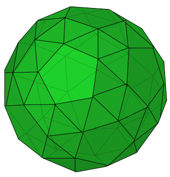
Projects in Mathematics and Applications

# GIẢI TÍCH

Trong Duong*
Linh Tran (Editor)†

23/07/2018



* International University, VNU, Vietnam
† National University of Singapore, Singapore

# Mục lục


**1 Giới hạn và hội tụ** 1
- 1.1 Giới thiệu giới hạn hàm số . . . 1
- 1.2 Giới hạn một bên và giới hạn tại vô cùng . . . 2
- 1.3 Tính liên tục của hàm số . . . 4
- 1.4 Định lý giá trị trung gian . . . 4

**2 Đạo hàm và cực trị hàm số** 6
- 2.1 Đạo hàm . . . 6
- 2.2 Các quy tắc tính đạo hàm - Chain Rule . . . 6
- 2.3 Tính tăng giảm của hàm số . . . 7
- 2.4 Cực đại và cực tiểu . . . 8
- 2.5 Tính lồi lõm và đạo hàm bậc hai . . . 10

**3 Nguyên hàm và Tích phân** 12
- 3.1 Tích phân - Diện tích phía dưới đồ thị . . . 12
- 3.2 Nguyên hàm - Định lý cơ bản của giải tích . . . 13
- 3.3 Nguyên hàm dưới dạng tích phân không xác định . . . 14
- 3.4 Một số ứng dụng của tích phân . . . 15

**4 Hàm số nhiều biến** 18
- 4.1 Hàm số nhiều biến . . . 18
- 4.2 Giới hạn, tính liên tục của hàm nhiều biến . . . 19

**5 Đạo hàm của hàm nhiều biến** 20
- 5.1 Đạo hàm riêng của hàm số thực . . . 20
- 5.2 Quy tắc mắt xích cho hàm nhiều biến . . . 20
- 5.3 Đạo hàm có hướng (directional derivative) - Gradient . . . 22
- 5.4 Đạo hàm của hàm nhận giá trị vector - Ma trận Jacobi . . . 23
- 5.5 Đạo hàm bậc hai - Ma trận Hessian . . . 24
- 5.6 Thứ tự của đạo hàm riêng bậc 2 - Định lý Schwarz . . . 26

**6 Cực trị hàm nhiều biến** 28
- 6.1 Khái niệm . . . 28
- 6.2 Định lý đạo hàm bậc nhất . . . 29
- 6.3 Định lý đạo hàm bậc 2 . . . 30

## Danh pháp tiếng Anh

1. Giải tích: Analysis/Calculus
2. Biến: Variable
3. Hàm số: Function
4. Liên tục: Continuous
5. Miền xác định: Domain
6. Miền giá trị: Image/Codomain
7. Số thực: Real/Real number
8. Vô cực: Infinity
9. Hội tụ/Tiến tới: Converge
10. Tổng: Sum
11. Tích: Product
12. Thương: Quotient/Ratio
13. Lũy thừa: Power
14. Giới hạn: Limit
15. Đạo hàm: Derivative
16. Hàm hợp: Composite
17. Quy tắc mắt xích: Chain rule
18. Nguyên hàm: Anti-derivative
19. Tích phân: Integral
20. Diện tích: Area
21. Đồ thị: Graph
22. Nhiều biến: Multi-variable
23. Đạo hàm riêng: Partial derivative
24. Đạo hàm có hướng: Directional derivative
25. Điểm cực đại/cực tiểu toàn cục: Global maxima/minima
26. Điểm cực đại/cực tiểu địa phương: Local maxima/minima
27. Cực trị toàn cục/địa phương: Local/Global extremum (số nhiều extrema)
28. Điểm yên ngựa: Saddle point

# 1 Giới hạn và hội tụ

Khái niệm giới hạn hàm số và sự hội tụ là trung tâm của ngành Giải Tích, cả cổ điển lẫn hiện đại. Là một trong những khái niệm nền móng của Giải Tích, giới hạn được ứng dụng nhiều trong việc "liên tục hóa" hàm số, xấp xỉ hàm số (nội suy) và khảo sát hàm số.

## 1.1 Giới thiệu giới hạn hàm số

Nếu $f$ là một hàm số từ $\mathbb{R} \to \mathbb{R}$ thì

$$\lim_{x\to a} f(x) = L,$$

gọi là "Giới hạn của $f(x)$ khi $x$ tiến tới $a$ bằng $L$". Điều đó có nghĩa rằng nếu ta chọn các giá trị $x$ rất gần nhưng không bằng $a$, thì $f(x)$ sẽ rất gần với giá trị $L$; ngoài ra, $f(x)$ càng lúc càng gần $L$ hơn nếu $x$ càng lúc càng gần $a$.

**Ví dụ 1.1.** Xét $f(x) = x + 3$. Khi đó:

$$\lim_{x\to 4} f(x) = 7.$$

Vì nếu ta thay các giá trị $x$ gần với 4 vào $f(x) = x + 3$ sẽ thu được các giá trị gần với 7.

Tuy nhiên không phải khi nào phép thử trực tiếp cũng mang lại kết quả.

**Ví dụ 1.2.** Tính giới hạn

$$\lim_{x\to 2} \frac{x^2 - 2x}{x^2 - 4}.$$

Thử thay $x = 2$ vào biểu thức thu được

$$f(2) = \frac{2^2 - 2\times 2}{2^2 - 4} = \frac{0}{0},$$

một tỷ số không có nghĩa. Trong khi đó, nếu thử thay các giá trị $x$ gần nhưng không bằng 2, ta thu được bảng 1 và nhận thấy $f(x)$ tiến dần đến 0.5.

| $x$ | $f(x)$ | $x$ | $f(x)$ |
|---|---|---|---|
| 3.000000 | 0.600000 | 1.999000 | 0.499875 |
| 2.500000 | 0.555556 | 1.990000 | 0.498747 |
| 2.100000 | 0.512195 | 1.900000 | 0.487179 |
| 2.010000 | 0.501247 | 1.500000 | 0.428571 |
| 2.001000 | 0.500125 | 1.000000 | 0.333333 |

**Bảng 1:** Tìm giới hạn bằng cách thế giá trị $x$ tiến gần đến $a$.

Rõ ràng không phải lúc nào giới hạn của một hàm số tại một điểm cũng tồn tại, và không phải lúc nào cũng tìm được bằng cách thử trực tiếp. Để xây dựng các lý thuyết khảo sát sự tồn tại và xác định giới hạn, ta cần một định nghĩa chặt chẽ về Toán học cho khái niệm này.

**Định nghĩa 1.3** (Giới hạn của hàm số)**.** Cho hàm số $y = f(x)$ xác định trong một lân cận của $a$, số thực $L$ hữu hạn được gọi là **giới hạn** của hàm số $f(x)$ khi $x \to a$ nếu:

$$\forall \varepsilon > 0, \exists \delta_\varepsilon > 0 : 0 < |x - a| < \delta_\varepsilon \implies |f(x) - L| < \varepsilon.$$

Để minh họa cho định nghĩa trên, cùng trở lại ví dụ 1.2. Xét biểu thức

$$\left| \frac{x^2 - 2x}{x^2 - 4} - \frac{1}{2} \right| = \left| \frac{x}{x+2} - \frac{1}{2} \right| = \left| \frac{x-2}{x+2} \right|, \quad \forall x \neq 2.$$

Với mọi $\varepsilon > 0$, chọn $\delta = \min\{2\varepsilon, 2\}$, khi đó:

$$0 < |x - 2| < \delta \implies |x + 2| \geq 4 - |x - 2| > 4 - \delta \geq 2.$$

Do đó

$$\left| \frac{x-2}{x+2} \right| < \frac{|x-2|}{2} \leq \frac{2\varepsilon}{2} = \varepsilon.$$

Như vậy theo định nghĩa 1.3, giới hạn của hàm số $f(x) = \dfrac{x^2 - 2x}{x^2 - 4}$ tại $x = 2$ là $\dfrac{1}{2}$.

Mặc dù định nghĩa 1.3 cho ta biết giới hạn là gì, nó không cho biết làm sao để tìm được giới hạn. **Liệu có xảy ra trường hợp có nhiều giới hạn khác nhau?** Định lý sau khẳng định giới hạn nếu tồn tại là duy nhất.

**Định lý 1.4.** *Nếu $L$ và $L'$ đều là giới hạn của hàm số $f$ khi $x \to a$ thì $L = L'$.*

**Lưu ý.** Các cách gọi sau là tương đương:

- ⋆ $L$ là giới hạn của hàm số $f(x)$ khi $x \to a$ ($x$ tiến tới $a$).
- ⋆ $L$ là giới hạn của hàm số $f(x)$ tại $x = a$.
- ⋆ $L$ là giới hạn của hàm số $f(x)$ tại điểm $a$.
- ⋆ $f(x)$ tiến tới $L$ ($f(x) \to L$) khi $x$ tiến tới $a$ ($x \to a$).

## 1.2 Giới hạn một bên và giới hạn tại vô cùng

Trong bảng 1, có thể thấy $f(x)$ tiến gần tới $1/2$ cả khi $x$ tiến gần tới 2 từ phía bên trái ($x < 2$) và bên phải ($x > 2$). Nếu $f(x)$ chỉ tiến gần tới $1/2$ khi $x$ tiếp cận 2 từ một bên thì giới hạn không tồn tại.

Trong nhiều trường hợp, ta chỉ quan tâm đến giá trị của $f(x)$ khi $x$ tiến đến $a$ từ một bên, chẳng hạn như khi $f$ chỉ xác định trên khoảng $(a, b)$. Khi đó việc khẳng định giới hạn không tồn tại không thể giúp ta kết luận gì về giá trị của $f(x)$. Do vậy khái niệm **giới hạn một bên** được ra đời.

**Định nghĩa 1.5** (Giới hạn một bên)**.** Cho hàm số $f$ nhận giá trị trong $\mathbb{R}$. Ta có:

1. Nếu $f$ xác định trên $(a-d, a)$ với $d>0$ nào đó và tồn tại $L \in \mathbb{R}$ sao cho

$$\forall \varepsilon > 0, \exists \delta_\varepsilon > 0 : a-\varepsilon < x < a \implies |f(x)-L| < \varepsilon,$$

thì $L$ được gọi là **giới hạn bên trái** của $f(x)$ tại $a$, ký hiệu $L = \lim\limits_{x \to a^-} f(x)$.

2. Nếu $f$ xác định trên $(a, a+d)$ với $d>0$ nào đó và tồn tại $L \in \mathbb{R}$ sao cho

$$\forall \varepsilon > 0, \exists \delta_\varepsilon > 0 : a < x < a+\varepsilon \implies |f(x)-L| < \varepsilon,$$

thì $L$ được gọi là **giới hạn bên phải** của $f(x)$ tại $a$, ký hiệu $L = \lim\limits_{x \to a^+} f(x)$.

Tương tự như giới hạn thông thường, các giới hạn một bên là duy nhất. Hơn nữa, khi chúng bằng nhau, giới hạn thực sự tồn tại và bằng giá trị chung của giới hạn 2 bên.

**Tính chất 1.6.** *Với hàm số f xác định trên một lân cận của a, ta có*

$$\exists L \in \mathbb{R} : L = \lim_{x \to a} f(x) \iff \begin{cases} \exists L^- : L^- = \lim\limits_{x \to a^-} f(x) \\ \exists L^+ : L^+ = \lim\limits_{x \to a^+} f(x) \\ L^- = L^+ \end{cases}.$$

Trong khảo sát hàm số, nhiều khi ta không chỉ quan tâm đến giá trị của hàm tại một điểm hữu hạn mà còn tại 2 điểm vô cực. Ta mở rộng định nghĩa giới hạn như sau:

**Định nghĩa 1.7** (Giới hạn tại vô cùng)**.** Cho $f(x)$ là một hàm số xác định trên khoảng $(x_0, \infty)$. Nếu tồn tại một số thực $L$ sao cho

$$\forall \varepsilon > 0 : \exists A : \forall x : x > A \implies |f(x)-L| < \varepsilon.$$

Khi đó ta gọi giới hạn của $f(x)$ khi $x \to +\infty$ là L. Ký hiệu $L = \lim\limits_{x \to +\infty} f(x)$.

