

VV285 RC Final

Finding Extrema

Xingjian Zhang

University of Michigan-Shanghai Jiao Tong University Joint Institute

August 2, 2020



JOINT INSTITUTE
交大密西根学院

- Free Extrema
- Constrained Extrema

We investigated how to find the extrema of a potential function without any constraints. In order to determine whether a point is critical, we focus on the second item of Eq.2. Furthermore, to determine whether a critical point is a local maximum/minimum, we focus on the third item of Eq.2.

Quadratic Approximation. Let $\Omega \subset \mathbb{R}^n$ be an open set and $f \in C^2(\Omega, \mathbb{R})$. Then for any $h \in \mathbb{R}^n$ small enough that $x + h \in \Omega$,

$$f(x + h) = f(x) + \langle \nabla f(x), h \rangle + \int_0^1 (1 - t) \langle \text{Hess } f(x + th)h, h \rangle dt. \quad (1)$$

Corollary. Let $\Omega \subset \mathbb{R}^n$ be an open set and $f \in C^2(\Omega, \mathbb{R})$. Then, as $h \rightarrow 0$,

$$f(x + h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \langle \text{Hess } f(x)h, h \rangle + o(h^2). \quad (2)$$

Remark: At the extrema point, the second item of Eq.2 vanishes for all directions. Then the variation trend of the function is dominated by the third item. You can make a comparison of it with what we have learned in VV186.

To better deal with the third item $\langle \text{Hess } f(x)h, h \rangle$, we introduce *quadratic form*.

Let $A \in \text{Mat}(n \times n, \mathbb{R})$. Then the *quadratic form induced by A* is defined as the map

$$Q_A := \langle \cdot, A(\cdot) \rangle, \quad x \mapsto \langle x, Ax \rangle = \sum_{j,k=1}^n a_{jk} x_j x_k, \quad x \in \mathbb{R}^n.$$

Clearly, $Q_A(\lambda x) = \lambda^2 Q_A(x)$ for any $\lambda \in \mathbb{R}$. Note also that Q_A is continuous, because it is a polynomial in x_1, x_2, \dots, x_n .

Definition. A quadratic form Q_A induced by a matrix $A \in \text{Mat}(n \times n, \mathbb{R})$ is called

- ▶ *positive definite* if $Q_A(x) > 0$ for all $x \neq 0$,
- ▶ *negative definite* if $Q_A(x) < 0$ for all $x \neq 0$,
- ▶ *indefinite* if $Q_A(x_0) > 0$ for some $x_0 \in \mathbb{R}^n$ and $Q_A(y_0) < 0$ for some $y_0 \in \mathbb{R}^n$.

A matrix A is said to be negative definite / positive definite / indefinite if the induced quadratic form Q_A has the corresponding property.

Remark:

- ▶ It is easy to see that not all quadratic forms fall into one of the above three categories.
- ▶ If A is positive definite, then $-A$ is negative definite.

A particular case in \mathbb{R}^2 :

Let $A \in \text{Mat}(2 \times 2, \mathbb{R})$ be symmetric, i.e.,

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Let $\Delta = \det A$. Then

- (i) A positive definite $\Rightarrow a > 0$ and $\Delta > 0$
- (ii) A negative definite $\Rightarrow a < 0$ and $\Delta > 0$
- (iii) A indefinite $\Rightarrow \Delta < 0$

Essential Condition. Let $\Omega \subset \mathbb{R}^n$, $f : \Omega \rightarrow \mathbb{R}$ and $\xi \in \text{int } \Omega$. Assume that all partial derivatives of f exist at ξ and that f has a local extremum (maximum or minimum) in ξ . Then

$$\nabla f(\xi) = 0.$$

If f is differentiable in ξ , this implies $Df|_{\xi} = 0$.

Hessian & Extrema Let $\Omega \subset \mathbb{R}^n$ be open, $f \in C^2(\Omega)$ and $\xi \in \Omega$. Let $\nabla f(\xi) = 0$ (i.e., $Df|_{\xi} = 0$).

- (i) If $\text{Hess } f|_{\xi}$ is positive definite, f has a strict local minimum at ξ .
- (ii) If $\text{Hess } f|_{\xi}$ is negative definite, f has a strict local maximum at ξ .
- (iii) If $\text{Hess } f|_{\xi}$ is indefinite, f has no extremum at ξ .

Remark: We often call $\xi \in \text{int } \Omega$ the critical points. Notice that

- ▶ a local extremum can exist at the boundary of Ω without being a critical point. (This situation is tricky, and we did not give a detailed discussion about it in the lecture.)
- ▶ the second theorem gives us a practical way to tell whether a critical point is an extremum.

Critical Point

A critical point is not necessarily an extrema. $(0,0)$ is critical point for both of the functions below. However, it is merely a *saddle point*.

