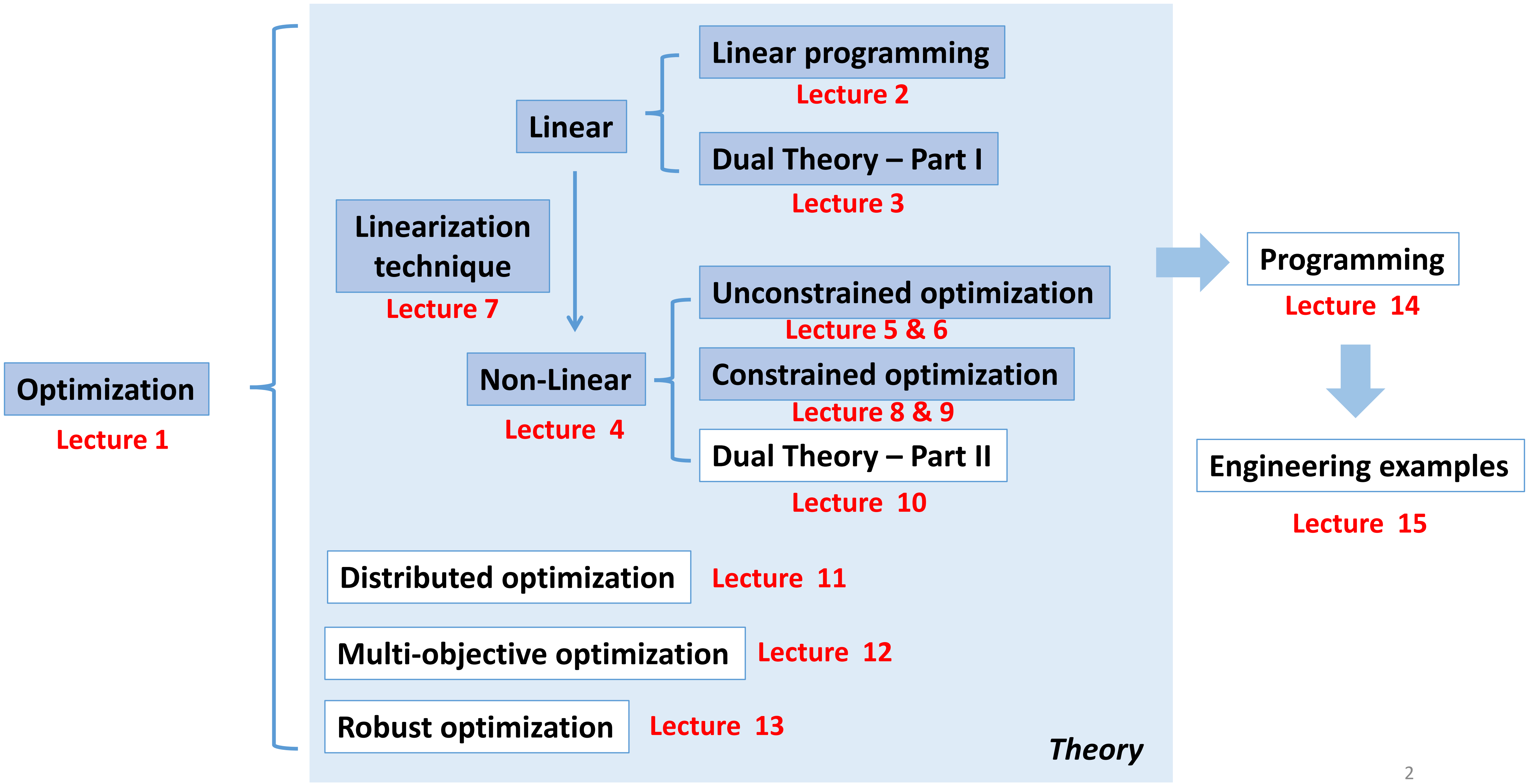
MAEG4070 Engineering Optimization

Lecture 8 Constrained Optimization
Lagrange Multiplier

Yue Chen
MAE, CUHK
email: yuechen@mae.cuhk.edu.hk
Oct 17, 2022

Content of this course (tentative)



Introduction

Nonlinear problems with constraints are quite common in practice.

