

Chapter 15

Lagrange Function

Constraint Optimization

Find the extrema of function

$$f(x, y)$$

subject to

$$g(x, y) = c$$

Example:

Find the extrema of function

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

subject to

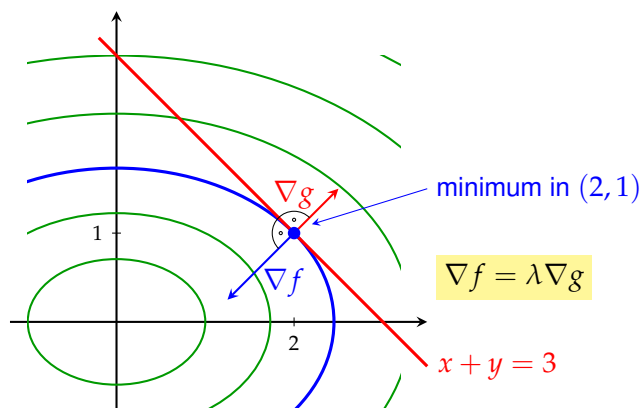
$$g(x, y) = x + y = 3$$

Graphical Solution

For the case of two variables we can find a solution graphically.

1. Draw the constraint $g(x, y) = c$ in the xy -plane.
(The *feasible region* is a curve in the plane)
2. Draw *appropriate* contour lines of objective function $f(x, y)$.
3. Investigate which contour lines of the objective function intersect with the feasible region.
Estimate the (approximate) location of the extrema.

Example – Graphical Solution



Extrema of $f(x, y) = x^2 + 2y^2$ subject to $g(x, y) = x + y = 3$

Lagrange Approach

Let \mathbf{x}^* be an extremum of $f(x, y)$ subject to $g(x, y) = c$.
Then $\nabla f(\mathbf{x}^*)$ and $\nabla g(\mathbf{x}^*)$ are proportional, i.e.,

$$\nabla f(\mathbf{x}^*) = \lambda \nabla g(\mathbf{x}^*)$$

where λ is some proportionality factor.

$$f_x(\mathbf{x}^*) = \lambda g_x(\mathbf{x}^*)$$

$$f_y(\mathbf{x}^*) = \lambda g_y(\mathbf{x}^*)$$

$$g(\mathbf{x}^*) = c$$

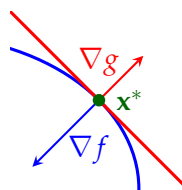
Transformation yields

$$f_x(\mathbf{x}^*) - \lambda g_x(\mathbf{x}^*) = 0$$

$$f_y(\mathbf{x}^*) - \lambda g_y(\mathbf{x}^*) = 0$$

$$c - g(\mathbf{x}^*) = 0$$

The l.h.s. is the gradient of $\mathcal{L}(x, y; \lambda) = f(x, y) + \lambda (c - g(x, y))$.



Lagrange Function

We create a new function from f , g and an auxiliary variable λ ,
called **Lagrange function**:

$$\mathcal{L}(x, y; \lambda) = f(x, y) + \lambda (c - g(x, y))$$

Auxiliary variable λ is called **Lagrange multiplier**.

Local extrema of f subject to $g(x, y) = c$ are critical points of
Lagrange function \mathcal{L} :

$$\mathcal{L}_x = f_x - \lambda g_x = 0$$

$$\mathcal{L}_y = f_y - \lambda g_y = 0$$

$$\mathcal{L}_\lambda = c - g(x, y) = 0$$

Example – Lagrange Function

Compute the local extrema of

$$f(x, y) = x^2 + 2y^2 \quad \text{subject to} \quad g(x, y) = x + y = 3$$

Lagrange function:

$$\mathcal{L}(x, y, \lambda) = (x^2 + 2y^2) + \lambda(3 - (x + y))$$

Critical points:

$$\begin{aligned}\mathcal{L}_x &= 2x - \lambda &= 0 \\ \mathcal{L}_y &= 4y - \lambda &= 0 \\ \mathcal{L}_\lambda &= 3 - x - y &= 0\end{aligned}$$

\Rightarrow unique critical point: $(\mathbf{x}_0; \lambda_0) = (2, 1; 4)$

Bordered Hessian Matrix

Matrix

$$\bar{\mathbf{H}}(\mathbf{x}; \lambda) = \begin{pmatrix} 0 & g_x & g_y \\ g_x & \mathcal{L}_{xx} & \mathcal{L}_{xy} \\ g_y & \mathcal{L}_{yx} & \mathcal{L}_{yy} \end{pmatrix}$$

is called the **bordered Hessian Matrix**.