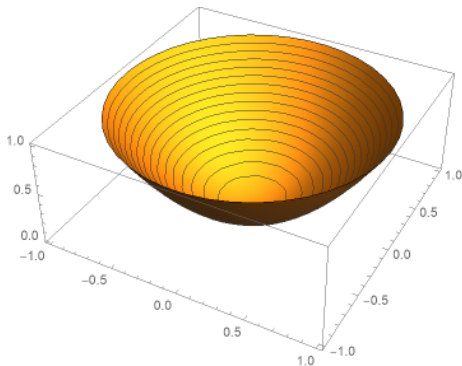


Figure: $z = x^2 + y^2$

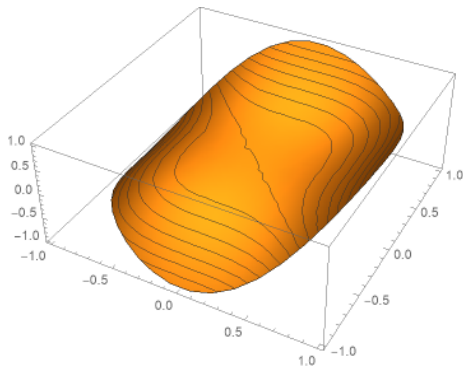


Figure: $z = x^3 + y^3$

Four Steps to Find Free Extrema



We follow a four-step process to find the extrema of a potential function $f \in \mathbb{C}^2(\Omega, \mathbb{R})$.

1. Check for critical points $\xi \in \text{int } \Omega$, i.e., those where $Df|_{\xi} = 0$.
2. Use definiteness of Hessian to check which of the critical points is an extremum.
3. Check the boundary $\partial\Omega$ separately for local extrema.
4. Identify the global extrema. Any finite global extremum must also be a local extremum, so it will be included among those found above.

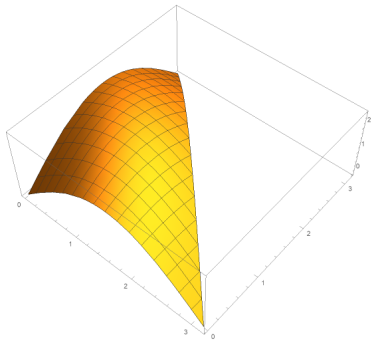
Remark: For step.3, we are not going to clarify all the local extremum at the boundary. Instead, we will merely find all the possible candidates that might be local extremum – directly calculate their corresponding function values – and compare them with local extremum inside the region. This is due to the difficulty of finding local extremum at the boundary.

Follow the four-step process, find all global extrema (if they exist) of the following real function on its domain.

$$\text{dom } f = \{(x, y) \in \mathbb{R}^2 : 0 \leq x, y, x + y \leq \pi\}, f(x, y) = \sin x + \sin y + \sin(x + y)$$

Follow the four-step process, find all global extrema (if they exist) of the following real function on its domain.

$$\text{dom } f = \{(x, y) \in \mathbb{R}^2 : 0 \leq x, y, x + y \leq \pi\}, f(x, y) = \sin x + \sin y + \sin(x + y)$$



Lagrange Multiplier Rule I



Let $\Omega \subset \mathbb{R}^n$ be open and $f \in C^1(\Omega, \mathbb{R})$, $g \in C^1(\Omega, \mathbb{R}^m)$, $m < n$. Assume that f has a local extremum on the set $E = \{x \in \mathbb{R}^n : g(x) = 0\}$ at $\xi \in E$. Assume further that in

$$Dg|_{\xi} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n} \end{pmatrix}$$

there exists a submatrix consisting of m columns whose determinant does not vanish. Then there exist m numbers (called Lagrange multipliers)

$\lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{R}$ such that

$$Df|_{\xi} + \sum_{i=1}^m \lambda_i Dg_i|_{\xi} = 0. \quad (3)$$

Remark: You need to show the bold pre-condition holds in your solution. Otherwise there will be deduction.

Lagrange Multiplier Rule II



In order to find the constrained extrema, we need to solve the $m + n$ equations

$$\begin{aligned}\frac{\partial f}{\partial x_i} + \lambda_1 \frac{\partial g_1}{\partial x_1} + \cdots + \lambda_m \frac{\partial g_m}{\partial x_i} &= 0, & i &= 1, \dots, n, \\ g_j(x) &= 0, & j &= 1, \dots, m.\end{aligned}$$

These equations are equivalent to the following: define

$$F(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) := f(x) + \lambda_1 g_1(x) + \cdots + \lambda_m g_m(x).$$

Then at an extremal point all partial derivatives of F will vanish.

Remark: The conditions above are easier to remember.

Find the extrema of the following, if they exist

$$f(x, y, z) = xyz \quad \text{s.t.} \quad x^2 + 2y^2 + 3z^2 = 6$$

$$f(x, y, z) = ye^{x-z} \quad \text{s.t.} \quad 9x^2 + 4y^2 + 36z^2 = 36 \text{ and } xy + yz = 1$$

One of the characteristic properties of constrained extrema problems is that it is often not necessary to calculate the second derivative to determine the nature of the extremum. This is the case, because in most applications one of two situations occur:

1. The constraint set $E = \{g(x) = 0\}$ is compact. Then there must exist a maximum and a minimum by Theorem 2.1.36.
2. The problem is one of finding the distance between a compact and a closed set. Then the distance between the sets is found by minimizing the distances between all points and this minimum exists.

In both cases, the strategy is to find all candidates for extrema points and evaluate the values of the function on these points. The largest value will be the maximum, the smallest value will be the minimum.

Prove that the triangle with maximum area A that has a given perimeter p is equilateral.

You can use Heron's formula where $s = p/2$

$$A(x, y, z) = \sqrt{s(s-x)(s-y)(s-z)}$$

Remark: How to prove it strictly? Pay attention to the preconditions we discussed just now.

Here are some general tips:

- ▶ Always pay attention to the preconditions when using Lagrange multiplier rule.
- ▶ Always show your effort to calculate something.
- ▶ Never directly give a conclusion without convincing explanation.
- ▶ Keep calm and believe yourself. All exercises are easy if you know what you are doing and hard if you don't.
- ▶ Have fun!