Let's look at an example:
A company produces product A and B, whose selling prices are 30 and 450, respectively. It takes 0.5 hours to sell product A and (2+0.3x_2) hours to sell product B. The operational time for the company is 800 hours. How to decide on the production plan to maximize the profit?

Solution: suppose quantities for A and B are x_1 and x_2, respectively.

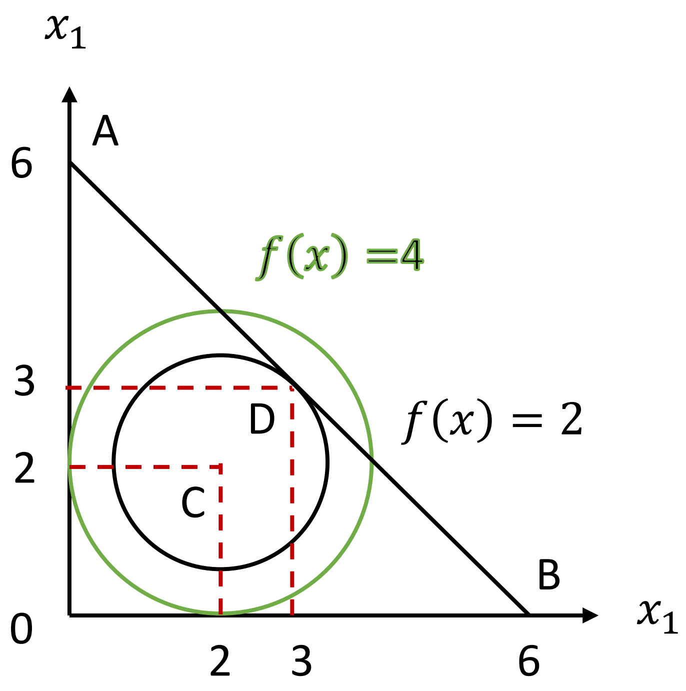
$$\begin{aligned} \max_{x_1, x_2} \ & 30x_1 + 450x_2 \\ \text{s.t.} \ & 0.5x_1 + 2x_2 + 0.3x_2^2 \leq 800 \\ & x_1 \geq 0, x_2 \geq 0 \end{aligned}$$

Geometrical method

Solve this optimization problem:

$$\min_{x_1,x_2} \; f(x) = (x_1 - 2)^2 + (x_2 - 2)^2$$
$$\text{s.t. } h(x) = x_1 + x_2 - 6 = 0$$

- The constraint is line AB
- We want to minimize the distance from a point at line AB to the point (2,2)
- Draw a circle centered at (2,2), increase its radius until the circle is tangent to the line

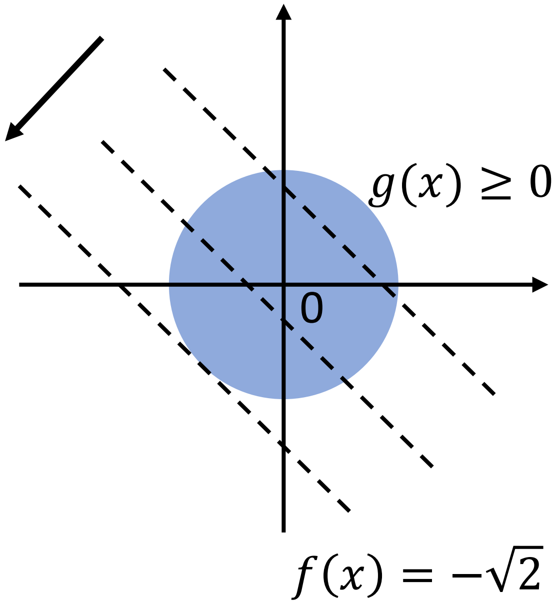


$$f(x^*) = 2, x^* = (3,3)$$

Geometrical method

Solve this optimization problem:

$$\min_{x_1,x_2} \; f(x) = x_1 + x_2$$
$$\text{s.t. } g(x) = 1 - x_1^2 - x_2^2 \geq 0$$