**Nhận xét 1.8.** Có thể thay đổi định nghĩa 1.7 một chút để có định nghĩa giới hạn tại âm vô cùng.

**Ví dụ 1.9** (Hàm dấu)**.** Hàm dấu

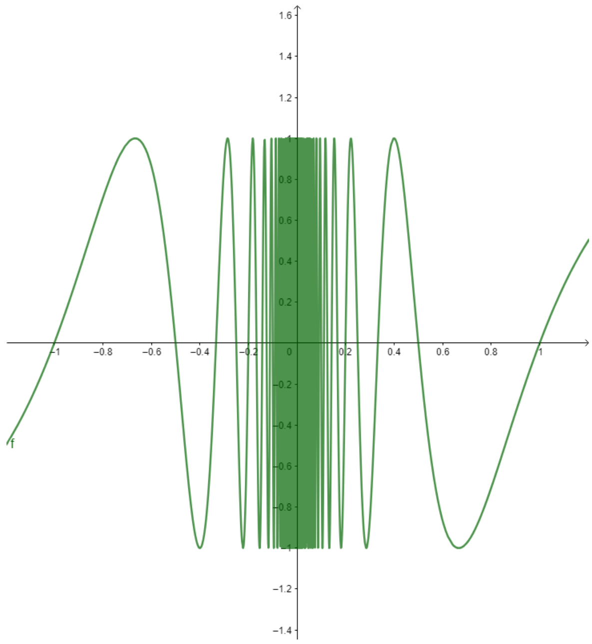
$$\operatorname{sgn}(x) = \begin{cases} \dfrac{x}{|x|} & \text{khi } x \neq 0 \\ 0 & \text{khi } x = 0 \end{cases}.$$

Dễ thấy $\lim\limits_{x \to 0^-} \operatorname{sgn}(x) = -1$ trong khi $\lim\limits_{x \to 0^+} \operatorname{sgn}(x) = 1$, do đó $\lim\limits_{x \to 0} \operatorname{sgn}(x)$ không tồn tại.

**Ví dụ 1.10.** Xét hàm số $f(x) = \sin \dfrac{\pi}{x}$. Dễ thấy:

$$\lim_{x \to -\infty} f(x) = \lim_{x \to +\infty} f(x) = \sin 0 = 0.$$

Nhưng $\lim\limits_{x \to 0} f(x)$ không tồn tại.



**Hình 1:** Đồ thị hàm số $y = \sin\left(\frac{\pi}{x}\right)$.

Thực chất, $\lim_{x\to 0^-} f(x)$ và $\lim_{x\to 0^+} f(x)$ đều không tồn tại!

### 1.3 Tính liên tục của hàm số

Cũng như giới hạn, khái niệm liên tục giữ vai trò trung tâm trong cả Giải Tích cổ điển lẫn hiện đại. Ta chỉ quan tâm đến sự liên tục của hàm thực một biến.

**Định nghĩa 1.11** (Hàm số liên tục)**.** Cho hàm số $f$ xác định trên một khoảng lân cận của $a$.

1. Hàm số $f$ liên tục tại $x = a$ khi $\lim_{x\to a} f(x) = f(a)$.
2. $f$ liên tục trên một khoảng/đoạn nếu nó liên tục tại tất cả các điểm trong khoảng/đoạn đó.

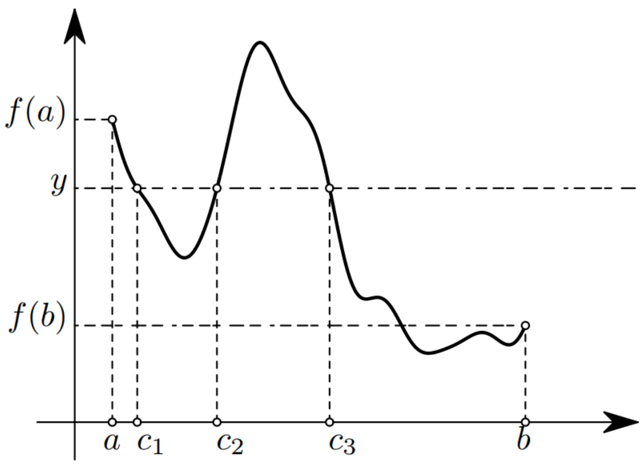
**Ví dụ 1.12.** Hàm số $f(x) = x^2 + 2x + 1$ liên tục trên $\mathbb{R}$.

### 1.4 Định lý giá trị trung gian

**Định lý 1.13** (Định lý giá trị trung gian)**.** *Nếu $f$ là một hàm số liên tục trên khoảng $[a, b]$ và nếu $y$ là một số nằm giữa $f(a)$ và $f(b)$ thì có một số $c$ với $a \leq c \leq b$ thỏa mãn $f(c) = y$.*

**Nhận xét 1.14.** Định lý này giải thích cho cái tên "liên tục", một hàm số $f$ liên tục trên $[a, b]$ khi đồ thị của nó có thể vẽ bằng một nét liên tục từ $(a, f(a))$ đến $(b, f(b))$.

**Nhận xét 1.15.** Nếu $f$ là một hàm số liên tục trên khoảng $[a, b]$ và nếu $f(x) \neq 0$ với mọi x thuộc khoảng này thì $f(x)$ hoặc dương với mọi $a < x < b$ hoặc âm với mọi $a < x < b$.




**Hình 2:** Định lý giá trị trung gian phát biểu rằng một hàm số liên tục phải đạt được bất kì giá trị y nào nằm giữa $f(a)$ và $f(b)$ tối thiểu một lần. Trong ví dụ này có ba giá trị $c$ mà $f(c) = y$.

**Bài tập**

**Bài tập 1.1.** Tính các giới hạn sau:

(a) $\lim\limits_{x\to 2} \dfrac{\sqrt{x}-\sqrt{2}}{x-2}$.

(b) $\lim\limits_{x\to\infty} \dfrac{3x^2+3}{5x^2+7x-39}$.

(c) $\lim\limits_{\theta\to 0} \dfrac{\sin\theta}{\theta}$.

(d) $\lim\limits_{\theta\to 0} \dfrac{1-\cos\theta}{\theta^2}$.

**Bài tập 1.2.** Các hàm số sau có liên tục không?

(a) $f(x) = \dfrac{1}{x}$ trên $[-1, 1]$.

(b) $f(x) = \begin{cases} e^{-1/x^2} & \text{khi } x \neq 0 \\ 0 & \text{khi } x = 0 \end{cases}$ trên $\mathbb{R}$.

# 2 Đạo hàm và cực trị hàm số

Khái niệm **đạo hàm** ra đời khi các nhà toán học quan tâm đến sự tăng giảm và tốc độ tăng giảm của hàm số. Đạo hàm được ứng dụng rộng rãi trong Vật Lý với các đại lượng vận tốc, gia tốc, công, . . .

## 2.1 Đạo hàm

**Định nghĩa 2.1** (Đạo hàm)**.** Cho hàm số $f$ xác định trên một khoảng $(c, d)$ và $c < a < d$. **Đạo hàm** của hàm số $f$ tại $a$ là giá trị của giới hạn sau

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}.$$

$f$ được gọi là **khả vi tại** $a$ nếu giới hạn trên tồn tại.

$f$ được gọi là **khả vi trên khoảng** $(c, d)$ nếu nó khả vi tại mọi điểm $a$ trên khoảng $(c, d)$.

**Nhận xét 2.2.** Ta thay $x = a + h$ vào giới hạn trên và cho $h \to 0$ thay vì $x \to a$ thu được công thức:

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.$$

Nếu thay $a$ bởi $x$ và $h$ bởi $\Delta x$ công thức trên trở thành

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

Đặt $\Delta y = f(x + \Delta x) - f(x)$:

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

Leibniz đề xuất một cách viết mới cho giới hạn này dưới dạng

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

**Ví dụ 2.3.** Xét hàm số $f(x) = x^2 + 2x + 3$ trên $\mathbb{R}$ tại điểm $a = -1$. Ta có $f(-1) = 2$ và

$$\lim_{x \to -1} \frac{f(x) - f(-1)}{x - (-1)} = \lim_{x \to -1} \frac{x^2 + 2x + 1}{x + 1} = \lim_{x \to -1} (x + 1) = 0.$$

Do vậy $f$ khả vi tại $-1$ và $f'(-1) = 0$.

**Nhận xét 2.4.** Nếu hàm số $f$ khả vi tại điểm $a$ thì nó cũng liên tục tại $a$.

## 2.2 Các quy tắc tính đạo hàm - Chain Rule

**Tính chất 2.5** (Công thức đạo hàm cơ bản)**.** *Với mọi hàm số $u, v$ khả vi trên cùng một khoảng, ta có:*

1. *Đạo hàm của hằng số bất kì luôn bằng 0:* $\frac{dc}{dx} = 0$.
2. *Đạo hàm của* $x^a$ *với a là số thực bất kì:* $\frac{dx^a}{dx} = ax^{a-1}$.
3. *Đạo hàm một bội của hàm số:* $(cu)' = cu'$.
4. *Đạo hàm của tổng (hiệu) hai hàm số:* $(u \pm v)' = u' \pm v'$.
5. *Đạo hàm của tích hai hàm số:* $(u \cdot v)' = u' \cdot v + u \cdot v'$.
6. *Đạo hàm của thương hai hàm số:* $\left(\frac{u}{v}\right)' = \frac{u' \cdot v - u \cdot v'}{v^2}$.

Quy tắc mắt xích được dùng để tính đạo hàm của hàm hợp của 2 hay nhiều hàm số.

**Tính chất 2.6** (Đạo hàm của hàm hợp)**.** *Nếu f và g đều khả vi thì $f \circ g$ cũng khả vi. Đạo hàm của $f \circ g$ được cho bởi:*

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x).$$

*Đặt $y = g(x)$ và $z = f(y)$ thì $z = (f \circ g)(x)$. Khi đó đạo hàm của z theo x chính là đạo hàm của $f \circ g$. Đạo hàm của z theo y là đạo hàm của hàm f và đạo hàm của y theo x là đạo hàm của hàm g. Ta có cách biểu diễn ngắn gọn như sau:*

$$\frac{dz}{dx} = (f \circ g)'(x), \ \frac{dz}{dy} = f'(y), \ \frac{dy}{dx} = g'(x).$$

*Quy tắc mắc xích phát biểu một cách ngắn gọn là:*

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}.$$

*Quy tắc này cũng áp dụng cho hàm số là hợp của nhiều hơn 2 hàm.*

**Sneak peek:** Bài tập 2.3 quan trọng để bạn đọc hiểu nền tảng toán học đằng sau thuật toán **Backpropagation** của mô hình **Neural Network**, một trong những mô hình được nghiên cứu và áp dụng nhiều nhất trong Machine Learning.

**Sneak peek:** Tất cả các quy tắc tính đạo hàm ở trên đều được ứng dụng trong các chương trình giải toán thông dụng hiện nay như **MATLAB**, **Maple** hay **Mathematica** (phiên bản online: **Wolfram Alpha**). Các chương trình này được nạp sẵn công thức tính đạo hàm của các hàm sơ cấp cơ bản như đa thức, sin, cos, lũy thừa, log, và dùng các quy tắc trên để tính đạo hàm của những hàm sơ cấp phức tạp hơn.