Sufficient condition for local extremum:

Let $(\mathbf{x}_0; \lambda_0)$ be a critical point of \mathcal{L} .

- ▶ $|\bar{\mathbf{H}}(\mathbf{x}_0; \lambda_0)| > 0 \Rightarrow \mathbf{x}_0$ is a *local maximum*
- ▶ $|\bar{\mathbf{H}}(\mathbf{x}_0; \lambda_0)| < 0 \Rightarrow \mathbf{x}_0$ is a *local minimum*
- ▶ $|\bar{\mathbf{H}}(\mathbf{x}_0; \lambda_0)| = 0 \Rightarrow$ no conclusion possible

Example – Bordered Hessian Matrix

Compute the local extrema of

$$f(x, y) = x^2 + 2y^2 \quad \text{subject to} \quad g(x, y) = x + y = 3$$

Lagrange function: $\mathcal{L}(x, y, \lambda) = (x^2 + 2y^2) + \lambda(3 - x - y)$

Critical point: $(\mathbf{x}_0; \lambda_0) = (2, 1; 4)$

Determinant of the bordered Hessian:

$$|\bar{\mathbf{H}}(\mathbf{x}_0; \lambda_0)| = \begin{vmatrix} 0 & g_x & g_y \\ g_x & \mathcal{L}_{xx} & \mathcal{L}_{xy} \\ g_y & \mathcal{L}_{yx} & \mathcal{L}_{yy} \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 4 \end{vmatrix} = -6 < 0$$

$\Rightarrow \mathbf{x}_0 = (2, 1)$ is a local minimum.

Many Variables and Constraints

Objective function

$$f(x_1, \dots, x_n)$$

and constraints

$$g_1(x_1, \dots, x_n) = c_1$$

\vdots

$$(k < n)$$

$$g_k(x_1, \dots, x_n) = c_k$$

Optimization problem: $\min / \max f(\mathbf{x})$ subject to $\mathbf{g}(\mathbf{x}) = \mathbf{c}$.

Lagrange Function:

$$\mathcal{L}(\mathbf{x}; \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T (\mathbf{c} - \mathbf{g}(\mathbf{x}))$$

Recipe – Critical Points

1. Create Lagrange Function \mathcal{L} :

$$\begin{aligned} \mathcal{L}(x_1, \dots, x_n; \lambda_1, \dots, \lambda_k) \\ = f(x_1, \dots, x_n) + \sum_{i=1}^k \lambda_i (c_i - g_i(x_1, \dots, x_n)) \end{aligned}$$

2. Compute all first partial derivatives of \mathcal{L} .

3. We get a system of $n + k$ equations in $n + k$ unknowns.
Find all solutions.

4. The first n components (x_1, \dots, x_n) are the elements of the critical points.

Example – Critical Points

Compute all critical points of

$$f(x_1, x_2, x_3) = (x_1 - 1)^2 + (x_2 - 2)^2 + 2x_3^2$$

subject to

$$x_1 + 2x_2 = 2 \quad \text{and} \quad x_2 - x_3 = 3$$

Lagrange function:

$$\begin{aligned} \mathcal{L}(x_1, x_2, x_3; \lambda_1, \lambda_2) = & ((x_1 - 1)^2 + (x_2 - 2)^2 + 2x_3^2) \\ & + \lambda_1 (2 - x_1 - 2x_2) + \lambda_2 (3 - x_2 + x_3) \end{aligned}$$