- The constraint represents all points within the unit circle centered at (0,0)
- We move the line with a slope of -1 until it is tangent to the circle

$$f(x^*) = -\sqrt{2}, x^* = (-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$$

Constrained Optimization with Equality

$$\min_x \; f(x)$$
$$\text{s.t. } h_i(x) = 0, i = 1, ..., m$$

where $f : \mathbb{R}^n \to R$, $h_i : \mathbb{R}^n \to R, \forall i = 1, ..., m.$

- We suppose both f and $h_i, \forall i$ are continuously differentiable functions
- Note that the theory also applies to case where f and $h_i, \forall i$ are continuously differentiable in a neighborhood of a local minimum.

Constrained Optimization with Equality

$\sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) = 0$
If and only if $\lambda_i^* = 0, \forall i$

Lagrange Multiplier Theorem

Let x^* be a local minimum and a regular point ($\nabla h_i(x^*), \forall i$ are linearly independent)
Then, there exists unique scalars $\lambda_1^*, \ldots, \lambda_m^*$ such that

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) = 0$$

Unconstrained case:
$\nabla f(x^*) = 0$

If in addition f and h are twice continuous differentiable

$$y^T \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(x^*) \right) y \geq 0 \quad \text{s.t. } \nabla h(x^*)^T y = 0$$

Unconstrained case: $y^T \nabla^2 f(x^*) y \geq 0, \forall y$

Constrained Optimization with Equality

Exercise on linearly independent

Consider two vectors $v_1 = (1,1)^T$ and $v_2 = (-3,2)^T$, are they linearly independent?

$$\lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} -3 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Therefore,

$$\begin{pmatrix} 1 & -3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 & -3 \\ 1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Vectors $v_1 = (1,1)^T$ and $v_2 = (-3,2)^T$ are linearly independent.

Constrained Optimization with Equality

Exercise on linearly independent

Consider three vectors $v_1 = (1,1)^T$, $v_2 = (-3,2)^T$, and $v_3 = (2,4)^T$

$$\lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} -3 \\ 2 \end{pmatrix} + \lambda_3 \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Therefore,

$$\begin{pmatrix} 1 & -3 & 2 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

We have $(\lambda_1, \lambda_2, \lambda_3) = (-16, -2, 5)$ satisfies the equation.

Vectors $v_1 = (1,1)^T$, $v_2 = (-3,2)^T$, and $v_3 = (2,4)^T$ are not linearly independent.

Constrained Optimization with Equality

Lagrange Multiplier Theorem (Necessary condition)

Define the Lagrangian function

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x)$$

Then, if x^* is a local minimum which is regular,
the Lagrange multiplier conditions are written

$$\nabla_x L(x^*, \lambda^*) = 0, \quad \nabla_\lambda L(x^*, \lambda^*) = 0$$

There are $n + m$ unknowns variables and $n + m$ equations.

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) = 0 \qquad h_i(x^*) = 0, i = 1, ..., m$$

Example-1

Consider the optimization problem:

$$\min_{x_1,x_2,x_3} \frac{1}{2}(x_1^2 + x_2^2 + x_3^2)$$
$$\text{s.t. } x_1 + x_2 + x_3 = 3$$

Linear independent

The Lagrangian function is

$$L(x,\lambda) = \frac{1}{2}(x_1^2 + x_2^2 + x_3^2) + \lambda(x_1 + x_2 + x_3 - 3)$$

Then the necessary condition is

$$x_1^* + \lambda^* = 0, \;\; x_2^* + \lambda^* = 0$$
$$x_3^* + \lambda^* = 0, \;\; x_1^* + x_2^* + x_3^* = 3$$

Therefore, $x^* = (1,1,1)$, $\lambda^* = -1$.

Example-2

Consider the optimization problem:

$$\min_{x_1, x_2} \ (x_1 - 2)^2 + (x_2 - 2)^2$$
$$\text{s.t. } x_1 + x_2 - 6 = 0$$

Linearly independent

The Lagrangian function is

$$L(x, \lambda) = (x_1 - 2)^2 + (x_2 - 2)^2 + \lambda(x_1 + x_2 - 6)$$

Then the necessary condition is

$$\begin{aligned} 2(x_1^* - 2) + \lambda^* &= 0 \\ 2(x_2^* - 2) + \lambda^* &= 0 \\ x_1^* + x_2^* - 6 &= 0 \end{aligned}$$

Therefore, $x^* = (3, 3)$, $\lambda^* = -2$.