## 2.3 Tính tăng giảm của hàm số

Như đã đề cập, khảo sát các đạo hàm cho ta biết sự tăng/giảm và tốc độc tăng/giảm của hàm số, do đó nó cũng giúp ta tìm cực trị hàm số. Điều này đặc biệt có ý nghĩa trong các bài toán tối ưu hóa.

**Định lý 2.7.** *Cho hàm số $f$ xác định và liên tục trên $[a, b]$, khả vi trên $(a, b)$. Khi đó:*

1. $f'(x) \geq 0, \quad \forall x \in (a, b) \iff f$ *không giảm trên đoạn* $[a, b]$.
2. $f'(x) > 0, \quad \forall x \in (a, b) \iff f$ *tăng ngặt trên đoạn* $[a, b]$.
3. $f'(x) \leq 0, \quad \forall x \in (a, b) \iff f$ *không tăng trên đoạn* $[a, b]$.
4. $f'(x) < 0, \quad \forall x \in (a, b) \iff f$ *giảm ngặt trên đoạn* $[a, b]$.

**Nhận xét 2.8.** Nếu $f$ khả vi và $f'$ liên tục trên lân cận của $a$ và tại điểm $a$, ta có $f' < 0$ trên $(a - \varepsilon, a)$ và $f' > 0$ trên $(a, a + \varepsilon)$ thì $f'(a) = 0$ và $f$ chuyển từ giảm sang tăng tại $a$.

Một bài toán cổ điển trong khảo sát hàm số là xác định những miền tăng và giảm của hàm số. Nhận xét trên cho thấy bài toán này có thể quy về việc khảo sát dấu của đạo hàm.

**Ví dụ 2.9.** Xác định tính tăng giảm của các hàm số sau:

1. $y = \dfrac{1}{x}$.
2. $y = x^3 - 3x$.

1. Đạo hàm của hàm số $f(x) = 1/x = x^{-1}$ là

$$f'(x) = -\frac{1}{x^2}.$$

Dễ thấy đạo hàm này luôn âm, nhưng hàm này không tăng nếu lấy $a = -1$ và $b = 1$:

$$a < b, \text{ nhưng } \frac{1}{a} = -1 < 1 = \frac{1}{b}!$$

Nguyên nhân là do định lý 2.7 chỉ áp dụng cho hàm số xác định trên một khoảng. Tuy nhiên, $f(x) = \dfrac{1}{x}$ không xác định trên khoảng $-1 < x < 1$ vì nó không xác định tại $x = 0$.

Điều đó dẫn tới việc không thể kết luận $f(x) = 1/x$ tăng từ $x = -1$ đến $x = 1$ Mặt khác, hàm số này xác định và khả vi trên khoảng $0 < x < \infty$, do đó định lý 2.7 nói rằng $f(x) = \dfrac{1}{x}$ giảm với $x > 0$.

2. Xét hàm số $y = f(x) = x^3 - 3x$. Đạo hàm của nó là

$$f'(x) = 3x^2 - 3 = 3(x - 1)(x + 1).$$

Từ đó ta thấy

$$\begin{aligned} f'(x) &> 0 \quad \text{với } x < -1, \\ f'(x) &< 0 \quad \text{với } -1 < x < 1, \\ f'(x) &> 0 \quad \text{với } x > 1. \end{aligned}$$

Do đó hàm $f$ tăng trên $(-\infty, -1)$, giảm trên $(-1, 1)$ và tăng trên $(1, \infty)$.

## 2.4 Cực đại và cực tiểu

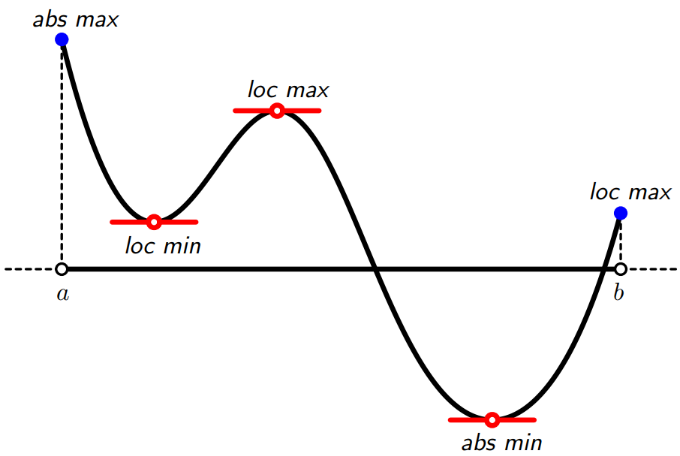
**Định nghĩa 2.10** (Cực đại).

1. Một hàm số có giá trị lớn nhất tại điểm $a$ thuộc miền xác định nếu $f(x) \leq f(a)$ với mọi x thuộc miền xác định của $f$. Khi đó $a$ là **cực đại toàn miền** của $f$.
2. Một hàm số có **cực đại địa phương** tại điểm $a$ thuộc miền xác định nếu có một giá trị nhỏ $\delta > 0$ thỏa mãn $f(x) \leq f(a)$ với mọi x với $a - \delta < x < a + \delta$ thuộc miền xác định của $f$.
3. Thay dấu $\leq$ bằng dấu $\geq$ trong hai định nghĩa trên, ta được tương ứng định nghĩa của **cực tiểu toàn miền** và **cực tiểu địa phương**.

Trong nhận xét 2.8, điểm $a$ thực chất là một cực đại địa phương. Có thể thấy việc tìm cực đại địa phương có thể thông qua giải phương trình $f'(x) = 0$.

**Định nghĩa 2.11** (Điểm dừng)**.** Bất kì giá trị $x$ nào mà $f'(x) = 0$ đều được gọi là một **điểm dừng** (stationary point) của hàm số $f$.

**Định lý 2.12.** *Giả sử $f$ là một hàm khả vi trên đoạn $[a, b]$. Mọi cực đại hoặc cực tiểu địa phương của $f$ hoặc là một đầu mút của đoạn $[a, b]$, hoặc là một điểm cố định của hàm số $f$.*




**Hình 3:** Một hàm số xác định trên đoạn $[a, b]$ với một giá trị nhỏ nhất, một cực tiểu và một cực đại ở trong, và hai cực đại tại hai đầu mút, một trong chúng thực chất là giá trị lớn nhất.

Nếu $f'(c) = 0$ thì $c$ là một điểm cố định, và nó có thể là cực đại hoặc cực tiểu địa phương. Ta có thể xác định tính chất của $c$ bằng cách nhận xét dấu của $f'(x)$ với $x$ lân cận $c$.

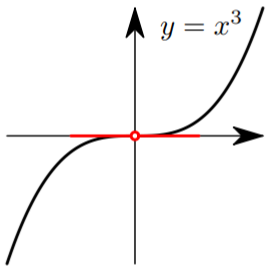
**Định lý 2.13** (Cách xác định cực trị địa phương)**.**

- ⋆ *Nếu trong một khoảng nhỏ $(c - \delta, c + \delta)$ ta có $f'(x) < 0$ với $x < c$ và $f'(x) > 0$ với $x > c$ thì $f$ có một cực tiểu địa phương tại $x = c$.*
- ⋆ *Nếu trong một khoảng nhỏ $(c - \delta, c + \delta)$ ta có $f'(x) > 0$ với $x < c$ và $f'(x) < 0$ với $x > c$ thì $f$ có một cực đại địa phương tại $x = c$.*

**Nhận xét 2.14.** Định lý 2.13 chỉ có thể áp dụng khi có sự **đổi dấu** của đạo hàm khi đi qua điểm $c$. Một điểm $c$ thỏa mãn $f'(c) = 0$ nhưng không có sự đổi dấu thì chưa thể khẳng định là cực trị địa phương!

**Ví dụ 2.15.** Hàm số $f(x) = x^3$ chỉ có duy nhất một điểm cố định $x = 0$. Đạo hàm $f'(x) = 3x^2$ không đổi dấu tại $x = 0$, do đó chưa thể xác định đây có phải cực trị hay không.

Thực chất, $x = 0$ không phải là cực đại lẫn cực tiểu vì $f(x) < f(0)$ với $x < 0$ và $f(x) > f(0)$ với $x > 0$.




**Định nghĩa 2.16** (Điểm yên ngựa)**.** Một điểm $c$ thỏa mãn $f'(c) = 0$ nhưng không phải là cực trị địa phương của $f$ được gọi là một **điểm yên ngựa** (saddle point) của $f$.

## 2.5 Tính lồi lõm và đạo hàm bậc hai

**Định nghĩa 2.17** (Tính lồi lõm)**.** Cho hàm số $f$ liên tục trên $[a, b]$.

1. Một hàm $f$ lồi trên khoảng $(a, b)$ nếu đoạn thẳng nối một cặp điểm bất kì trên đồ thị nằm phía trên phần đồ thị giữa hai điểm đó.
2. Một hàm $f$ lõm trên khoảng $(a, b)$ nếu đoạn thẳng nối một cặp điểm bất kì trên đồ thị nằm phía dưới phần đồ thị giữa hai điểm đó.

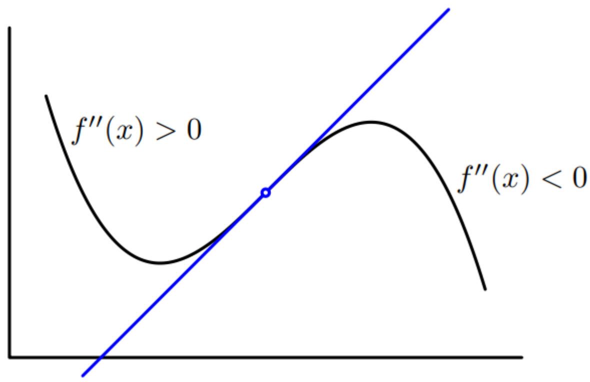

**Hình 4:** Nếu một hàm số là lồi thì tất cả dây cung phải nằm phía trên đồ thi. Nếu nó không lồi thì một số dây cung sẽ cắt hoặc nằm phía dưới đồ thị.

Đối với lớp hàm số khả vi bậc hai (đa số các hàm sơ cấp thường dùng), định lý sau giúp việc xác định tính lồi lõm trở nên dễ dàng.

**Định lý 2.18.** *Cho hàm số $f$ liên tục trên $[a, b]$ và khả vi bậc 2 trên $(a, b)$.*

1. *Một hàm số $f$ lồi trên khoảng $(a, b)$ khi và chỉ khi $f''(x) \geq 0, \quad \forall x \in (a, b)$.*
2. *Một hàm số $f$ lõm trên khoảng $(a, b)$ khi và chỉ khi $f''(x) \leq 0, \quad \forall x \in (a, b)$.*

**Định nghĩa 2.19** (Điểm uốn)**.** Điểm trên đồ thị hàm số $f$ mà tại đó $f''(x)$ đổi dấu được gọi là **điểm uốn** (inflection point). Như vậy nếu $c$ là điểm uốn thì $f''(c) = 0$ (không có chiều ngược lại).




**Hình 5:** Tiếp tuyến cắt đồ thị tại một điểm uốn.

**Định lý 2.20** (Cách xác định cực trị)**.** *Cho hàm $f$ liên tục trên $[a, b]$, khả vi bậc 2 trên $(a, b)$.*

1. *Nếu $c$ là một điểm cố định của hàm số $f$ và $f''(c) < 0$ thì $f$ đạt cực đại tại $x = c$.*
2. *Nếu $c$ là một điểm cố định của hàm số $f$ và $f''(c) > 0$ thì $f$ đại cực tiểu tại $x = c$.*
3. *Không có kết luận nào trong trường hợp $f''(c) = 0$.*

**Bài tập**

**Bài tập 2.1.** Chứng minh toàn bộ tính chất phía trên, sử dụng định nghĩa của đạo hàm.

**Bài tập 2.2.** Phát biểu và chứng minh công thức đạo hàm của tính $n$ hàm số:

$$F(x) = f_1(x) \cdot f_2(x) \cdot \ldots \cdot f_n(x).$$

**Bài tập 2.3.** Phát biểu và chứng minh quy tắc mắt xích cho hàm hợp của $n$ hàm số:

$$F(x) = f_1(f_2(\cdots(f_n(x)))\cdots).$$

**Bài tập 2.4.** Tính đạo hàm hàm số

$$f(x) = \sqrt[4]{1 - x^4}, \quad -1 \leq x \leq 1.$$

**Bài tập 2.5.** Người ta đang bơm nước vào một quả bóng hình cầu có bán kính và thể tích lần lượt là $R$ và $V$. Giả sử nước được bơm vào bóng sao cho mỗi giây bán kính quả bóng tăng thêm $4cm$. Hỏi tốc độ gia tăng thể tích của quả bóng khi bán kính quả bóng đạt $20cm$.

