Lagrange Multipliers

Optimization Methods in Management Science

Master in Management

HEC Lausanne

Dr. Rodrigue Oeuvray

Fall 2019 Semester

Outline

- Equality constraints
- Equality and inequality constraints
- Sufficient conditions

Optimization with Equality Constraints

General formulation:

$$\begin{array}{rrcl} & \underset{\boldsymbol{x} \in \mathbb{R}^n}{opt} & f(\boldsymbol{x}) & \\ & h_1(\boldsymbol{x}) & = & b_1 \\ \text{s.t.} & & \vdots & \\ & h_l(\boldsymbol{x}) & = & b_l \end{array}$$

opt means *min* or *max*

Optimization with Equality Constraints (Cont'd)

- The lagrangian is defined by:

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{j=1}^{l} \lambda_j (h_j(\boldsymbol{x}) - b_j)$$

- Each constraint function is multiplied by a variable $\lambda_j \in \mathbb{R}$, called a lagrange multiplier

Optimization with Equality Constraints (Cont'd)

The reason L is of interest is the following:

Proposition

Assume that $\boldsymbol{x}^* = (x_1^*, \ldots, x_n^*)$ *maximizes or minimizes* $f(\boldsymbol{x})$ *subject to the constraints* $h_i(\boldsymbol{x}) = b_i, i = 1, \ldots, l$. *If the vectors* $\nabla h_1(\boldsymbol{x}^*), \ldots, \nabla h_l(\boldsymbol{x}^*)$ *are linearly independent, then there exists a vector* $\boldsymbol{\lambda} = (\lambda_1^*, \ldots, \lambda_l^*)$ *such that* $\nabla L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = 0$, *i.e*

$$\frac{\partial L}{\partial x_1}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = \cdots = \frac{\partial L}{\partial x_n}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = 0$$

$$\frac{\partial L}{\partial \lambda_1}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = \cdots = \frac{\partial L}{\partial \lambda_l}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = 0$$

Example 1

- We consider the following example:

$$\begin{array}{ll} min & x_1 + x_2 + x_3^2 \\ \text{s.t.} & \begin{array}{rcl} x_1 & = & 1 \\ x_1^2 + x_2^2 & = & 1 \end{array} \end{array}$$

- It is easy to check that the minimum is achieved at $(x_1, x_2, x_3) = (1, 0, 0)$
- The associated lagrangian is

$$L(x_1, x_2, x_3, \lambda_1, \lambda_2) = x_1 + x_2 + x_3^2 + \lambda_1(x_1 - 1) + \lambda_2(x_1^2 + x_2^2 - 1)$$

Example 1 (Cont'd)

- Observe that

$$\frac{\partial L}{\partial x_2}(1,0,0,\lambda_1,\lambda_2) = 1 \quad \forall \lambda_1, \lambda_2$$

- Consequently, $\frac{\partial L}{\partial x_2}$ does not **vanish** at the **optimal** solution !
- **Case (i)** of the former proposition occurs:

$$\nabla h_1(1,0,0) = (1\ 0\ 0)^T \quad \text{and} \quad \nabla h_2(1,0,0) = (2\ 0\ 0)^T$$

are **linearly dependent** vectors !

Optimization with Equality Constraints (Cont'd)

- Once we have found a **candidate** solution $\boldsymbol{x}^*$, it is not always easy to figure out whether they correspond to a minimum, a maximum or neither
- But in the two following situations, one can conclude

Proposition

1. *If $f(\boldsymbol{x})$ is concave and all of the $h_i(\boldsymbol{x})$ are linear, then any feasible $\boldsymbol{x}^*$ with a corresponding $\boldsymbol{\lambda}^*$ making $\nabla L(\boldsymbol{x}^*, \boldsymbol{\lambda}) = 0$ maximizes $f(\boldsymbol{x})$ subject to the constraints*
2. *Similarly, If $f(\boldsymbol{x})$ is convex and all of the $h_i(\boldsymbol{x})$ are linear, then any feasible $\boldsymbol{x}^*$ with a corresponding $\boldsymbol{\lambda}^*$ making $\nabla L(\boldsymbol{x}^*, \boldsymbol{\lambda}) = 0$ minimizes $f(\boldsymbol{x})$ subject to the constraints*

Example 2

- We consider the following example:

$$\begin{aligned} \mathit{min} \quad & 2x_1^2 + x_2^2 \\ \text{s.t.} \quad & x_1 + x_2 = 1 \end{aligned}$$