Example-3

Portfolio Selection

We plan to invest in n different assets, indexed by $i = 1, ..., n$. The total wealth is 1.
The return e_i is random with an expectation of $\bar{e}_i$ and the covariance matrix $Q \in \mathbb{R}^{n\times n}$ with its (ij)-th item $Q_{ij} = \mathbb{E}[(e_i - \bar{e}_i)(e_j - \bar{e}_j)]$.
The expected return is R, and we want to minimize the variance of return $x^T Q x$.

$$\min_x \; x^T Q x$$

$$\text{s.t.} \; \sum_{i=1}^n x_i = 1 \qquad \boldsymbol{\lambda_1}$$

$$\sum_{i=1}^n \bar{e}_i x_i = R \qquad \boldsymbol{\lambda_2}$$

* Suppose $(\bar{e}_1, \ldots, \bar{e}_n) \neq a(1, \ldots, 1)$

Example-3

Portfolio Selection

The Lagrangian function is

$$L(x, \lambda) = x^T Qx + \lambda_1(\sum_{i=1}^{n} x_i - 1) + \lambda_2(\sum_{i=1}^{n} \bar{e}_i x_i - R)$$

Then the necessity condition is

$$\begin{aligned} 2Qx^* + \lambda_1^* u + \lambda_2^* \bar{e} &= 0 \\ u^T x^* &= 1 \\ \bar{e}^T x^* &= R \end{aligned}$$

When $\bar{e} \neq au$
Linearly independent

Replace x^* we can obtain λ_1^*, λ_2^* as the solution of

$$\begin{aligned} \frac{\lambda_1^*}{2} u^T Q^{-1} u + \frac{\lambda_2^*}{2} u^T Q^{-1} \bar{e} &= -1 \\ \frac{\lambda_1^*}{2} \bar{e}^T Q^{-1} u + \frac{\lambda_2^*}{2} \bar{e}^T Q^{-1} \bar{e} &= -R \end{aligned}$$

where $u = (1, ..., 1)^T, \bar{e} = (\bar{e}_1, ..., \bar{e}_n)^T$.

Example-3

Portfolio Selection

First, according to definition $Q_{ij} = E[(e_i - \bar{e}_i)(e_j - \bar{e}_j)]$

$$= E[(e_j - \bar{e}_j)(e_i - \bar{e}_i)] = Q_{ji}$$

Q is a symmetric matrix

$$L(x, \lambda) = x^T Q x + \lambda_1 \left(\sum_{i=1}^{n} x_i - 1\right) + \lambda_2 \left(\sum_{i=1}^{n} \bar{e}_i x_i - R\right)$$

$$= (x_1, x_2 \ldots x_n) \begin{pmatrix} Q_{11} & Q_{12} & \cdots & Q_{1n} \\ Q_{21} & Q_{22} & \cdots & Q_{2n} \\ \vdots & & & \\ Q_{n1} & Q_{n2} & \cdots & Q_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \lambda_1 \left(\sum_{i=1}^{n} x_i - 1\right) + \lambda_2 \left(\sum_{i=1}^{n} \bar{e}_i x_i - R\right)$$

$$= \left(\sum_{i=1}^{n} x_i Q_{i1}, \sum_{i=1}^{n} x_i Q_{i2}, \ldots, \sum_{i=1}^{n} x_i Q_{in}\right) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \lambda_1 \left(\sum_{i=1}^{n} x_i - 1\right) + \lambda_2 \left(\sum_{i=1}^{n} \bar{e}_i x_i - R\right)$$

$$= x_1 \sum_{i=1}^{n} x_i Q_{i1} + x_2 \sum_{i=1}^{n} x_i Q_{i2} + \cdots + x_n \sum_{i=1}^{n} x_i Q_{in} + \lambda_1 \left(\sum_{i=1}^{n} x_i - 1\right) + \lambda_2 \left(\sum_{i=1}^{n} \bar{e}_i x_i - R\right)$$