**Bài tập 2.6.** Khảo sát và vẽ đồ thị hàm số

$$f(x) = \frac{x(1 - x)}{1 + x^2}.$$

**Bài tập 2.7.** Tìm hình chữ nhật có diện tích lớn nhất, biết chu vi của nó không đổi và bằng 1.

# 3 Nguyên hàm và Tích phân

## 3.1 Tích phân - Diện tích phía dưới đồ thị

Để hiểu rõ ý nghĩa tích phân, ta bắt đầu với bài toán tính diện tích phía dưới đồ thị.

Cho $f$ là một hàm số xác định trên đoạn $[a, b]$ và mang giá trị dương, nói một cách khác đồ thị hàm số nằm phía trên trục $x$. Tính diện tích phần diện tích của vùng nằm giữa trục hoành, đồ thị hàm số $y = f(x)$ và các trục tung $y = a$ và $y = b$. Ta có thể tính diện tích này bằng cách xấp xỉ diện tích của nhiều hình chữ nhật nhỏ trên đoạn $[a, b]$ bằng cách chọn các số $x_1 < \ldots < x_n$ sao cho

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b.$$

Những số này chia đoạn $[a, b]$ thành $n$ đoạn nhỏ

$$[x_0, x_1],\ [x_1, x_2],\ \ldots,\ [x_{n-1}, x_n],$$

với độ dài là

$$\Delta x_1 = x_1 - x_0,\ \Delta x_2 = x_2 - x_1,\ \ldots,\ \Delta x_n = x_n - x_{n-1}$$

Trong mỗi đoạn ta chọn một điểm $c_k$: trong đoạn đầu tiên ta chọn $x_0 \le c_1 \le x_1$, trong đoạn thứ hai ta chọn $x_1 \le c_2 \le x_2$,..., trong đoạn con cuối cùng ta chọn $x_{n-1} \le c_n \le x_n$.

Sau đó ta xác định $n$ hình chữ nhật: hình chữ nhật thứ $k$ có đáy là đoạn $[x_{k-1}, x_k]$ trên trục hoành, chiều cao là $f(c_k)$, với $k = 1, 2, \ldots, n$.

Diện tích của hình chữ nhật thứ $k$ là tích của chiều cao và độ dài đáy, tức là $f(c_k)\Delta x_k$. Tổng diện tích của tất cả hình chữ nhật là:

$$R = f(c_1)\Delta x_1 + f(c_2)\Delta x_2 + \ldots + f(c_n)\Delta x_n.$$

Tổng này được gọi là **Tổng Riemann** (**Riemann sum**).

Bằng một số cách phân chia hợp lý (chia khoảng, $n$ lớn hơn) có thể cải thiện xấp xỉ này.

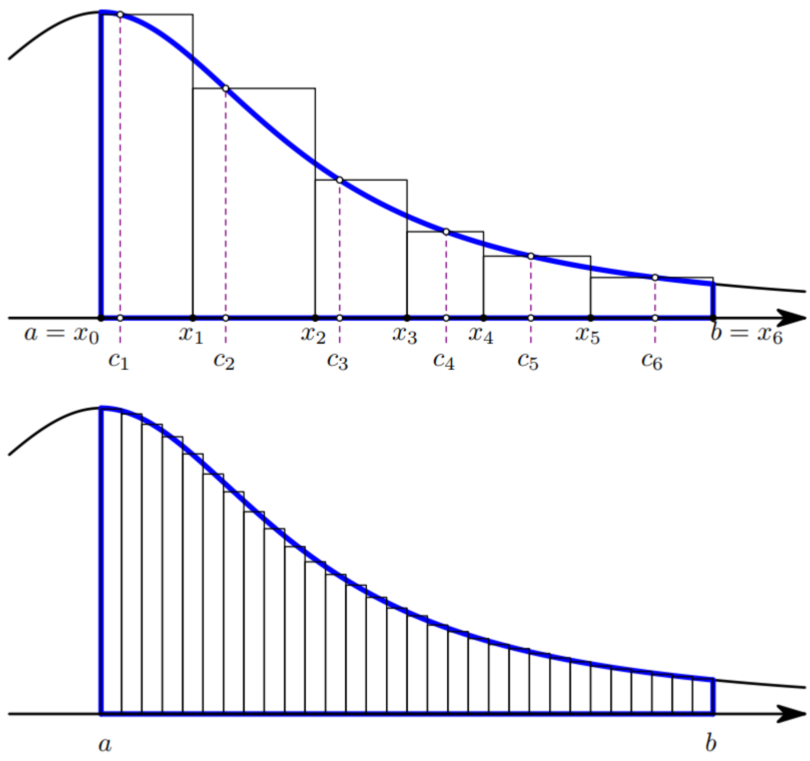
**Định nghĩa 3.1** (Tích phân)**.** Nếu $f$ là một hàm số xác định trên đoạn $[a, b]$, ta phát biểu rằng

$$\int_b^a f(x)dx = I,$$

nghĩa là tích phân "$f(x)$ với $x$ từ $a$ đến $b$" bằng I, nếu với mỗi $\varepsilon > 0$ có thể tìm một $\delta > 0$ thỏa mãn

$$|f(c_1)\Delta x_1 + f(c_2)\Delta x_2 + \ldots + f(c_n)\Delta x_n - I| < \varepsilon,$$

với mọi cách chia mà tất cả các đoạn có độ dài $\delta x_k < \delta$.




**Hình 6:** Tính diện tích miền giới hạn bằng tổng Riemann

Có thể giá trị tích phân từ $a$ đến $b$ của $f$ như một giới hạn của tổng Riemann khi số các hình chữ nhật tiến đến vô cùng và độ dài đáy của chúng tiến đến 0. Do đó, nó cũng là duy nhất:

**Định lý 3.2.** *Nếu $I$ và $I'$ cùng thỏa mãn điều kiện tích phân trong định nghĩa 3.1 thì $I = I'$.*

## 3.2 Nguyên hàm - Định lý cơ bản của giải tích

**Định nghĩa 3.3** (Nguyên hàm)**.** Một hàm số $F$ được gọi là một **nguyên hàm** của $f$ trên khoảng $[a, b]$ nếu $F'(x) = f(x)$ với mỗi $x$ thỏa $a < x < b$.

**Ví dụ 3.4.** $F(x) = \frac{1}{2}x^2$ và $G(x) = \frac{1}{2}x^2 - 2$ là các nguyên hàm của $f(x) = x$.

**Nhận xét 3.5.** Một hàm số liên tục luôn có vô số nguyên hàm và các nguyên hàm đều khác nhau một hằng số.

**Định lý 3.6** (Định lý cơ bản của giải tích)**.** *Cho $f$ là một hàm thực liên tục trên đoạn $[a, b]$.*

*1. Cho $F$ được xác định trên $[a, b]$, bởi*

$$F(x) = \int_a^x f(t)dt, \quad \forall x \in [a, b].$$

*Khi đó ta có $F$ liên tục đều đặn (uniformly continuous) trên đoạn $[a, b]$, khả vi trên khoảng $(a, b)$ và*

$$F'(x) = f(x), \quad \forall x \in (a, b).$$

2. *(Hệ quả) Với $F$ là một nguyên hàm bất kì của $f$, ta có:*

$$\int_a^b f(x)dx = F(b) - F(a).$$

*Có thể viết gọn như sau*

$$F(b) - F(a) \equiv [F(x)]_{x=a}^b = [F(x)]_a^b = F(x)\mid_a^b.$$

**Nhận xét 3.7.** Sử dụng hệ quả 2 của định lý 3.6, chỉ cần tìm một nguyên hàm của hàm số cần tính tích phân, không cần sử dụng đúng hàm ở điều kiện 1. Điều đó cũng dẫn tới khái niệm **tích phân không xác định** sau.

## 3.3 Nguyên hàm dưới dạng tích phân không xác định

Định lý 3.6 nói rằng để tính tích phân của hàm $f$ trên đoạn $[a, b]$ trước tiên ta cần tìm một nguyên hàm $F$ của $f$. Nguyên hàm thường được biểu diễn như sau:

$$F(x) = \int f(x)dx.$$

Ví dụ,

$$\int x^2 = \frac{1}{3}x^3 + C, \int \sin 5x = -\frac{1}{5}\cos 5x + C, \ \ldots$$

**Ví dụ 3.8** (Một số nguyên hàm đơn giản)**.**

1. $\displaystyle\int x^n = \frac{x^{n+1}}{n+1} + C.$
2. $\displaystyle\int \frac{1}{x} = \ln|x| + C.$
3. $\displaystyle\int e^x = e^x + C.$
4. $\displaystyle\int a^x = \frac{a^x}{\ln a} + C.$

**Tính chất 3.9** (Các tính chất của tích phân)**.**

1. $\displaystyle\int [f(x) + g(x)]dx = \int f(x)dx + \int g(x)dx.$
2. $\displaystyle\int cf(x)dx = c\int f(x)dx.$
3. $\displaystyle\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx.$
4. $\displaystyle\int_a^b f(x)dx = -\int_b^a f(x)dx.$

**Ví dụ 3.10.** Tính

$$I = \int_1^2 (6x^5 - 2x^{-4} - 7x)dx.$$

Trước tiên cần tìm một nguyên hàm, sử dụng $\int x^n = \dfrac{x^{n+1}}{n+1}$:

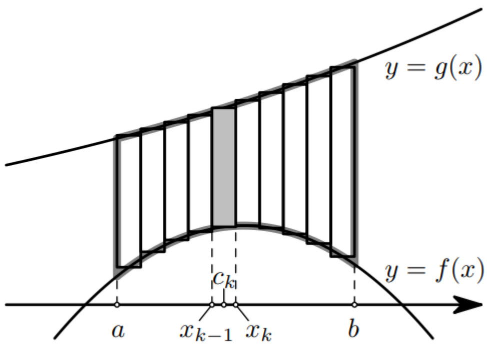
$$\begin{aligned} I &= \left(6 \cdot \frac{x^6}{6} - 2 \cdot \frac{x^{-3}}{-3} - 7 \cdot \frac{x^2}{2}\right)\Bigg|_1^2 \\ &= \left(x^6 + \frac{2}{3}x^{-3} - \frac{7}{2}x^2\right)\Bigg|_1^2. \end{aligned}$$

Thay số hai cận vào thu được kết quả

$$\begin{aligned} I &= \left(2^6 + \frac{2}{3} \cdot 2^{-3} - \frac{7}{2} \cdot 2^2\right) - \left(1^6 + \frac{2}{3} \cdot 1^{-3} - \frac{7}{2} \cdot 1^2\right) \\ &= \frac{601}{12} - \left(-\frac{11}{6}\right) \\ &= \frac{623}{12}. \end{aligned}$$

## 3.4 Một số ứng dụng của tích phân

### Diện tích giữa các đồ thị




Dễ thấy phần diện tích giữa đồ thị của 2 hàm số $f$ và $g$ là hiệu giữa phần lớn hơn và phần nhỏ hơn.

