_1 + 5x_2$
(4, 3) = local maximum
$Z = 27 = 3x_1 + 5x_2$
Feasible region (not a convex set)
x_1
x_2
0
2
4
6
8

EXAMPLE of Constrained NLP: Portfolio Selection with Risky Securities

$$
\begin{aligned}
\text{minimize} \quad & V(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j \\
\text{subject to} \quad & \sum_{j=1}^{n} p_j x_j \le B \\
& \sum_{j=1}^{n} \mu_j x_j \ge L \\
& x_j \ge 0 \quad \text{for all } j
\end{aligned}
$$

This is a constrained NLP problem. In fact it is **linearly** constrained.

Minimization vs Maximization

We will focus on minimization type problems, since maximization problems can be transformed to minimization problems

The optimal solutions (if any exist) of the problem

$$\begin{array}{rl} \textbf{maximize} & f(x) \\ \textbf{subject to} & g_j(x) \le 0, \quad j = 1, \ldots, m \\ & h_\ell(x) = 0, \quad \ell = 1, \ldots, r \end{array} \tag{1}$$

coincide with the optimal solutions of the following problem

$$\begin{array}{rl} \textbf{minimize} & -f(x) \\ \textbf{subject to} & g_j(x) \le 0, \quad j = 1, \ldots, m \\ & h_\ell(x) = 0, \quad \ell = 1, \ldots, r \end{array} \tag{2}$$

NOTE: The optimal values of the problems (1) and (2) are obviously **not the same** (differ in sign)

Furthermore, the local maxima of problem (1) coincide with the local minima of problem (2).

Necessary Optimality Conditions

Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable over $\mathbb{R}^n$

Let each $g_j : \mathbb{R}^n \to \mathbb{R}$ and each $h_\ell : \mathbb{R}^n \to \mathbb{R}$.

Unconstrained NLP problem

$$\begin{array}{rl} \textbf{minimize} & f(x) \\ \textbf{no restrictions} & x \in \mathbb{R}^n \end{array}$$

Necessary Optimality Condition: If x^* is an optimal solution for the problem, then x^* satisfies

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

Constrained NLP problem

$$\begin{array}{rl} \textbf{minimize} & f(x) \\ \textbf{subject to} & g_j(x) \le 0, \quad j = 1, \ldots, m \\ & h_\ell(x) = 0, \quad \ell = 1, \ldots, r \end{array}$$

Necessary Optimality Condition:

???

Necessary Optimality Conditions: Equality Constrained Problems

Consider the equality constrained problem:

$$\begin{array}{rl} \textbf{minimize} & f(x) \\ \textbf{subject to} & h_\ell(x) = 0, \quad \ell = 1, \ldots, r \end{array} \tag{3}$$

The functions f and all h_ℓ are continuously differentiable over $\mathbb{R}^n$. The optimality conditions are specified through the use of Lagrangian Function:

- Introduce a multiplier (shadow price) λ_ℓ for each constraint $h_\ell(x) = 0$.

- Lagrangian Function is given by

$$L(x, \lambda) = f(x) + \sum_{\ell=1}^{r} \lambda_\ell h_\ell(x)$$

where $\lambda = (\lambda_1, \ldots, \lambda_r)$

Necessary Optimality Condition:
Assuming some regularity conditions for problem (3), if x^* is an optimal solution of the problem, then there exists a Lagrange multiplier (optimal shadow price) $\lambda^* = (\lambda_1^*, \ldots, \lambda_r^*)$ such that

$$\left\{ \begin{array}{l} \nabla_x L(x^*, \lambda^*) = 0 \\ \\ \nabla_\lambda L(x^*, \lambda^*) = 0 \end{array} \right\} \quad \begin{array}{c} \Longleftrightarrow \\ \\ \Longleftrightarrow \end{array} \quad \left\{ \begin{array}{l} \nabla f(x^*) + \sum_{\ell=1}^r \lambda_\ell^* \nabla h_\ell(x^*) = 0 \\ \\ h_\ell(x^*) = 0 \text{ for } \ell = 1, \ldots, r \end{array} \right\}$$

λ^* is the optimal shadow price associated with the solution x^*
This condition is known as KKT condition
IMPORTANT: The KKT condition can be satisfied at a local minimum, a global minimum (solution of the problem) as well as at a saddle point.
Question:
We want to determine the optimal solutions of the problem (global minima of the constrained problem)? How can we use the KKT condition?