Example-3

Portfolio Selection

when taking the partial derivative with respet to x_1. For exampe:

$$\frac{\partial L}{\partial x_1} = \sum_{i=1}^{n} x_i \theta_{i1} + x_1\theta_{i1} + x_2\theta_{i2} + \cdots + x_n\theta_{in} + \lambda_1 + \bar{e}_1\lambda_2$$

because θ is symmetric

$\theta_{i1} = \theta_{1i}$

$\theta_{i2} = \theta_{2i}$

$\vdots$

$\theta_{in} = \theta_{ni}$

$$\Rightarrow\ = 2\sum_{i=1}^{n} Q_{1i} x_i + \lambda_1 + \bar{e}_1\lambda_2$$

Similarly. $\frac{\partial L}{\partial x_2} = 2\sum_{i=1}^{n} \theta_{2i} x_i + \lambda_1 + \bar{e}_2\lambda_2$

$\vdots$

$$\frac{\partial L}{\partial x_n} = 2\sum_{i=1}^{n} \theta_{ni} x_i + \lambda_1 + \bar{e}_n\lambda_2$$

$$\therefore \begin{pmatrix} \frac{\partial L}{\partial x_1} \\ \frac{\partial L}{\partial x_2} \\ \vdots \\ \frac{\partial L}{\partial x_n} \end{pmatrix} = 2\begin{pmatrix} \sum_{i=1}^{n}\theta_{1i}x_i \\ \sum_{i=1}^{n}\theta_{2i}x_i \\ \vdots \\ \sum_{i=1}^{n}\theta_{ni}x_i \end{pmatrix} + \lambda_1\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + \lambda_2\begin{pmatrix} \bar{e}_1 \\ \bar{e}_2 \\ \vdots \\ \bar{e}_n \end{pmatrix}$$

$$= 2\begin{pmatrix} \theta_{11} & \theta_{12} & \cdots & \theta_{1n} \\ \theta_{21} & \theta_{22} & \cdots & \theta_{2n} \\ \vdots & \vdots & & \vdots \\ \theta_{n1} & \theta_{n2} & \cdots & \theta_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \lambda_1\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + \lambda_2\begin{pmatrix} \bar{e}_1 \\ \bar{e}_2 \\ \vdots \\ \bar{e}_n \end{pmatrix}$$

$$= 2Q\chi + \lambda_1 u + \lambda_2\bar{e}$$

Constrained Optimization with Equality

Sufficiency condition

When f and $h_i, \forall i$ are *twice* continuously differentiable.
If $x^* \in \mathbb{R}^n$ and $\lambda^* \in \mathbb{R}^m$ satisfy

$$\nabla_x L(x^*, \lambda^*) = 0, \;\; \nabla_\lambda L(x^*, \lambda^*) = 0$$

$$y^T \nabla^2_{xx} L(x^*, \lambda^*) y > 0, \forall y \neq 0, \nabla h(x^*)^T y = 0$$

Then x^* is a strict local minimum.

$$\left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(x^*) \right)$$

If $>$, then *strict*
If $\geq$, then *relative*

How about *local* v.s. *global*? Next lecture

Example

Consider the optimization problem:

$$\min_{x_1,x_2,x_3} \quad -(x_1x_2 + x_2x_3 + x_1x_3)$$
$$\text{s.t. } x_1 + x_2 + x_3 = 3$$

(Review) **First**, we obtain the x^* and λ^* according to the *necessary condition.* The Lagrangian function is

$$L(x, \lambda) = -(x_1x_2 + x_2x_3 + x_1x_3) + \lambda(x_1 + x_2 + x_3 - 3)$$

Then the necessity condition is

$$\begin{aligned} -x_2^* - x_3^* + \lambda^* &= 0 \\ -x_1^* - x_3^* + \lambda^* &= 0 \\ -x_1^* - x_2^* + \lambda^* &= 0 \\ x_1^* + x_2^* + x_3^* &= 3 \end{aligned}$$

Therefore, $x^* = (1, 1, 1)^T$, $\lambda^* = 2$.