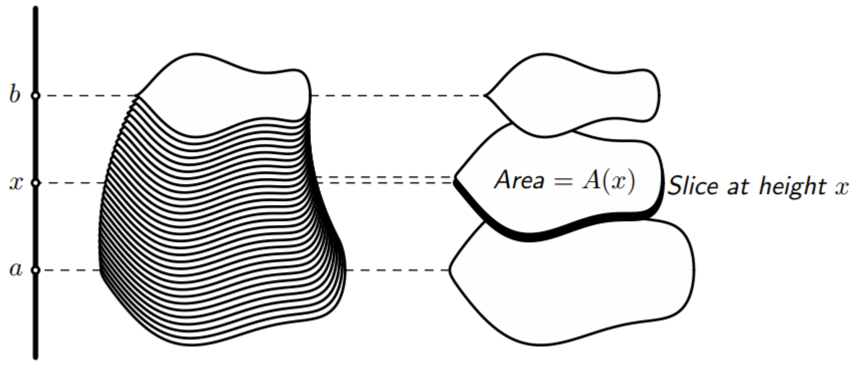
**Định lý 3.11.** *Cho 2 hàm số $f$, $g$ liên tục trên đoạn $[a, b]$ sao cho $f(x) \geq g(x)$, $\forall x \in (a, b)$. Khi đó diện tích nằm dưới đồ thị của $f$ và trên đồ thị của $g$ được tính bằng*

$$\int_a^b [f(x) - g(x)]\, dx.$$

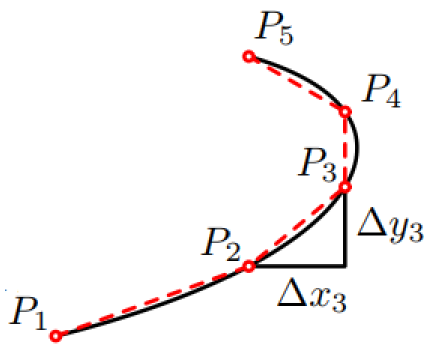
### Thể tích vật rắn

**Định lý 3.12.** *Giả sử vật cần tìm thể tích có chiều cao $h$, và tính được diện tích mặt cắt vuông góc với đường cao ứng với độ cao $x \in [0, h]$ là $A(x)$, khi đó:*

$$\int_0^h A(x) dx.$$




### Chiều dài của một đường cong



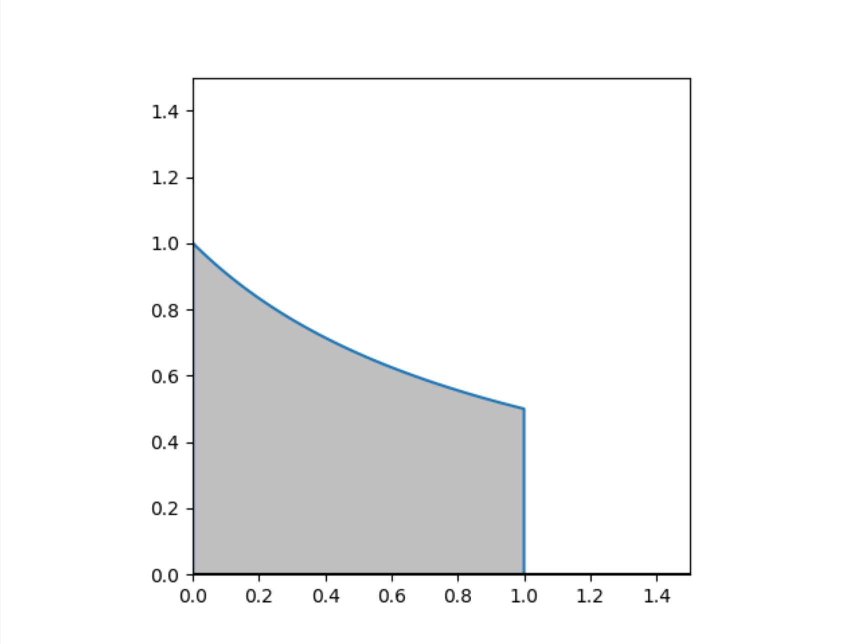

**Định lý 3.13.** *Cho hàm số $f$ liên tục trên $[a, b]$ và khả vi trên $(a, b)$. Chiều dài của đường cong đồ thị của $f$ từ $a$ đến $b$ được tính theo:*

$$L = \int_a^b \sqrt{1 + (f'(x))^2} dx.$$

Chứng minh sẽ được dành cho bạn đọc.

**Bài tập**

**Bài tập 3.1.** Cho hàm số $f$ liên tục không âm trên $[0, a]$. Gọi $A$ là mảng mặt phẳng được giới hạn bởi đồ thị của $f$ và các đoạn thẳng nối $(f(0), 0)$, $(0, 0)$, $(0, a)$ và $(a, f(a))$.




Tìm công thức tính thể tích của các hình sau:

(a) Hình khối được tạo bằng cách xoay mảng $A$ quanh trục tung.

(b) Hình khối được tạo bằng cách xoay mảng $A$ quanh trục hoành.

**Bài tập 3.2.** Tính diện tích của phần mặt phẳng **bị giới hạn** bởi 2 đồ thị $x^2 + x + 1$ và $2x + 3$.

**Bài tập 3.3** (**Khó**)**.** Chứng minh định lý 3.13. Gợi ý: dùng định nghĩa tích phân qua tổng Riemann.

# 4 Hàm số nhiều biến

## 4.1 Hàm số nhiều biến

**Định nghĩa 4.1** (Hàm số nhiều biến)**.** Hàm số $f$ gán vào mỗi $\mathbf{x} \in \mathbb{R}^n$ một điểm $\mathbf{y} \in \mathbb{R}^m$.

$$\mathbf{y} = f(\mathbf{x}).$$

Tập hợp của tất cả điểm $\mathbf{y}$ là **tập giá trị** của hàm số.

Mỗi thành phần của $\mathbf{y} = (y_1, ..., y_m)$ là một hàm trả về giá trị thực với $\mathbf{x} \in \mathbb{R}^n$, được viết dưới dạng $y_i = f_i(x)$:

$$\begin{array}{ccc} y_1 & = & f_1(\mathbf{x}), \\ y_2 & = & f_2(\mathbf{x}), \\ \vdots & \vdots & \vdots \\ y_m & = & f_m(\mathbf{x}). \end{array}$$

Nếu viết ra tất cả thành phần của $\mathbf{x} = (x_1, ..., x_n)$ thì hàm số ban đầu có thể được viết thành $m$ hàm số của $n$ biến:

$$\begin{array}{ccc} y_1 & = & f_1(x_1, ..., x_n), \\ y_2 & = & f_2(x_1, ..., x_n), \\ \vdots & \vdots & \vdots \\ y_m & = & f_m(x_1, ..., x_n). \end{array}$$

**Đồ thị** của hàm số là tất cả các cặp $(\mathbf{x}, \mathbf{y})$ với $\mathbf{y} = f(\mathbf{x})$.

**Ví dụ 4.2.**

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y+z \\ xyz \end{pmatrix},$$

xác định một hàm số từ $\mathbb{R}^3$ đến $\mathbb{R}^2$.

Ta cũng có thể viết $f : \mathbb{R}^3 \to \mathbb{R}^2$ dưới dạng

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} f_1(x,y,z) \\ f_2(x,y,z) \end{pmatrix}.$$

hoặc $f(x,y,z) = (f_1(x,y,z), f_2(x,y,z)) = f_1(x,y,z)\vec{i} + f_2(x,y,z)\vec{j}$.
Trong ví dụ này, các hàm

$$f_1\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x+y+z \text{ và } f_2\begin{pmatrix} x \\ y \\ z \end{pmatrix} = xyz,$$

được gọi là các hàm thành phần của $f$ và có thể viết

$$f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}.$$

## 4.2 Giới hạn, tính liên tục của hàm nhiều biến

**Định nghĩa 4.3** (Giới hạn)**.** Xét một hàm số $\mathbf{y} = f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^m$. Lấy $\mathbf{x^0} = (x_1^0, \ldots, x_n^0)$ là một điểm trong $\mathbb{R}^n$ và $\mathbf{y^0} = (y_1^0, \ldots, y_m^0)$ là một điểm trong $\mathbb{R}^m$. Ta nói rằng $\mathbf{y^0}$ là **giới hạn** của $f$ khi $\mathbf{x}$ tiến tới $\mathbf{x^0}$

$$\lim_{\mathbf{x}\to\mathbf{x}^0} f(\mathbf{x}) = \mathbf{y^0},$$

nếu

$$\forall \varepsilon > 0, \exists \delta > 0 : \left|\mathbf{x} - \mathbf{x^0}\right| < \delta \to \left|f(\mathbf{x}) - \mathbf{y^0}\right| < \varepsilon.$$

Tương tự với hàm một biến, giới hạn của hàm nhiều biến, nếu tồn tại, là duy nhất.

**Định lý 4.4.** *Nếu* $\mathbf{l}$ *và* $\mathbf{l}'$ *là giới hạn của* $\mathbf{f}$ *khi* $\mathbf{x}$ *tiến tới* $\mathbf{x^0}$*, thì* $\mathbf{l} = \mathbf{l}'$*.*

Định lý sau cho phép việc xác định giới hạn của hàm nhận giá trị trong $\mathbb{R}^m$ được quy về xác định giới hạn của hàm thực nhiều biến.

**Định lý 4.5.** *Cho hàm số* $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ *sao cho*

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{pmatrix},$$

*với* $f_i : \mathbb{R}^n \to \mathbb{R}$ *với mọi* $i$*. Khi đó*

$$\mathbf{l} = (l_1, l_2, \cdots, l_m) \in \mathbb{R}^m = \lim_{\mathbf{x}\to\mathbf{x}^0} \mathbf{f}(\mathbf{x}) \iff l_i = \lim_{\mathbf{x}\to\mathbf{x}^0} f_i(\mathbf{x}), \quad \forall i \leq n.$$

Khái niệm liên tục của hàm nhiều biến cũng là tổng quát từ khái niệm cho hàm một biến.

**Định nghĩa 4.6** (Tính liên tục của hàm nhiều biến)**.** Hàm số liên tục tại $\mathbf{x^0}$ nếu

$$\lim_{\mathbf{x}\to\mathbf{x}^0} f(\mathbf{x}) = f(\mathbf{x^0}).$$

Hàm số được gọi là liên tục nếu nó liên tục tại mọi điểm trên miền xác định của nó.

# 5 Đạo hàm của hàm nhiều biến

Để tổng quát hóa từ đạo hàm cho hàm một biến lên hàm nhiều biến, ta chú ý tới cách hiểu **"Đạo hàm là số đo sự biến thiên của hàm số theo chiều trục** $x$**"**. Như vậy, nếu hàm một biến thì có một trục $x$ để xem xét, còn hàm $n$ biến thì có $n$ trục. Do đó, ta dự đoán **"đạo hàm"** của hàm $n$ biến phải là một biểu thức phụ thuộc vào $n$ **"đạo hàm nhỏ"**, mỗi **"đạo hàm nhỏ"** thể hiện sự biến thiên của hàm qua một trục.

## 5.1 Đạo hàm riêng của hàm số thực

Ta bắt đầu với việc xây dựng khái niệm đạo hàm cho hàm thực nhiều biến

**Định nghĩa 5.1** (Đạo hàm riêng)**.** Cho $f : \mathbb{R}^n \to \mathbb{R}$ là một hàm số thực $n$ biến liên tục trên một lân cận của điểm $\mathbf{x}^0 = \left(x_1^0, x_2^0, \cdots, x_n^0\right)$. **Đạo hàm riêng theo** $x_i$ của $f$ tại $\mathbf{x}^0$ được tính theo công thức sau:

$$f_{x_i}\left(\mathbf{x}^0\right) = \frac{\partial f}{\partial x_i} = \lim_{h \to 0} \frac{f\left(x_1^0, \cdots, x_i^0 + h, x_{i+1}^0, \cdots, x_n^0\right) - f\left(x_1^0, \cdots, x_i^0, x_{i+1}^0, \cdots, x_n^0\right)}{h}.$$

Nói cách khác, $f_{x_i}\left(\mathbf{x}^0\right)$ là giá trị đạo hàm của $f$ tại $\mathbf{x}^0$ khi coi những biến ngoài $x_i$ là hằng số.