Answer:

We can set up a system of linear equations using the KKT condition:

$$\nabla f(x) + \sum_{\ell=1}^{r} \lambda_\ell \nabla h_\ell(x) = 0$$
$$h_\ell(x) = 0 \text{ for } \ell = 1, \ldots, r$$

We have $n + r$ unknown variables (x of size n and λ of size r) and $n + r$ equations

We can solve the system and find the points that satisfy the equations (KKT condition)

These points are known as *stationary points* (or *KKT points*)

The optimal solutions (if any) are among these points.

NOTE:

We need additional information to characterize the stationary points as global or local minimum or other by using the second order information for the objective $f(x)$ at the KKT points (see Lecture 25).

Example

Determine all the stationary points of the following constrained problem

$$
\begin{array}{rl}
\text{minimize} & f(x) = x_1^2 + x_2^2 + x_3^2 \\
\text{subject to} & x_1 + x_2 + 3x_3 - 2 = 0 \\
 & 5x_1 + 2x_2 + x_3 - 5 = 0
\end{array}
$$

We construct the Lagrangian function for the problem:

$$L(x, \lambda) = x_1^2 + x_2^2 + x_3^2 + \lambda_1(x_1 + x_2 + 3x_3 - 2) + \lambda_2(5x_1 + 2x_2 + x_3 - 5)$$

We set up the equations:

$$\frac{\partial L(x,\lambda)}{\partial x_1} = 2x_1 + \lambda_1 + 5\lambda_2 = 0$$
$$\frac{\partial L(x,\lambda)}{\partial x_2} = 2x_2 + \lambda_1 + 2\lambda_2 = 0$$
$$\frac{\partial L(x,\lambda)}{\partial x_3} = 2x_3 + 3\lambda_1 + \lambda_2 = 0$$
$$\frac{\partial L(x,\lambda)}{\partial \lambda_1} = x_1 + x_2 + 3x_3 - 2 = 0$$
$$\frac{\partial L(x,\lambda)}{\partial \lambda_2} = 5x_1 + 2x_2 + x_3 - 5 = 0$$

We solve them: $x = [0.8043\ \ 0.3478\ \ 0.2826]^T$ and $\lambda = -[0.0870\ \ 0.3044]^T$

So we have only one KKT point, namely $x = [0.8043\ \ 0.3478\ \ 0.2826]^T$
- but we still do not know if this is optimal or not

We can check the Hessian of f. We will find that the Hessian is given by a diagonal matrix with diagonal entries equal to 2.

The Hessian is positive definite, therefore the point x^* is a global minimum.

Necessary Optimality Conditions: Inequality and Equality Constrained Problems

Consider the following constrained problem:

$$\begin{array}{rl} \textbf{minimize} & f(x) \\ \textbf{subject to} & g_j(x) \le 0, \quad j = 1, \ldots, m \\ & h_\ell(x) = 0, \quad \ell = 1, \ldots, r \end{array} \tag{4}$$

The functions f and all g_j and h_ℓ are continuously differentiable over $\mathbb{R}^n$. The optimality conditions are specified through the use of Lagrangian Function:

- Introduce a multiplier (shadow price) per constraint:
  $\mu_j \ge 0$ for each constraint $g_j(x) \le 0$ and
  λ_ℓ for each constraint $h_\ell(x) = 0$

- Lagrangian Function is given by

$$L(x, \mu, \lambda) = f(x) + \sum_{j=1}^{m} \mu_j g_j(x) + \sum_{\ell=1}^{r} \lambda_\ell h_\ell(x)$$

where $\mu = (\mu_1, \ldots, \mu_m)$ and $\lambda = (\lambda_1, \ldots, \lambda_r)$

Necessary Optimality Condition:

Assuming some regularity conditions for problem (4), if x^* is an optimal solution of the problem, then there exist Lagrange multipliers (optimal shadow prices) $\mu^* = (\mu_1^*, \ldots, \mu_m^*) \geq 0$ and $\lambda^* = (\lambda_1^*, \ldots, \lambda_r^*)$ such that

$$\nabla f(x^*) + \sum_{j=1}^{m} \mu_j^* \nabla g_j(x^*) + \sum_{\ell=1}^{r} \lambda_\ell^* \nabla h_\ell(x^*) = 0$$

$$g_j(x^*) \leq 0 \quad \text{for all } j = 1, \ldots, m$$

$$h_\ell(x^*) = 0 \text{ for all } \ell = 1, \ldots, r$$

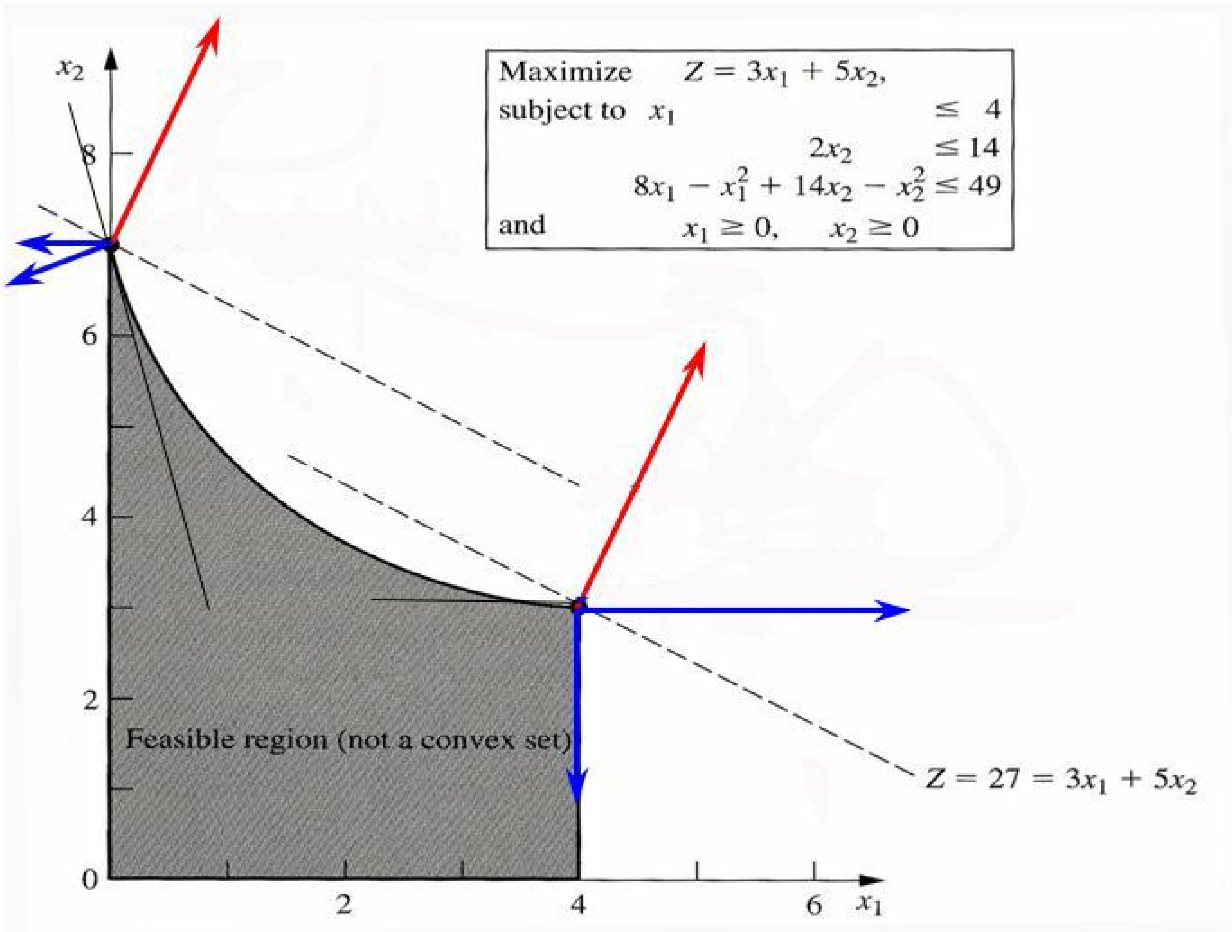
$$\mu_j^* \geq 0 \quad \text{for all } j = 1, \ldots, m$$

$$\mu_j^* g_j(x^*) = 0 \quad \text{for all } j = 1, \ldots, m$$

μ^* and λ^* is the optimal shadow price associated with the solution x^*

This is the KKT condition

Graphical Illustration of the KKT Condition



IMPORTANT: The KKT condition can be satisfied at a local minimum, a global minimum (solution of the problem) as well as at a saddle point.