Example

Second, we check the *sufficiency condition.*

$$\nabla^2_{xx} L(x^*, \lambda^*) = \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}$$

Let $y \neq 0$ be a vector that satisfies $\nabla h(x^*)^T y = 0$, i.e. $y_1 + y_2 + y_3 = 0$. Therefore,

$$\begin{aligned} & y^T \nabla^2_{xx} L(x^*, \lambda^*) y \\ = & -y_1(y_2 + y_3) - y_2(y_1 + y_3) - y_3(y_1 + y_2) \\ = & -y_1(-y_1) - y_2(-y_2) - y_3(-y_3) \\ = & \ y_1^2 + y_2^2 + y_3^2 > 0 \end{aligned}$$

Hence, x^* is a strict local optimum.

Sensitivity Theorem (Extended Reading)*

Consider a family of optimization problems:

$$\min_{x} \ f(x)$$
$$\text{s.t. } h(x) = u$$

where u is a parameter.

Suppose when $u = 0$, the above problem has a local minimum x^* that is regular and associated with a unique λ^* satisfying the sufficiency condition.

Then, there exists an open sphere $\mathcal{S}$ centered at $u = 0$ such that:
(1) for every $u \in \mathcal{S}$, there is a local-minimum-Lagrange multiplier $x(u), \lambda(u)$.
(2) $x(u)$ and $\lambda(u)$ are continuously differentiable with $x(0) = x^*$ and $\lambda(0) = \lambda^*$.
(3)Denote $f(x(u))$ as $\mathcal{F}(u)$, we have

$$\nabla \mathcal{F}(u) = -\lambda(u), \forall u \in \mathcal{S}$$

The impact of one unit change of u on the optimal objective value

Sensitivity Theorem (Extended Reading)*

Consider this optimization problem:

$$\min_x \; f(x) = \frac{1}{2}(x_1^2 - x_2^2) - x_2$$
$$\text{s.t. } h(x) = x_2 = u$$

The Lagrangian function is

$$L(x, \lambda) = \frac{1}{2}(x_1^2 - x_2^2) - x_2 + \lambda(x_2 - u)$$

Then

$$x_1^* = 0, -x_2^* - 1 + \lambda^* = 0, x_2^* - u = 0$$

Therefore $x(u) = (0, u)$, $\lambda(u) = 1 + u$.

$$\mathcal{F}(u) = \; -\frac{1}{2}u^2 - u$$

and $\nabla\mathcal{F}(0) = -u - 1 = -\lambda(0) = -1$, consistent with the sensitivity theorem.

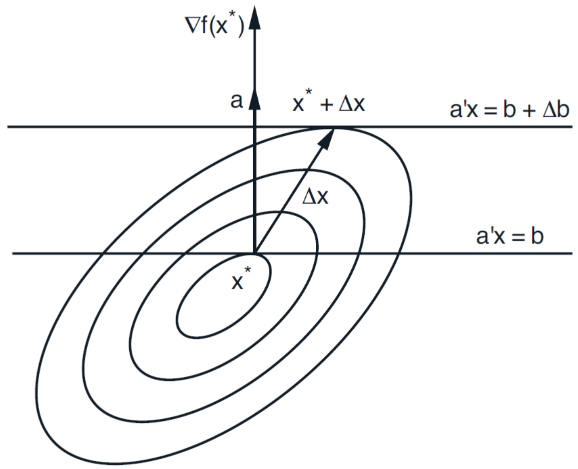
Sensitivity Theorem* (Extended Reading)

Consider this problem:

$$\min_x \ f(x)$$
$$\text{s.t. } a^T x = b \ : \ \lambda$$

If coefficient b changes to $b + \Delta b$, the minimum x^* will change to $x^* + \Delta x$.
Since $a^T x^* = b$ and $a^T(x^* + \Delta x) = a^T x^* + a^T \Delta x = b + \Delta b$.
So we have $a^T \Delta x = \Delta b$. According to the condition $\nabla f(x^*) = -\lambda^* a$.