**Ví dụ 5.2.** Nếu $f(x, y) = x^2 + xy + y^3 - 1$ thì $f_x(x, y) = 2x + y$, $f_y(x, y) = x + 3y^2$.

**Nhận xét 5.3** (Đạo hàm riêng bậc cao)**.** Có thể đạo hàm riêng $f_x, f_y$ theo $x$ và $y$ để thu được 4 đạo hàm bậc hai:

$$f_{xx} = \frac{\partial^2 f}{\partial x^2}, \quad f_{yy} = \frac{\partial^2 f}{\partial y^2}, \qquad f_{xy} = \frac{\partial^2 f}{\partial x \partial y}, \quad f_{yx} = \frac{\partial^2 f}{\partial y \partial x}.$$

**Ví dụ 5.4.** Với $f(x, y) = x^4 + x^2 y + y^3 - 1$ như trên, khi đó

$$f_{xx} = 12x^2 + 2y, \quad f_{yy} = 6y, \quad f_{xy} = f_{yx} = 2x.$$

## 5.2 Quy tắc mắt xích cho hàm nhiều biến

Nhớ lại quy tắc mắt xích cho đạo hàm hàm một biến: nếu $f = f(u)$ và $u = u(t)$ thì $\dfrac{df}{dt} = \dfrac{df}{du} \cdot \dfrac{du}{dt}$. Quy tắc mắt xích tổng quát cho hàm số $n$ biến phát biểu như sau:

**Định lý 5.5** (Quy tắc mắt xích cho đạo hàm toàn phần)**.** *Cho hàm số khả vi* $f : \mathbb{R}^n \to \mathbb{R}$ *và các hàm số khả vi* $x_1, x_2, \cdots, x_n : \mathbb{R} \to \mathbb{R}$ *đều là hàm theo biến* $t$*, khi đó:*

$$\frac{\mathrm{d}f}{\mathrm{d}t} = f_{x_1}\frac{\mathrm{d}x_1}{\mathrm{d}t} + f_{x_2}\frac{\mathrm{d}x_2}{\mathrm{d}t} + \cdots + f_{x_n}\frac{\mathrm{d}x_n}{\mathrm{d}t}.$$

**Ví dụ 5.6.** Xét $f = f(x, y)$ với $x, y$ đều là các hàm của cùng một biến số $t$: $x = x(t)$, $y = y(t)$. Khi đó:

$$\frac{df}{dt} = f_x \cdot \frac{dx}{dt} + f_y \cdot \frac{dy}{dt}.$$

Nếu $f(x, y) = x^2 + y^2$ và $x(t) = \cos t, \quad y(t) = \sin t$, ta có:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = 2x(t) \cdot (-\sin t) + 2y(t) \cdot \cos t = -2\cos t \sin t + 2\sin t \cos t = 0.$$

Viết lại $f$ theo $t$ trước rồi lấy đạo hàm, ta có:

$$f(x(t), y(t)) = (\cos t)^2 + (\sin t)^2 = 1 \implies \frac{\mathrm{d}f}{\mathrm{d}t} = 0.$$

**Ví dụ 5.7.** Tương tự đối với hàm $f(x, y, z)$ với $x = x(t)$, $y = y(t)$, $z = z(t)$ ta có

$$\frac{df}{dt} = f_x \cdot \frac{dx}{dt} + f_y \cdot \frac{dy}{dt} + f_z \cdot \frac{dz}{dt}.$$

Trong trường hợp các hàm thành phần của $\mathbf{x}$ có nhiều hơn một biến, $f$ sẽ là một *hàm nhiều biến*, khi đó quy tắc mắt xích để tính đạo hàm riêng theo từng biến của $f$ như sau:

**Định lý 5.8** (Quy tắc mắt xích cho đạo hàm riêng)**.** *Cho hàm số khả vi $f : \mathbb{R}^n \to \mathbb{R}$ và các hàm số khả vi $x_1, x_2, \cdots, x_n : \mathbb{R}^m \to \mathbb{R}$ đều là hàm theo các biến $t_1, t_2, \cdots, t_m$, khi đó:*

$$\frac{\partial f}{\partial t_i} = f_{x_1}\frac{\partial x_1}{\partial t_i} + f_{x_2}\frac{\partial x_2}{\partial t_i} + \cdots + f_{x_n}\frac{\partial x_n}{\partial t_i}, \quad \forall i \in \{1, 2, \cdots, m\}.$$

**Ví dụ 5.9.** Nếu $f = f(x, y, z)$ với $x = x(s, t)$, $y = y(s, t)$, $z = z(s, t)$ thì

$$\frac{\partial f}{\partial s} = f_x \cdot \frac{\partial x}{\partial s} + f_y \cdot \frac{\partial y}{\partial s} + f_z \cdot \frac{\partial z}{\partial s},$$

tương tự

$$\frac{\partial f}{\partial t} = f_x \cdot \frac{\partial x}{\partial t} + f_y \cdot \frac{\partial y}{\partial t} + f_z \cdot \frac{\partial z}{\partial t}.$$

Nếu $f(x, y) = x^2 + y^2 - z^2$ với

$$\begin{aligned} x(s, t) &= s^2 - t^2, \\ y(s, t) &= 2st, \\ z(s, t) &= s^2 + t^2. \end{aligned}$$

Khi đó

$$\begin{aligned} \frac{\partial f}{\partial s} &= 2x \cdot 2s + 2y \cdot 2t - 2z \cdot 2s = (4s^3 - 4st^2) + 8st^2 - (4s^3 + 4st^2) = 0. \\ \frac{\partial f}{\partial t} &= -2x \cdot 2t + 2y \cdot 2s - 2z \cdot 2t = -(4s^2t - 4t^3) + 8s^2t - (4s^2t + 4t^3) = 0. \end{aligned}$$

Khai triển $f$ theo $s$ và $t$, ta được:

$$f(x(s, t), y(s, t), z(s, t)) = (s^2 - t^2)^2 + 4s^2t^2 - (s^2 + t^2)^2 = 0.$$

Do đó

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial t} = 0.$$

### 5.3 Đạo hàm có hướng (directional derivative) - Gradient

Như đã đề cập, đạo hàm riêng của hàm số $f$ theo biến $x_i$ là độ biến thiên của $f$ theo trục thứ $i$ trong không gian $n$ chiều. Một cách tự nhiên, ta đặt câu hỏi: **Liệu sự biến thiên của $f$ theo một hướng bất kì, không trùng với các trục có sẵn được đo như thế nào?** Khái niệm **Đạo hàm có hướng** giúp ta giải quyết câu hỏi trên.

**Định nghĩa 5.10** (Đạo hàm có hướng)**.** Cho hàm số $f : \mathbb{R}^n \to \mathbb{R}$, điểm $\mathbf{x}^* = (x_1^*, x_2^*, \cdots, x_n^*) \in \mathbb{R}^n$ và vector $\mathbf{u} \in \mathbb{R}^n$. Khi đó **Đạo hàm theo hướng u** của $f$ tại $\mathbf{x}^0$ được tính theo công thức:

$$D_{\mathbf{u}} f\left(\mathbf{x}^*\right) = \lim_{h \to 0} \frac{f\left(\mathbf{x}^* + h\mathbf{u}\right) - f\left(\mathbf{x}^*\right)}{h}.$$

nếu giới hạn bên phải tồn tại. Nếu không, ta nói hàm số $f$ **không khả vi** theo hướng $\mathbf{u}$ tại $\mathbf{x}^*$.

**Định nghĩa 5.11.** Hàm số $f : \mathbb{R}^n \to \mathbb{R}$ được gọi là **khả vi** (bậc 1) tại $\mathbf{x}^*$ nếu như $f$ khả vi theo mọi hướng và $D_{\mathbf{u}} f$ liên tục tại $\mathbf{x}^*$ với mọi $\mathbf{u}$.

**Nhận xét 5.12.** Nếu định nghĩa hàm $f_{\mathbf{u}} : \mathbb{R} \to \mathbb{R}$ bởi $f_{\mathbf{u}}(t) = f\left(\mathbf{x}^* + t\mathbf{u}\right)$, thì

$$D_{\mathbf{u}} f\left(\mathbf{x}^*\right) = \frac{\mathrm{d} f_{\mathbf{u}}}{\mathrm{d} t}(0).$$

**Nhận xét 5.13.** Đây là một hướng tổng quát tự nhiên cho khái niệm đạo hàm riêng. Thực chất, nếu $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ là hệ cơ sở chính tắc[1] của $\mathbb{R}^n$, tức là

$$\begin{aligned} e_{ij} &= 0, \quad \forall i \neq j; \\ e_{ii} &= 1, \quad \forall i \leq n. \end{aligned}$$

Khi đó:

$$D_{\mathbf{e}_i} f = \frac{\partial f}{\partial x_i}, \quad \forall i \leq n.$$

**Ví dụ 5.14.** Đạo hàm theo hướng $(a, b)$ hàm số $f(x, y) = x^2 + y^2$ là:

$$\begin{aligned} \lim_{h \to 0} \frac{(x + ah)^2 + (y + bh)^2 - x^2 - y^2}{h} &= \lim_{h \to 0} \frac{2axh + h^2 + 2byh + h^2}{h} \\ &= 2ax + 2by. \end{aligned}$$

**Nhận xét 5.15.** Trong ví dụ trên, có thể thấy với mọi $\mathbf{u} = (a, b)$,

$$D_{\mathbf{u}} f(x, y) = \begin{bmatrix} 2x \\ 2y \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} f_x(x, y) \\ f_y(x, y) \end{bmatrix} \cdot \mathbf{u}.$$

Điều xảy ra ở nhận xét 5.15 thực chất đúng với mọi hàm số khả vi và vector $\mathbf{u}$. Ta có định lý quan trọng sau:

**Định lý 5.16** (Công thức đạo hàm có hướng)**.** *Cho $f : \mathbb{R}^n \to \mathbb{R}$ là một hàm số thực n biến có đủ đạo hàm riêng trên một lân cận của điểm $\mathbf{x}^0 = (x_1^0, x_2^0, \cdots, x_n^0)$. Khi đó với mọi vector $\mathbf{u} \in \mathbb{R}^n$, $D_{\mathbf{u}} f$ tồn tại và:*

$$D_{\mathbf{u}} f = \begin{bmatrix} f_{x_1}\left(\mathbf{x}^0\right) \\ f_{x_2}\left(\mathbf{x}^0\right) \\ \vdots \\ f_{x_n}\left(\mathbf{x}^0\right) \end{bmatrix} \cdot \mathbf{u} = \nabla f \cdot \mathbf{u}.$$

1 Khái niệm cơ sở và hệ cơ sở chính tắc nằm trong bài giảng Đại số Tuyến tính

Định lý trên còn cho ta điều kiện cần và đủ để một hàm số khả vi bậc 1:

**Hệ quả 5.17.** Hàm số $f : \mathbb{R}^n \to \mathbb{R}$ khả vi tại $\mathbf{x}^* \in \mathbb{R}^n$ khi và chỉ khi $f_{x_i}$ xác định và liên tục tại $\mathbf{x}^*$ với mọi $i = 1, 2, \ldots, n$.