Example – Critical Points

Partial derivatives (gradient):

$$\begin{aligned}\mathcal{L}_{x_1} &= 2(x_1 - 1) - \lambda_1 &= 0 \\ \mathcal{L}_{x_2} &= 2(x_2 - 2) - 2\lambda_1 - \lambda_2 &= 0 \\ \mathcal{L}_{x_3} &= 4x_3 + \lambda_2 &= 0 \\ \mathcal{L}_{\lambda_1} &= 2 - x_1 - 2x_2 &= 0 \\ \mathcal{L}_{\lambda_2} &= 3 - x_2 + x_3 &= 0\end{aligned}$$

We get the critical points of \mathcal{L} by solving this system of equations.

$$x_1 = -\frac{6}{7}, x_2 = \frac{10}{7}, x_3 = -\frac{11}{7}; \lambda_1 = -\frac{26}{7}, \lambda_2 = \frac{44}{7}.$$

The unique critical point of f subject to these constraints is

$$\mathbf{x}_0 = \left(-\frac{6}{7}, \frac{10}{7}, -\frac{11}{7}\right).$$

Bordered Hessian Matrix

$$\bar{\mathbf{H}}(\mathbf{x}; \boldsymbol{\lambda}) = \begin{pmatrix} 0 & \dots & 0 & \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \frac{\partial g_k}{\partial x_1} & \dots & \frac{\partial g_k}{\partial x_n} \\ \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_k}{\partial x_1} & \mathcal{L}_{x_1 x_1} & \dots & \mathcal{L}_{x_1 x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial x_n} & \dots & \frac{\partial g_k}{\partial x_n} & \mathcal{L}_{x_n x_1} & \dots & \mathcal{L}_{x_n x_n} \end{pmatrix}$$

For $r = k + 1, \dots, n$

let $B_r(\mathbf{x}; \boldsymbol{\lambda})$ denote the $(k + r)$ -th leading principle minor of $\bar{\mathbf{H}}(\mathbf{x}; \boldsymbol{\lambda})$.

Sufficient Condition for Local Extrema

Assume that $(\mathbf{x}_0; \boldsymbol{\lambda}_0)$ is a critical point of \mathcal{L} . Then

- ▶ $(-1)^k B_r(\mathbf{x}_0; \boldsymbol{\lambda}_0) > 0$ for all $r = k + 1, \dots, n$
⇒ \mathbf{x}_0 is a *local minimum*
- ▶ $(-1)^r B_r(\mathbf{x}_0; \boldsymbol{\lambda}_0) > 0$ for all $r = k + 1, \dots, n$
⇒ \mathbf{x}_0 is a *local maximum*

(n is the number of variables x_i and k is the number of constraints.)

Example – Sufficient Condition for Local Extrema

Compute all extrema of $f(x_1, x_2, x_3) = (x_1 - 1)^2 + (x_2 - 2)^2 + 2x_3^2$
subject to constraints $x_1 + 2x_2 = 2$ and $x_2 - x_3 = 3$

Lagrange Function:

$$\mathcal{L}(x_1, x_2, x_3; \lambda_1, \lambda_2) = ((x_1 - 1)^2 + (x_2 - 2)^2 + 2x_3^2) + \lambda_1(2 - x_1 - 2x_2) + \lambda_2(3 - x_2 + x_3)$$

Critical point of \mathcal{L} :

$$x_1 = -\frac{6}{7}, x_2 = \frac{10}{7}, x_3 = -\frac{11}{7}; \lambda_1 = -\frac{26}{7}, \lambda_2 = \frac{44}{7}.$$

Example – Sufficient Condition for Local Extrema

Bordered Hessian matrix:

$$\bar{\mathbf{H}}(\mathbf{x}; \boldsymbol{\lambda}) = \begin{pmatrix} 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 1 & 0 & 2 & 0 & 0 \\ 2 & 1 & 0 & 2 & 0 \\ 0 & -1 & 0 & 0 & 4 \end{pmatrix}$$

3 variables, 2 constraints: $n = 3, k = 2 \Rightarrow r = 3$

$$B_3 = |\bar{\mathbf{H}}(\mathbf{x}; \boldsymbol{\lambda})| = 14$$

$$(-1)^k B_r = (-1)^2 B_3 = 14 > 0 \quad \text{condition satisfied}$$

$$(-1)^r B_r = (-1)^3 B_3 = -14 < 0 \quad \text{not satisfied}$$

Critical point $\mathbf{x}_0 = (-\frac{6}{7}, \frac{10}{7}, -\frac{11}{7})$ is a *local minimum*.