- The associated lagrangian is

$$L(x_1, x_2, \lambda_1) = 2x_1^2 + x_2^2 + \lambda_1(x_1 + x_2 - 1)$$

- First-order conditions:

$$\begin{aligned} \frac{\partial L}{\partial x_1}(x_1^*, x_2^*, \lambda^*) &= 4x_1^* + \lambda_1^* = 0 \\ \frac{\partial L}{\partial x_2}(x_1^*, x_2^*, \lambda^*) &= 2x_2^* + \lambda_1^* = 0 \\ \frac{\partial L}{\partial \lambda}(x_1^*, x_2^*, \lambda^*) &= x_1^* + x_2^* - 1 = 0 \end{aligned}$$

Example 2 (Cont'd)

- We get the following candidate solution: $x_1^* = 1/3, x_2^* = 2/3$ and $\lambda_1^* = -4/3$ with function value 2/3
- The hessian of f is given by:

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$$

- Since $f(x_1, x_2)$ is **convex** (its hessian is positive semi-definite and even positive definite) and as the constraint is linear, then we conclude that this candidate solution is the **unique global minimum** of this problem

Minimization with Equality and Inequality Constraints

We now consider the general case for a **minimization** problem:

$$
\begin{aligned}
min_{\boldsymbol{x} \in \mathbb{R}^n} \quad & f(\boldsymbol{x}) \\
s.t. \quad & g_i(\boldsymbol{x}) \leq 0, \quad i = 1, \dots, k \\
& h_i(\boldsymbol{x}) = 0, \quad i = 1, \dots, l
\end{aligned}
$$

Its lagrangian is defined by

$$L(\boldsymbol{x}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = f(\boldsymbol{x}) + \sum_{i=1}^{k} \alpha_i g_i(\boldsymbol{x}) + \sum_{i=1}^{l} \beta_i h_i(\boldsymbol{x})$$

We say that the ith constraint $g_i(\boldsymbol{x}) \leq 0$ is active at a point $\bar{\boldsymbol{x}}$ if $g_i(\bar{\boldsymbol{x}}) = 0$

Min with Equality and Inequality Constraints (Cont'd)

The fundamental result is the following:

Proposition

*Assume that $\boldsymbol{x}^* = (x_1^*, \ldots, x_n^*)$ minimizes $f(\boldsymbol{x})$ subject to the constraints $g_i(\boldsymbol{x}) \leq 0, i = 1, \ldots, k$ and $h_i(\boldsymbol{x}) = 0, i = 1, \ldots, l$. If the gradient at $\boldsymbol{x}^*$ of the **active** constraints (including the equality constraints h_i) are linearly independent, then there exists vectors $\boldsymbol{\alpha}^* = (\alpha_1^*, \ldots, \alpha_k^*)$ and $\boldsymbol{\beta}^* = (\beta_1^*, \ldots, \beta_l^*)$ such that*

$$
\begin{aligned}
\frac{\partial}{\partial x_i} L(\boldsymbol{x}^*, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) &= 0, \quad i = 1, \ldots, n \\
h_i(\boldsymbol{x}^*) &= 0, \quad i = 1, \ldots, l \\
\alpha_i^* g_i(\boldsymbol{x}^*) &= 0, \quad i = 1, \ldots, k \text{ (dual complementary cond.)} \\
g_i(\boldsymbol{x}^*) &\leq 0, \quad i = 1, \ldots, k \\
\alpha_i^* &\geq 0, \quad i = 1 \ldots, k
\end{aligned}
$$

Example 3

- We consider the following example:

$$\begin{array}{llcll} min & (x-2)^2+2(y-1)^2 & & & \\ \text{s.t.} & x+4y & \leq & 3 & \\ & x & \geq & y & \iff \quad (\text{-}x+y\leq 0) \end{array}$$

- The associated lagrangian is

$$L(x,y,\alpha_1,\alpha_1)=(x-2)^2+2(y-1)^2+\alpha_1(x+4y-3)+\alpha_2(\text{-}x+y)$$

- System of equations to solve:

$$\begin{aligned} {}''\frac{\partial L}{\partial x}'' &: 2(x-2)+\alpha_1-\alpha_2=0 \\ {}''\frac{\partial L}{\partial y}'' &: 4(y-1)+4\alpha_1+\alpha_2=0 \\ {}''\alpha_1 g_1'' &: \alpha_1(x+4y-3)=0 \end{aligned}
\qquad
\begin{aligned} {}''\alpha_2 g_2'' &: \alpha_2(\text{-}x+y)=0 \\ {}''g_1'' &: x+4y-3\leq 0 \\ {}''g_2'' &: \text{-}x+y\leq 0 \\ & \alpha_1,\alpha_2\geq 0 \end{aligned}$$

Example 3 (Cont'd)