**Ví dụ 5.18.** Nếu $f : \mathbb{R}^2 \to \mathbb{R}$ thì đạo hàm có hướng của nó là:

$$D_{\mathbf{u}} f = \frac{\partial f}{\partial x} u_1 + \frac{\partial f}{\partial y} u_2 = \nabla \cdot \mathbf{u},$$

với $\mathbf{u} = (u_1, u_2)$ là vector đơn vị.

Sự xuất hiện bất biến của đại lượng $\nabla$ trong mọi công thức đạo hàm có hướng dẫn đến định nghĩa sau:

**Định nghĩa 5.19** (Gradient)**.** Đại lượng $\nabla f$ trong định lý 5.16 được gọi là **gradient** của $f$ tại điểm $\mathbf{x}^0$:

$$\nabla f = \begin{bmatrix} f_{x_1}\left(\mathbf{x}^0\right) \\ f_{x_2}\left(\mathbf{x}^0\right) \\ \vdots \\ f_{x_n}\left(\mathbf{x}^0\right) \end{bmatrix}.$$

**Lưu ý.** $\nabla f$ là một hàm từ $\mathbb{R}^n$ đến $\mathbb{R}^n$.

**Ví dụ 5.20.** Hàm số $f(x, y) = x^2 + y^2$ sẽ có gradient là $\nabla f = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$.

**Ví dụ 5.21.** Xét hàm số $f : \mathbb{R}^3 \to \mathbb{R},\ f(\mathbf{x}) = \left(x_1^2 + 2x_2^2 + 3x_3^2\right)^{1/2}$. Đạo hàm bậc nhất theo $\mathbf{x}$ của hàm số đó là

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f(\mathbf{x})}{\partial x_1} \\ \dfrac{\partial f(\mathbf{x})}{\partial x_2} \\ \dfrac{\partial f(\mathbf{x})}{\partial x_3} \end{bmatrix} = \begin{bmatrix} x_1\left(x_1^2 + 2x_2^2 + 3x_3^2\right)^{-1/2} \\ 2x_2\left(x_1^2 + 2x_2^2 + 3x_3^2\right)^{-1/2} \\ 3x_3\left(x_1^2 + 2x_2^2 + 3x_3^2\right)^{-1/2} \end{bmatrix}.$$

### 5.4 Đạo hàm của hàm nhận giá trị vector - Ma trận Jacobi

Xét $\mathbf{f} : \mathbb{R} \to \mathbb{R}^m$:

$$\mathbf{f}(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{bmatrix}.$$

Đạo hàm của hàm số này theo $x$ là một vector cột như sau:

$$\frac{\mathrm{d}\mathbf{f}}{\mathrm{d}x} \triangleq \begin{bmatrix} f_1'(x) \\ f_2'(x) \\ \vdots \\ f_m'(x) \end{bmatrix}.$$

Kết hợp biểu diễn trên với khái niệm gradient, ta được công thức tổng quát cho "đạo hàm" của một hàm số từ $\mathbb{R}^n$ đến $\mathbb{R}^m$ như sau:

**Định nghĩa 5.22** (Ma trận Jacobi)**.** Cho hàm số $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ với các hàm thành phần $f_i : \mathbb{R}^n \to \mathbb{R}, i = 1, 2, \cdots, m$. Các đạo hàm riêng bậc nhất của hàm này (nếu tồn tại) sẽ được xếp thành một ma trận có kích thước $m \times n$, được gọi là **ma trận Jacobi** (**Jacobian**) của **f**:

$$\mathbf{J_f}(\mathbf{x}) \triangleq \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{x})}{\partial x_1} & \frac{\partial f_m(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{m\times n}.$$

**Nhận xét 5.23.** Ma trận Jacobi còn có thể được viết dưới dạng sau:

$$\mathbf{J_f}(\mathbf{x}) = \begin{bmatrix} \nabla f_1(\mathbf{x})^T \\ \nabla f_2(\mathbf{x})^T \\ \vdots \\ \nabla f_m(\mathbf{x})^T \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{f}}{\partial x_1}(\mathbf{x}) & \frac{\partial \mathbf{f}}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial \mathbf{f}}{\partial x_n}(\mathbf{x}) \end{bmatrix}.$$

**Nhận xét 5.24.** Cho hàm $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ khả vi tại $\mathbf{x}$ và vector $\mathbf{u} \in \mathbb{R}^n$. Đạo hàm theo hướng **u** của **f** tại **x** chính là:

$$\mathbf{D_u f}(\mathbf{x}) = \begin{bmatrix} D_\mathbf{u} f_1(\mathbf{x}) \\ D_\mathbf{u} f_2(\mathbf{x}) \\ \vdots \\ D_\mathbf{u} f_m(\mathbf{x}) \end{bmatrix} = \mathbf{J_f}(\mathbf{x}) \cdot \mathbf{u}.$$

## 5.5 Đạo hàm bậc hai - Ma trận Hessian

Từ định nghĩa 5.22, ta có thể suy ra công thức tìm *đạo hàm bậc hai* của một hàm số $\mathbb{R}^n \to \mathbb{R}$ bằng cách coi gradient (đạo hàm bật nhất) như một hàm $\mathbb{R}^n \to \mathbb{R}^n$.

**Định nghĩa 5.25** (Ma trận Hessian)**.** Đạo hàm bậc hai (Hessian) của hàm số $f : \mathbb{R}^n \to \mathbb{R}$, nếu tất cả đạo hàm riêng bậc hai của $f$ tồn tại và liên tục trên miền xác định, được định nghĩa và sắp xếp như sau:

$$\mathbf{H}_f(\mathbf{x}) \triangleq \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_2^2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x}) \end{bmatrix}.$$

Trong trường hợp $\mathbf{f}$ là một hàm vector $\mathbb{R}^n \to \mathbb{R}^m$, tập hợp các đạo hàm riêng bậc hai không phải là một ma trận $n \times n$ mà là một tensor bậc ba. Nó có thể được biểu diễn đơn giản dưới dạng một vector gồm $m$ ma trận Hessian, mỗi ma trận con là một hàm thành phần của $\mathbf{f}$

$$\mathbf{H}(\mathbf{f}) = \begin{bmatrix}\mathbf{H}(f_1) & \mathbf{H}(f_2) & \dots & \mathbf{H}(f_m)\end{bmatrix}.$$

Cũng như đạo hàm bậc một, ta cũng có khái niệm đạo hàm bậc 2 theo hướng, được định nghĩa một cách tương tự:

**Định nghĩa 5.26.** Cho hàm số $f : \mathbb{R}^n \to \mathbb{R}$ khả vi bậc 1 tại điểm $\mathbf{x}^* \in \mathbb{R}^n$ và vector $\mathbf{u} \in \mathbb{R}^n$. **Đạo hàm bậc 2** tại $\mathbf{x}^*$ của $f$ được tính theo công thức:

$$D^2_{\mathbf{u}} f\left(\mathbf{x}^*\right) = \lim_{h \to 0} \frac{D_{\mathbf{u}} f\left(\mathbf{x}^* + h\mathbf{u}\right) - D_{\mathbf{u}} f\left(\mathbf{x}^*\right)}{h} = \frac{\mathrm{d}^2 f\left(\mathbf{x}^* + t\mathbf{u}\right)}{\mathrm{d}t^2}(0)$$

nếu giới hạn trên tồn tại.

**Định nghĩa 5.27.** Hàm số $f : \mathbb{R}^n \to \mathbb{R}$ **khả vi bậc 2** tại $\mathbf{x}^* \in \mathbb{R}^n$ nếu $D^2_{\mathbf{u}} f$ xác định và liên tục tại $\mathbf{x}^*$ với mọi $\mathbf{u} \in \mathbb{R}^n$.

**Ví dụ 5.28.** Hàm số $f(x, y, z) = x^3 + y^3 + z^3$ xác định trên $\mathbb{R}$ có $\nabla f(x, y, z) = \begin{bmatrix}3x^2 & 3y^2 & 3z^2\end{bmatrix}^T$ Với $\mathbf{u} = \begin{bmatrix}a & b & c\end{bmatrix}^T$, ta có

$$D_{\mathbf{u}} f(x, y, z) = 3ax^2 + 3by^2 + 3cz^2$$

Do đó

$$\nabla D_{\mathbf{u}} f(x, y, z) = \begin{bmatrix}6ax & 6by & 6cz\end{bmatrix}^T$$

Vậy

$$D^2_{\mathbf{u}} f(x, y, z) = \nabla D_{\mathbf{u}} f(x, y, z) \cdot \mathbf{u} = 6a^2x + 6b^2y + 6c^2z$$

Định lý quan trọng sau liên hệ ma trận Hessian và khái niệm đạo hàm bậc 2.

**Định lý 5.29.** *Cho hàm số $f : \mathbb{R}^n \to \mathbb{R}$, điểm $\mathbf{x}^* \in \mathbb{R}^n$ và vector $\mathbf{u} \in \mathbb{R}^n$. Nếu $D^2_{\mathbf{u}} f\left(\mathbf{x}^*\right)$ tồn tại thì*

$$D^2_{\mathbf{u}} f\left(\mathbf{x}^*\right) = \mathbf{u}^T \mathbf{H}(f) \mathbf{u}$$

Chứng minh sẽ dành cho bạn đọc.

**Ví dụ 5.30.** Hàm số $f(x, y, z) = x^3 + y^3 + z^3$ xác định trên $\mathbb{R}$ có

$$\mathbf{H}(f) = \begin{bmatrix}6x & 0 & 0 \\ 0 & 6y & 0 \\ 0 & 0 & 6z\end{bmatrix}$$

Do đó, với $\mathbf{u} = \begin{bmatrix}a & b & c\end{bmatrix}^T$,

$$D^2_{\mathbf{u}} f(x, y, z) = \begin{bmatrix}a & b & c\end{bmatrix} \begin{bmatrix}6x & 0 & 0 \\ 0 & 6y & 0 \\ 0 & 0 & 6z\end{bmatrix} \begin{bmatrix}a \\ b \\ c\end{bmatrix} = 6a^2x + 6b^2y + 6c^2z$$

Ta thu được cùng kết quả với ví dụ 5.28.

Từ định lý 5.29, ta có hệ quả tương tự với hệ quả 5.17:

**Hệ quả 5.31.** Một hàm số $f : \mathbb{R}^n \to \mathbb{R}$ khả vi bậc 2 tại $\mathbf{x}^* \in \mathbb{R}^n$ khi và chỉ khi $f_{x_i x_j}$ xác định và liên tục tại $\mathbf{x}^*$ với mọi $i, j = 1, 2, \ldots, n$.

**Nhận xét 5.32.** Không giống như đạo hàm bậc 1, việc các đạo hàm riêng bậc 2 $f_{x_i x_i}$ xác định và liên tục là không đủ để hàm số khả vi tại một điểm. Ta cần các "đạo hàm bậc 2 kết hợp" dạng $f_{x_i x_j}$.

Đây cũng là một trong những khó khăn chính cho việc xây dựng khái niệm đạo hàm bậc 2 từ các đạo hàm riêng. Ta có thể dễ dàng định nghĩa *đạo hàm riêng bậc 2 theo $x_i$ và $x_j$* bằng công thức đạo hàm riêng 2 lần, tuy nhiên ta không thể tìm một giải thích trực quan như "coi các biến khác là hằng số, đạo hàm theo $x_i$" như trường hợp bậc 1. Việc định nghĩa Hessian trước thông qua ma trận Jacobi, mà bắt nguồn là đạo hàm của hàm vector, cho ta một con đường "thẳng thớm" hơn dẫn tới khái niệm đạo hàm bậc 2. Một cách trực quan, đạo hàm bậc 2 chính là "đạo hàm của đạo hàm bậc 1".