$$\begin{aligned}\Delta f &= f(x^* + \Delta x) - f(x^*) \\ &= \nabla f(x^*)^T \Delta x + o(||\Delta x||) \\ &= -\lambda^* a^T \Delta x + o(||\Delta x||) \\ &= -\lambda^* \Delta b + o(||\Delta x||)\end{aligned}$$



Therefore

$$\lambda^* = -\frac{\Delta f}{\Delta b}$$

A generalization of “Shadow price” in Lecture 3 (next slide)

Lecture-3: Economic Interpretation

According to Optimality criterion, we have

$$f^* = c^T x^* = b^T \lambda^* = \lambda_1^* b_1 + \cdots + \lambda_N^* b_N$$

If parameter b_n changes, what is the impact on the optimal value f^*?

$$\frac{\partial f^*}{\partial b_1} = \lambda_1^*, ..., \frac{\partial f^*}{\partial b_N} = \lambda_N^*$$

Therefore, λ_n^* can be interpreted as the change of f^* should there be 1 unit change of b_n. We call it "*shadow price*" in economics.

The scarcer the resource, the greater the impact of its changes on the objective function (cost), and therefore the higher the shadow price.

Constrained Optimization with Inequality

General Form:

$$\min_x \; f(x)$$
$$\text{s.t. } h(x) = 0, \boxed{g(x) \leq 0}$$

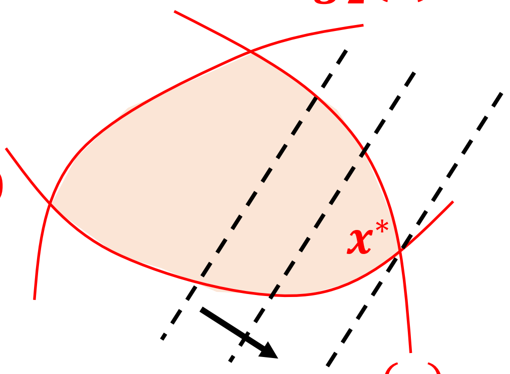
where $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^r$ are continuously differentiable. Here,

$$h = (h_1, ..., h_m)$$
$$g = (g_1, ..., g_r)$$

Constrained Optimization with Inequality

Let's look at the inequality constraint $g(x) \le 0$. If x^* is a local minimum, then

- If $g_j(x^*) = 0$, then the j-th constraint is active **Treated as equality**
- If $g_j(x^*) < 0$, then the j-th constraint is inactive **Doesn't matter**



Let $\mathcal{A}(x) \coloneqq \{j | g_j(x) = 0\}$ be the set of active constraints

Assume that x^* is regular, similarly, we can write down the Lagrangian function

$$L(x, \lambda, \mu) = f(x^*) + \sum_{i=1}^{m} \lambda_i^* h_i(x^*) + \boxed{\sum_{j=1}^{r} \mu_j^* g_j(x^*)}$$

- Let the μ_j^* corresponds to inactive constraint equal to zero. $\mu_j^* = 0, \forall j \notin \mathcal{A}(x)$
- Let $\mu_j^* \ge 0, \forall j$ (explain later)

Constrained Optimization with Inequality

We try to explain the logic behind "$\mu_j^* \geq 0, \forall j$" using sensitivity theorem.
Relax the j-th constraint to $g_j(x) = u_j, u_j > 0$. Since $\Delta f \leq 0$, we have

$$\mu_j^* = -(\Delta f)/u_j \geq 0$$

$$\lambda^* = -\frac{\Delta f}{\Delta x}$$

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^{r} \mu_j \nabla g_j(x^*) = 0$$

$$h_i(x^*) = 0, \forall i = 1, \ldots, m$$

Point that satisfies these conditions is called Karush-Kuhn-Tucker (KKT) point

$$g_j(x^*) = 0, \mu_j^* \geq 0, \forall j \in \mathcal{A}(x^*)$$

$$g_j(x^*) < 0, \mu_j^* = 0, \forall j \notin \mathcal{A}(x^*)$$

$$g_j(x^*)\mu_j^* = 0, g_j(x^*) \leq 0, \mu_j^* \geq 0, \forall j = 1, \ldots, r$$

or $0 \leq -g_j(x^*) \perp \mu_j^* \geq 0$

Constrained Optimization with Inequality

Steps to write down the KKT condition

1. Turn it into standard form:

$$\min_x \ f(x)$$
$$\text{s.t. } h(x) = 0, g(x) \leq 0$$

2. Write down the Lagrangian function

$$L(x, \lambda, \mu) = f(x^*) + \sum_{i=1}^{m} \lambda_i^* h_i(x^*) + \sum_{j=1}^{r} \mu_j^* g_j(x^*)$$