Since they are **two** complementary conditions ($\alpha_i g_i(\boldsymbol{x}) = 0, i = 1, 2$), there are **four** cases to check : 1. $\alpha_1 = \alpha_2 = 0$, 2. $\alpha_1 = g_2(\boldsymbol{x}) = 0$, 3. $\alpha_2 = g_1(\boldsymbol{x}) = 0$, 4. $g_1(\boldsymbol{x}) = g_2(\boldsymbol{x}) = 0$

1. $\alpha_1 = 0, \alpha_2 = 0$: we get $x = 2, y = 1$ which is not feasible $(x + 4y - 3 > 0)$
2. $\alpha_1 = 0, \text{-}x + y = 0$: we get $x = 4/3, y = 4/3, \alpha_2 = \text{-}4/3$ which is not feasible $(\alpha_2 < 0)$
3. $\alpha_2 = 0, x + 4y - 3 = 0$: we get $x = 5/3, y = 1/3, \alpha_1 = 2/3$. This solution is feasible
4. $x + 4y - 3 = 0$, $\text{-}x + y = 0$: we get $x = 3/5, y = 3/5, \alpha_1 = 22/25$, $\alpha_2 = \text{-}48/25$ which is not feasible $(\alpha_2 < 0)$

We conclude that the **optimal** solution is given by $x = 5/3$ and $y = 1/3$

Sufficient Conditions

- The KKT conditions give us candidate optimal solutions. When are these conditions sufficent for optimality ?

Sufficent Conditions

When f and the $g_i, i = 1, \ldots, k$ are **convex**, the $h_i, i = 1, \ldots, l$ are **affine**, and the problem **satisfies** the Slater's condition, then any point that satisfies the KKT conditions gives a point that **minimizes** $f(\boldsymbol{x})$ subject to the constraints

The Slater's condition corresponds to the existence of some **feasible** $\boldsymbol{x}$ so that $g_i(\boldsymbol{x}) < 0, \ \forall i$

Breach of Slater's Condition

- We consider the following example:

$$\begin{aligned} &\mathit{min} \quad x \\ &\text{s.t.} \quad x^2 \leq 0 \end{aligned}$$

- The **optimal** solution is obvioulsy $x = 0$
- The lagrangian of this problem is

$$L(x, \alpha) = x + \alpha x^2$$

- The equation $\frac{\partial L}{\partial x}(x, \alpha) = 1 + 2\alpha x = 0$ is **not satisfied** at the optimum !
- KKT conditions are not satisfied since **Slater's condition does not hold** ! The condition $x^2 < 0$ is simply not possible

Maximization with Equality and Inequality Constraints

We now consider the general case for a **maximization** problem:

$$\begin{aligned} max_{\boldsymbol{x}\in\mathbb{R}^n} \quad & f(\boldsymbol{x}) \\ s.t. \quad & g_i(\boldsymbol{x}) \leq 0, \quad i = 1, \ldots, k \\ & h_i(\boldsymbol{x}) = 0, \quad i = 1, \ldots, l \end{aligned}$$

Its lagrangian is defined by

$$L(\boldsymbol{x}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = f(\boldsymbol{x}) - \sum_{i=1}^{k} \alpha_i g_i(\boldsymbol{x}) + \sum_{i=1}^{l} \beta_i h_i(\boldsymbol{x}), \quad \alpha_i \geq 0, i = 1, \ldots, k$$

Max with Equality and Inequality Constraints (Cont'd)

The fundamental result is the following:

Proposition

Assume that $\boldsymbol{x}^* = (x_1^*, \ldots, x_n^*)$ *maximizes* $f(\boldsymbol{x})$ *subject to the constraints* $g_i(\boldsymbol{x}) \leq 0, i = 1, \ldots, k$ *and* $h_i(\boldsymbol{x}) = 0, i = 1, \ldots, l$*. If the gradient at* $\boldsymbol{x}^*$ *of the active constraints (including the equality constraints* h_i*) are linearly independent, then there exists vectors* $\boldsymbol{\alpha}^* = (\alpha_1^*, \ldots, \alpha_k^*)$ *and* $\boldsymbol{\beta}^* = (\beta_1^*, \ldots, \beta_l^*)$ *such that*

$$
\begin{aligned}
\frac{\partial}{\partial x_i} L(\boldsymbol{x}^*, \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*) &= 0, \quad i = 1, \ldots, n \\
h_i(\boldsymbol{x}^*) &= 0, \quad i = 1, \ldots, l \\
\alpha_i^* g_i(\boldsymbol{x}^*) &= 0, \quad i = 1, \ldots, k \text{ (dual complementary cond.)} \\
g_i(\boldsymbol{x}^*) &\leq 0, \quad i = 1, \ldots, k \\
\alpha_i^* &\geq 0, \quad i = 1 \ldots, k
\end{aligned}
$$