We can use the KKT condition to characterize all the stationary points of the problem, and then perform some additional testing to determine the optimal solutions of the problem (global minima of the constrained problem).

Determining KKT points: we set up a KKT system for problem (4):

$$\nabla f(x) + \sum_{j=1}^{m} \mu_j \nabla g_j(x) + \sum_{\ell=1}^{r} \lambda_\ell \nabla h_\ell(x) = 0$$

$$g_j(x) \le 0 \quad \text{for all } j = 1, \ldots, m$$

$$h_\ell(x) = 0 \text{ for all } \ell = 1, \ldots, r$$

$$\mu_j \ge 0 \quad \text{for all } j = 1, \ldots, m$$

$$\mu_j g_j(x) = 0 \quad \text{for all } j = 1, \ldots, m$$ **complementarity slackness**

We may solve this (nonlinear) system in unknown variables x, μ and λ, and find all the points satisfying the KKT condition

These points are *stationary points* (or *KKT points*) of the problem

The optimal points are among the KKT points

We need additional information to characterize the KKT points as global or local minimum or other - we use the second order information, as discussed in Lecture 25.

Sufficient Optimality Conditions: Convex Inequality Constrained Problems

Consider the following constrained problem:

$$\begin{array}{rl} \textbf{minimize} & f(x) \\ \textbf{subject to} & g_j(x) \leq 0, \quad j = 1, \ldots, m \end{array} \tag{5}$$

Functions f and all g_j are **convex** and continuously differentiable over $\mathbb{R}^n$.

We say that the problem is **convex** For this convex problem, the KKT conditions are also **sufficient** for optimality.

Sufficient Optimality Condition:
Assuming some additional regularity conditions for **convex** problem (5), x^* is an optimal solution of the problem, if and only if there exists a Lagrange multiplier (optimal shadow price) $\mu^* = (\mu_1^*, \ldots, \mu_m^*)$ such that

$$\nabla f(x^*) + \textstyle\sum_{j=1}^m \mu_j^* \nabla g_j(x^*)$$

$$g_j(x^*) \leq 0 \quad \text{for all } j = 1, \ldots, m$$

$$\mu_j^* \geq 0 \quad \text{for all } j = 1, \ldots, m$$

$$\mu_j^* g_j(x^*) = 0 \quad \text{for all } j = 1, \ldots, m$$

μ^* is the optimal shadow price associated with the solution x^*

We can use the KKT condition to fully recover optimal solutions of the convex problem (global minima of the constrained problem).

Looking for the solutions: we set up a KKT system for problem (5),

$$\nabla f(x) + \sum_{j=1}^{m} \mu_j \nabla g_j(x) + \sum_{\ell=1}^{r} \lambda_\ell \nabla h_\ell(x^*) = 0$$

$$g_j(x) \le 0 \quad \text{for all } j = 1, \ldots, m$$

$$h_\ell(x) = 0 \text{ for all } \ell = 1, \ldots, r$$

$$\mu_j \ge 0 \quad \text{for all } j = 1, \ldots, m$$

$$\mu_j g_j(x) = 0 \quad \text{for all } j = 1, \ldots, m$$ **complementarity slackness**

We solve this system in unknown variables x and μ, and find all the KKT points

IMPORTANT By the assumed convexity of problem (5), all the KKT points are global minima of the problem (optimal solutions).

EXAMPLE

Consider the following problem

$$\begin{aligned} \text{minimize} \quad & f(x) = -\ln(x_1 + 1) - x_2 \\ \text{subject to} \quad & 2x_1 + x_2 \leq 3 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{aligned}$$