3. The KKT condition is

$$\nabla_x L(x^*, \lambda^*, \mu^*) = 0$$
$$h_i(x^*) = 0, \forall i = 1, ..., m$$
$$0 \leq -g_j(x^*) \perp \mu_j^* \geq 0, \forall j = 1, ..., r$$

Example-1

Consider this optimization problem:

$$\begin{aligned}\min_{x_1,x_2} \ & (x_1-2)^2+x_2^2 \\ \text{s.t. } & x_1-x_2^2\geq 0 \\ & -x_1+x_2\geq 0\end{aligned}$$

Are $x^{(1)}=(0,0)^T$, $x^{(2)}=(1,1)^T$ KKT points?

Solution: For both points, the equality holds for all constraints.

$$L(x,\mu)=(x_1-2)^2+x_2^2+\mu_1(-x_1+x_2^2)+\mu_2(x_1-x_2)$$

Therefore, the KKT conditions are

$$\begin{aligned}2x_1-4-\mu_1+\mu_2&=0\\ 2x_2+2\mu_1x_2-\mu_2&=0\\ 0\leq\mu_1\perp(x_1-x_2^2)&\geq 0\\ 0\leq\mu_2\perp(-x_1+x_2)&\geq 0\end{aligned}$$

For point $x^{(1)}=(0,0)^T$, we have $\mu_2=0$ and $\mu_1=-4<0$, no feasible μ.
For point $x^{(2)}=(1,1)^T$, $\mu=(0,2)^T$.

Example-2

Find the KKT point of problem:

$$\begin{aligned} \min_{x} \ & f(x) = (x_1 - 1)^2 + x_2 \\ \text{s.t. } & g_1(x) = -x_1 - x_2 + 2 \geq 0 \\ & g_2(x) = x_2 \geq 0 \end{aligned}$$

Solution: The Lagrange function is

$$L(x, \mu) = (x_1 - 1)^2 + x_2 + \mu_1(x_1 + x_2 - 2) + \mu_2(-x_2)$$

Then the KKT condition is

$$\begin{aligned} 2(x_1 - 1) + \mu_1 &= 0 \\ 1 + \mu_1 - \mu_2 &= 0 \\ 0 \leq (-x_1 - x_2 + 2) \perp \mu_1 &\geq 0 \\ 0 \leq x_2 \perp \mu_2 &\geq 0 \end{aligned}$$

We can get $x^* = (1, 0)$, $\mu^* = (0, 1)$.

Kuhn-Tucker Necessary Conditions

Let x^* be a local minimum and a regular point.
Then there exist unique Lagrange multiplier vectors $\lambda^* = (\lambda_1^*, ..., \lambda_m^*)$, $\mu^* = (\mu_1^*, ..., \mu_r^*)$

$$\nabla_x L(x^*, \lambda^*, \mu^*) = 0$$

$$h_i(x^*) = 0, \forall i = 1, ..., m$$

$$\boxed{0 \leq -g_j(x^*) \perp \mu_j^* \geq 0, \forall j = 1, ..., r}$$

Complementary Slackness

If f, h, and g are twice continuously differentiable, then

$$y^T \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) y \geq 0, \forall y \in V(x^*)$$

where

$$V(x^*) = \{y | \nabla h_i(x^*)^T y = 0, \forall i = 1, ..., m; \nabla g_j(x^*)^T y = 0, \forall j \in \mathcal{A}(x^*)\}$$

A local minimum is a KKT point

How about sufficiency condition? Convex optimization (next lecture)

Thanks！