Sufficient Condition for Global Extrema

Let $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ be a critical point of the Lagrange function \mathcal{L} of optimization problem

$$\min / \max f(\mathbf{x}) \quad \text{subject to} \quad \mathbf{g}(\mathbf{x}) = \mathbf{c}$$

If $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}^*)$ is *concave* (convex) in \mathbf{x} , then \mathbf{x}^* is a **global maximum** (global minimum) of $f(\mathbf{x})$ subject to $\mathbf{g}(\mathbf{x}) = \mathbf{c}$.

Example – Sufficient Condition for Global Extrema

$(x^*, y^*, \lambda^*) = (2, 1; 4)$ is a critical point of the Lagrange function \mathcal{L} of optimization problem

$$\min / \max f(x, y) = x^2 + 2y^2 \quad \text{subject to} \quad g(x, y) = x + y = 3$$

Lagrange function:

$$\mathcal{L}(x, y, \lambda^*) = (x^2 + 2y^2) + 4 \cdot (3 - (x + y))$$

Hessian matrix:

$$\mathbf{H}_{\mathcal{L}}(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} \quad \begin{array}{l} H_1 = 2 > 0 \\ H_2 = 8 > 0 \end{array}$$

\mathcal{L} is convex in (x, y) .

Thus $(x^*, y^*) = (2, 1)$ is a global minimum.

Example – Sufficient Condition for Global Extrema

$$(\mathbf{x}^*, \lambda^*) = \left(-\frac{6}{7}, \frac{10}{7}, -\frac{11}{7}; -\frac{26}{7}, \frac{44}{7}\right)$$

is a critical point of the Lagrange function of optimization problem

$$\min / \max f(x_1, x_2, x_3) = (x_1 - 1)^2 + (x_2 - 2)^2 + 2x_3^2$$

$$\text{subject to} \quad g_1(x_1, x_2, x_3) = x_1 + 2x_2 = 2$$

$$g_2(x_1, x_2, x_3) = x_2 - x_3 = 3$$

Lagrange function:

$$\mathcal{L}(\mathbf{x}; \lambda^*) = ((x_1 - 1)^2 + (x_2 - 2)^2 + 2x_3^2) - \frac{26}{7}(2 - x_1 - 2x_2) + \frac{44}{7}(3 - x_2 + x_3)$$

Example – Sufficient Condition for Global Extrema

Hessian matrix:

$$\mathbf{H}_{\mathcal{L}}(x_1, x_2, x_3) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \quad \begin{array}{l} H_1 = 2 > 0 \\ H_2 = 4 > 0 \\ H_3 = 16 > 0 \end{array}$$

\mathcal{L} is convex in \mathbf{x} .

$\mathbf{x}^* = \left(-\frac{6}{7}, \frac{10}{7}, -\frac{11}{7}\right)$ is a global minimum.

Interpretation of Lagrange Multiplier

Extremum \mathbf{x}^* of optimization problem

$$\min / \max f(\mathbf{x}) \text{ subject to } \mathbf{g}(\mathbf{x}) = \mathbf{c}$$

depends on \mathbf{c} , $\mathbf{x}^* = \mathbf{x}^*(\mathbf{c})$, and so does the extremal value

$$f^*(\mathbf{c}) = f(\mathbf{x}^*(\mathbf{c}))$$

How does $f^*(\mathbf{c})$ change with varying \mathbf{c} ?

$$\frac{\partial f^*}{\partial c_j}(\mathbf{c}) = \lambda_j^*(\mathbf{c})$$

That is, Lagrange multiplier λ_j is the derivative of the extremal value w.r.t. exogeneous variable c_j in constraint $g_j(\mathbf{x}) = c_j$.