Important Remark

For the following **maximization** problem:

$$\begin{aligned} max_{\boldsymbol{x} \in \mathbb{R}^n} \quad & f(\boldsymbol{x}) \\ s.t. \quad & g_i(\boldsymbol{x}) \leq 0, \quad i = 1, \ldots, k \\ & h_i(\boldsymbol{x}) = 0, \quad i = 1, \ldots, l, \end{aligned}$$

we could equivalently define its lagrangian by:

$$L(\boldsymbol{x}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = f(\boldsymbol{x}) + \sum_{i=1}^{k} \alpha_i g_i(\boldsymbol{x}) + \sum_{i=1}^{l} \beta_i h_i(\boldsymbol{x})$$

but with

$$\alpha_i \leq 0, i = 1, \ldots, k$$

Example 4

- We consider the following example:

$$\begin{aligned} max \quad & x^3 - 3x \\ \text{s.t.} \quad & x \leq 2 \end{aligned}$$

- The associated lagrangian is

$$L(x, \alpha_1) = x^3 - 3x - \alpha_1(x - 2)$$

- System of equations to solve:

$$"\frac{\partial L}{\partial x}" : 3x^2 - 3 - \alpha_1 = 0 \qquad "\alpha_1 g_1" : \alpha_1(x-2) = 0$$

$$"g_1" : x - 2 \leq 0 \qquad \alpha_1 \geq 0$$

Example 4 (Cont'd)

For the complementary condition, there are two cases : 1. $\alpha_1 = 0$,
2. $x = 2$ to check :

1. If $\alpha_1 = 0$ then $3x^2 - 3 = 0$ so $x = 1$ or $x = -1$. Both solutions are feasible. $f(1) = -2, f(-1) = 2$
2. If $x = 2$, then $\alpha_1 = 9$. We obtain a new feasible solution. $f(2) = 2$

We conclude that we have **two global** maximas: $x = -1$ and $x = 2$

Sufficient Conditions

- The KKT conditions give us candidate optimal solutions. When are these conditions sufficent for optimality ?

Sufficent Conditions

When f is **concave**, the $g_i, i = 1, \ldots, k$ are **convex**, the $h_i, i = 1, \ldots, l$ are **affine**, and the problem satisfies the Slater's condition, then any point that satisfies the KKT conditions gives a point that **maximizes** $f(\boldsymbol{x})$ subject to the constraints

Review of Optimality Conditions - 1

Review of Optimality Conditions

Single Variable - Unconstrained

- Solve $f'(x) = 0$ to get candidates x^*
- If $f''(x^*) > 0$, it is a local min
- If $f''(x^*) < 0$, it is a local max
- If $f(x)$ is convex, a local min is a global min
- If $f(x)$ is concave, a local max is a global max

Review of Optimality Conditions - 2

Review of Optimality Conditions

Multipe Variables - Unconstrained

- Solve $\nabla f(x) = 0$ to get candidates $\boldsymbol{x}^*$
- If the hessian at $\boldsymbol{x}^*$ is positive definite, it is a local min
- If the hessian at $\boldsymbol{x}^*$ is negative definite, it is a local max
- If $f(\boldsymbol{x})$ is convex, a local min is a global min
- If $f(\boldsymbol{x})$ is concave, a local max is a global max

Review of Optimality Conditions - 3

Review of Optimality Conditions

Multipe Variables - Constrainted

- Solve KKT conditions to get candidates $\boldsymbol{x}^*$
- The best candidate is the optimum **if** it is feasible **and** if the optimum **exists**
- If max problem, any $\boldsymbol{x}^*$ is an optimum **if** f concave, g_i convex, h_i affine, and the problem satisfies the Slater's condition
- If min problem, any $\boldsymbol{x}^*$ is an optimum **if** f convex, g_i convex, and h_i affine, and the problem satisfies the Slater's condition

Remark. It is not obvious to determine if an optimization has an optimum or not

Existence of an Optimum

- The Weierstrass extreme value theorem states that a **continuous** function on a **closed** and **bounded** set obtains its extreme values (min or max)
  - Examples of closed sets: $[a, b], x^2 + y^2 \leq 1, x^2 + y^2 = 1, \ldots$
  - Examples of non-closed sets: $[a, b[, x^2 + y^2 < 1, \ldots$
- When the domain is not bounded, then we don't have the guarantee that an optimum exists
  - The problem $max\ f(x) = x$ subject to $x \geq 0$ has no finite optimum
  - But the solution of $min\ f(x) = x$ subject to $x \geq 0$ is $x^* = 0$