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

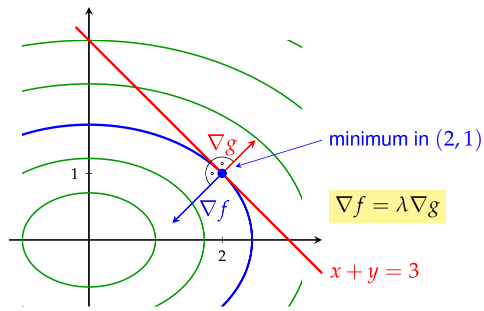
$$g(x,y)=x+y=3$$

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-plain.
   (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 $g(x,y)=x+y=3$

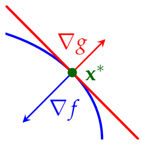
Lagrange Approach

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.

$$\begin{aligned} f_x(\mathbf{x}^*)&=\lambda\,g_x(\mathbf{x}^*)\\ f_y(\mathbf{x}^*)&=\lambda\,g_y(\mathbf{x}^*)\\ g(\mathbf{x}^*)&=c \end{aligned}$$



Transformation yields

$$\begin{aligned} f_x(\mathbf{x}^*)-\lambda\,g_x(\mathbf{x}^*)&=0\\ f_y(\mathbf{x}^*)-\lambda\,g_y(\mathbf{x}^*)&=0\\ c-g(\mathbf{x}^*)&=0 \end{aligned}$$

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 $\mathcal{L}$:

$$\begin{aligned} \mathcal{L}_x&=f_x-\lambda\,g_x&=0\\ \mathcal{L}_y&=f_y-\lambda\,g_y&=0\\ \mathcal{L}_\lambda&=c-g(x,y)&=0 \end{aligned}$$

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 $\mathcal{L}$.

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

Example – Bordered Hessian Matrix

Compute the local extrema of

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

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

$$\begin{aligned} g_1(x_1, \dots, x_n) &= c_1 \\ &\vdots \qquad\qquad (k < n) \\ g_k(x_1, \dots, x_n) &= c_k \end{aligned}$$

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

Lagrange Function:

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

Recipe – Critical Points

1. Create Lagrange Function $\mathcal{L}$:

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

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

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

subject to

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

Lagrange function:

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

Example – Critical Points

Partial derivatives (gradient):

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

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

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

The unique critical point of f subject to these constraints is $\mathbf{x}_0 = (-\tfrac{6}{7}, \tfrac{10}{7}, -\tfrac{11}{7})$.

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 $r = k+1, \dots, n$
  $\Rightarrow$ $\mathbf{x}_0$ is a *local minimum*
- $(-1)^r\, B_r(\mathbf{x}_0; \boldsymbol{\lambda}_0) > 0$ for all $r = k+1, \dots, n$
  $\Rightarrow$ $\mathbf{x}_0$ 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 $f(x_1, x_2, x_3) = (x_1 - 1)^2 + (x_2 - 2)^2 + 2\,x_3^2$
subject to constraints $x_1 + 2\,x_2 = 2$ and $x_2 - x_3 = 3$

Lagrange Function:

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

Critical point of $\mathcal{L}$:

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

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: $n = 3,\ k = 2 \quad \Rightarrow \quad r = 3$

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

$(-1)^k B_r = (-1)^2 B_3 = 14$ > 0 condition satisfied

$(-1)^r B_r = (-1)^3 B_3 = -14$ < 0 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

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

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

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

Lagrange function:

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

Hessian matrix:

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

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

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

Example – Sufficient Condition for Global Extrema

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

is a critical point of the Lagrange function of optimization problem

$$\begin{array}{ll} \min / \max & f(x_1, x_2, x_3) = (x_1 - 1)^2 + (x_2 - 2)^2 + 2\,x_3^2 \\ \text{subject to} & g_1(x_1, x_2, x_3) = x_1 + 2\,x_2 = 2 \\ & g_2(x_1, x_2, x_3) = x_2 - x_3 = 3 \end{array}$$

Lagrange function:

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

Example – Sufficient Condition for Global Extrema

Hessian matrix:

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

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

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

Interpretation of Lagrange Multiplier

Extremum $\mathbf{x}^*$ of optimization problem

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

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 $\mathbf{c}$?

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

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 extemum $\mathbf{x}^*$.

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

Example – Lagrange Multiplier

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

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

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

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

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

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

Example – Roy's Identity

Maximize utility function

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

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

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

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

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

and thus

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

Example – Shephard's Lemma

Minimize expenses

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

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

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

$$\frac{\partial e}{\partial p_j} = \frac{\partial \mathcal{L}}{\partial p_j} = x_j^*$$ [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