Proof Idea

Lagrange function \mathcal{L} and objective function f coincide in extremum \mathbf{x}^* .

$$\begin{aligned} \frac{\partial f^*(\mathbf{c})}{\partial c_j} &= \frac{\partial \mathcal{L}(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\lambda}(\mathbf{c}))}{\partial c_j} \quad [\text{chain rule}] \\ &= \sum_{i=1}^n \underbrace{\mathcal{L}_{x_i}(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\lambda}(\mathbf{c}))}_{=0} \cdot \frac{\partial x_i^*(\mathbf{c})}{\partial c_j} + \frac{\partial \mathcal{L}(\mathbf{x}, \mathbf{c})}{\partial c_j} \Big|_{(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\lambda}^*(\mathbf{c}))} \\ &\quad \text{as } \mathbf{x}^* \text{ is a critical point} \\ &= \frac{\partial \mathcal{L}(\mathbf{x}, \mathbf{c})}{\partial c_j} \Big|_{(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\lambda}^*(\mathbf{c}))} \\ &= \frac{\partial}{\partial c_j} \left(f(\mathbf{x}) + \sum_{i=1}^k \lambda_i (c_i - g_i(\mathbf{x})) \right) \Big|_{(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\lambda}^*(\mathbf{c}))} \\ &= \lambda_j^*(\mathbf{c}) \end{aligned}$$

Example – Lagrange Multiplier

$(x^*, y^*) = (2, 1)$ is a minimum of optimization problem

$$\begin{aligned} \min / \max \quad & f(x, y) = x^2 + 2y^2 \\ \text{subject to} \quad & g(x, y) = x + y = c = 3 \end{aligned}$$

with $\lambda^* = 4$.

How does the minimal value $f^*(c)$ change with varying c ?

$$\frac{df^*}{dc} = \lambda^* = 4$$

Envelope Theorem

What is the derivative of the extremal value f^* of optimization problem

$$\min / \max f(\mathbf{x}, \mathbf{p}) \text{ subject to } \mathbf{g}(\mathbf{x}, \mathbf{p}) = \mathbf{c}$$

w.r.t. parameters (exogeneous variables) \mathbf{p} ?

$$\frac{\partial f^*(\mathbf{p})}{\partial p_j} = \frac{\partial \mathcal{L}(\mathbf{x}, \mathbf{p})}{\partial p_j} \Big|_{(\mathbf{x}^*(\mathbf{p}), \lambda^*(\mathbf{p}))}$$

Example – Roy's Identity

Maximize utility function

$$\max U(\mathbf{x}) \text{ subject to } \mathbf{p}^T \cdot \mathbf{x} = w$$

The maximal utility depends on prices \mathbf{p} and income w ab:

$$U^* = U^*(\mathbf{p}, w) \quad [\text{indirect utility function}]$$

Lagrange function $\mathcal{L}(\mathbf{x}, \lambda) = U(\mathbf{x}) + \lambda (w - \mathbf{p}^T \cdot \mathbf{x})$

$$\frac{\partial U^*}{\partial p_j} = \frac{\partial \mathcal{L}}{\partial p_j} = -\lambda^* x_j^* \quad \text{and} \quad \frac{\partial U^*}{\partial w} = \frac{\partial \mathcal{L}}{\partial w} = \lambda^*$$

and thus

$$x_j^* = -\frac{\partial U^* / \partial p_j}{\partial U^* / \partial w} \quad [\text{Marshallian demand function}]$$

Example – Shephard's Lemma

Minimize expenses

$$\min \mathbf{p}^T \cdot \mathbf{x} \text{ subject to } U(\mathbf{x}) = \bar{u}$$

The *expenditure function* (minimal expenses) depend on prices \mathbf{p} and level \bar{u} of utility: $e = e(\mathbf{p}, \bar{u})$

Lagrange function $\mathcal{L}(\mathbf{x}, \lambda) = \mathbf{p}^T \cdot \mathbf{x} + \lambda (\bar{u} - U(\mathbf{x}))$

$$\frac{\partial e}{\partial p_j} = \frac{\partial \mathcal{L}}{\partial p_j} = x_j^* \quad [\text{Hicksian demand function}]$$

Summary

- ▶ constraint optimization
- ▶ graphical solution
- ▶ Lagrange function and Lagrange multiplier
- ▶ extremum and critical point
- ▶ bordered Hessian matrix
- ▶ global extremum
- ▶ interpretation of Lagrange multiplier
- ▶ envelope theorem