## 5.6 Thứ tự của đạo hàm riêng bậc 2 - Định lý Schwarz

Trong ví dụ 5.4, hàm số $f(x, y) = x^4 + x^2 y + y^3 - 1$ có $f_{xy} = f_{yx} = 2x$. Sự đối xứng này không phải ngẫu nhiên. Ta có định lý sau:

**Định lý 5.33** (Schwarz)**.** *Nếu hàm số $f : \mathbb{R}^n \to \mathbb{R}$ khả vi bậc 2 tại $\mathbf{x}^* \in \mathbb{R}^n$,ta có*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}^*) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}^*) \quad \forall i, j = 1, 2, \ldots, n$$

Chứng minh sẽ dành cho bạn đọc.

**Hệ quả 5.34.** Ma trận Hessian của hàm khả vi bậc 2 là ma trận đối xứng

Sự đối xứng của ma trận Hessian giúp ta dễ dàng tìm các trị riêng của nó hơn bình thường, và chúng đều là số thực. Tính chất này sẽ được nhắc đến trong phần 6, Cực trị hàm nhiều biến, cụ thể là trong Định lý đạo hàm bậc 2.

**Bài tập**

**Bài tập 5.1.** Tìm $\nabla f$ và $D_{\mathbf{u}} f(x_0, y_0)$ với:

(a) $f(x, y) = x^2 + xy + y^2$,
$\mathbf{u} = (1, 2)$,
$(x_0, y_0) = (1, -1)$.

(b) $f(x, y) = \sin x \cos y + \sin y \cos x$,
$\mathbf{u} = (1, 1)$,
$(x_0, y_0) = (0, 0)$.

**Bài tập 5.2** (**Khó**)**.** Phát biểu và chứng minh cách tìm phương trình mặt phẳng tiếp xúc đồ thị hàm số $f : \mathbb{R}^2 \to \mathbb{R}$ tại điểm $(x_0, y_0)$. Gợi ý: sử dụng gradient.

**Bài tập 5.3** (**Khó**)**.** Chứng minh định lý 5.16 sử dụng định nghĩa 5.10 của đạo hàm có hướng và định nghĩa 5.1 của đạo hàm riêng.

**Bài tập 5.4.** Tìm ma trận Jacobi của hàm số $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^2$ và tính giá trị của nó tại $(1, 2, 3)$:

$$\mathbf{f}(x, y, z) = \begin{bmatrix} x + y + z \\ xyz \end{bmatrix}.$$

**Bài tập 5.5.** Tìm ma trận Hessian của hàm số $f : \mathbb{R}^3 \to \mathbb{R}$:

$$f(x, y, z) = x^3 + y^3 + z^3 - 3xyz.$$

# 6 Cực trị hàm nhiều biến

## 6.1 Khái niệm

Tương tự với hàm một biến, đối với hàm nhiều biến ta cũng có các khái niệm cực đại/cực tiểu toàn cục/địa phương.

Việc tìm cực trị của hàm nhiều biến là cốt lõi của các bài toán tối ưu hóa trong Toán ứng dụng, Machine Learning cũng như nhiều ngành khoa học/kỹ thuật khác.

Trong các định nghĩa và kết quả sau, $D$ là một tập hợp con của $\mathbb{R}^n$.

**Định nghĩa 6.1.** Cho hàm số $f : D \to \mathbb{R}$. Một điểm $\mathbf{x}^* \in D$ được gọi là một **điểm cực đại toàn cục** của $f$ nếu $f(\mathbf{x}^*) \geq f(\mathbf{x})$ với mọi $\mathbf{x} \in D$.

**Định nghĩa 6.2.** Cho hàm số $f : D \to \mathbb{R}$. Một điểm $\mathbf{x}^* \in D$ được gọi là một **điểm cực tiểu toàn cục** của $f$ nếu $f(\mathbf{x}^*) \leq f(\mathbf{x})$ với mọi $\mathbf{x} \in D$.

Trước khi định nghĩa cực trị địa phương, ta bắt đầu với khái niệm **lân cận**. Như đã biết, lân cận của $x \in \mathbb{R}$ là một khoảng $(x - \varepsilon, x + \varepsilon)$ nào đó. Trong không gian $n$ chiều, ta không có khái niệm so sánh giữa 2 điểm, do vậy không thể định nghĩa khoảng và đoạn như trong $\mathbb{R}$. Tuy nhiên, nếu chỉ nhìn vào bản chất của khái niệm lân cận là tập hợp các điểm *nằm gần* $x$, ta có thể định nghĩa lân cận trên không gian $n$ chiều bằng quan hệ khoảng cách.

**Định nghĩa 6.3.** Một **lân cận** của điểm $\mathbf{x}^* \in \mathbb{R}^n$ là một tập hợp $B_r(\mathbf{x}^*) = \{\mathbf{x} : \|\mathbf{x} - \mathbf{x}^*\| < r\}$.

**Lưu ý.** Khái niệm lân cận được định nghĩa khác trong topo và giải tích chuyên sâu. Trong tài liệu này, chúng tôi chỉ sử dụng một định nghĩa có phần "tạm bợ" cho từ lân cận nhằm giúp bạn đọc tiếp thu các kiến thức về cực trị một cách dễ dàng. Nếu sử dụng khái niệm chuẩn của lân cận ta vẫn được các kết quả tương đồng với các kết quả trong tài liệu này.

**Định nghĩa 6.4.** Cho hàm số $f : D \to \mathbb{R}$. Một điểm $\mathbf{x}^* \in D$ được gọi là một **điểm cực đại địa phương** của $f$ nếu tồn tại $r > 0$ sao cho $B_r(\mathbf{x}^*) \subset D$ và $f(\mathbf{x}^*) \geq f(\mathbf{x})$ với mọi $\mathbf{x} \in B_r(\mathbf{x}^*)$.

**Định nghĩa 6.5.** Cho hàm số $f : D \to \mathbb{R}$. Một điểm $\mathbf{x}^* \in D$ được gọi là một **điểm cực tiểu địa phương** của $f$ nếu tồn tại $r > 0$ sao cho $B_r(\mathbf{x}^*) \subset D$ và $f(\mathbf{x}^*) \leq f(\mathbf{x})$ với mọi $\mathbf{x} \in B_r(\mathbf{x}^*)$.

Trước khi đến với các phương pháp tìm cực trị, ta định nghĩa khái niệm *miền trong* của một tập hợp.

**Định nghĩa 6.6. Miền trong** của một tập hợp $D \subseteq \mathbb{R}^n$ là tập hợp sau:

$$D^o = \{\mathbf{x} \in D : \exists r > 0 : B_f(\mathbf{x}) \subset D\}.$$

**Ranh giới** của $D$ là phần còn lại: $\partial D = D \setminus D^o$.

**Nhận xét 6.7.** Rõ ràng cực trị địa phương chỉ có thể xuất hiện với các điểm ở miền trong của tập xác định. Trên thực tế, các hàm số cần được tối ưu hóa xuất hiện trong các bài toán ứng dụng, bao gồm Machine Learning, đều có tập xác định mà miền trong chiếm phần lớn. Do đó bạn đọc không cần quá bận tâm về việc tìm miền trong của tập xác định. Trong đại đa số các trường hợp, việc xác định miền trong là tương đối dễ dàng.

## 6.2 Định lý đạo hàm bậc nhất

Định lý sau cho phép ta xác định mọi điểm thuộc miền trong *có khả năng* là cực trị địa phương. Tuy chỉ áp dụng được cho hàm khả vi bậc 1, đại đa số các hàm được dùng trong thực tế đều khả vi bậc 1 nên có thể áp dụng được.

**Định lý 6.8** (Định lý đạo hàm bậc nhất)**.** *Cho hàm số $f : D \to \mathbb{R}$ khả vi bậc 1. Nếu $\mathbf{x}^* \in D^o$ là một điểm cực trị địa phương thì $\nabla f(\mathbf{x}^*) = 0$.*

**Ví dụ 6.9.** Hàm số $f(x,y) = x^2 + y^2$ xác định trên $\mathbb{R}$ có duy nhất một điểm cực trị địa phương (cũng là cực tiểu toàn cục) là $(0,0)$. Ta có:

$$\nabla f(0,0) = \begin{bmatrix} 2x \\ 2y \end{bmatrix} \Bigg|_{(0,0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

*Ý tưởng chứng minh.* Nếu $f(\mathbf{x}^*)$ là cực đại địa phương (tương tự với cực tiểu), xét hàm số $f_1 : \mathbb{R} \to \mathbb{R}$ sao cho

$$f_1(x) = f\left(x, x_2^*, x_3^*, \dots, x_n^*\right).$$

Khi đó dễ thấy $x_1^*$ là một điểm cực đại địa phương của $f_1$, do đó

$$0 = \frac{\mathrm{d}f_1}{\mathrm{d}x}\left(x_1^*\right) = \frac{\partial f}{\partial x_1}\left(\mathbf{x}^*\right).$$

Làm tương tự với $x_2^*, x_3^*, \dots, x_n^*$, ta có điều phải chứng minh. □

Đến đây, ta có thể nghĩ rằng, mọi điểm thỏa mãn $\nabla f = 0$ đều là một cực trị địa phương. Tuy nhiên, tương tự như trường hợp 1 biến, một số điểm có thể làm cho $\nabla f = 0$ nhưng vẫn không phải cực trị địa phương.

**Ví dụ 6.10.** Hàm số $f(x,y) = x^3 + y^3$ thỏa mãn $\nabla f(0,0) = 0$ nhưng điểm $(0,0)$ không phải là cực trị địa phương.

**Ví dụ 6.11.** Hàm số $f(x,y) = x^2 - y^2$ thỏa mãn $\nabla f(0,0) = 0$ nhưng điểm $(0,0)$ không phải là cực trị địa phương.

Vậy sẽ có những điểm thỏa mãn gradient bằng 0 nhưng không phải là cực trị địa phương. Những điểm này gọi là **Điểm yên ngựa**.

**Định nghĩa 6.12.** Cho hàm số $f : D \to \mathbb{R}$ khả vi bậc 1. Điểm $\mathbf{x}^* \in D^o$ là một **điểm yên ngựa** nếu $\nabla f(\mathbf{x}^*) = 0$ nhưng $\mathbf{x}^*$ không phải là một cực trị địa phương.

Sau đây ta tìm hiểu một loại điểm yên ngựa đặc biệt, sẽ có ý nghĩa trong việc tìm hiểu các tính chất quan trọng của cực trị địa phương.

Trong ví dụ 6.11, $\nabla f = \begin{bmatrix} 2x & -2y \end{bmatrix}^T$.

Đạo hàm theo hướng $\mathbf{e}_1 = \begin{bmatrix}1\\0\end{bmatrix}$ là $D_{\mathbf{e}_1}f(x,y) = 2x$. $D_{\mathbf{e}_1}f$ đổi dấu từ âm sang dương khi đi qua $(0,0)$, do đó $(0,0)$ là một điểm **cực tiểu địa phương** *theo hướng* $\mathbf{e}_1$.

Đạo hàm theo hướng $\mathbf{e}_2 = \begin{bmatrix}0\\1\end{bmatrix}$ là $D_{\mathbf{e}_2}f(x,y) = -2y$. $D_{\mathbf{e}_2}f$ đổi dấu từ dương sang âm khi đi qua $(0,0)$, do đó $(0,0)$ là một điểm **cực đại địa phương** *theo hướng* $\mathbf{e}_2$.

Sau đây ta làm rõ khái niệm cực trị theo hướng, và những hệ quả quan trọng của nó trong việc tìm cực trị địa phương.

**Định nghĩa 6.13.** Cho hàm số $f : D \to \mathbb{R}$ khả vi bậc 1 và vector $\mathbf{u} \in \mathbb{R}^n \setminus \{0\}$. Điểm $\mathbf{x}^* \in D^o$ là một cực đại/tiểu địa phương **theo hướng u** nếu hàm số sau

$$f_{\mathbf{u}}(t) = f\left(\mathbf{x}^* + t\mathbf{u}\right),$$

đạt cực đại/tiểu địa phương tại $t = 0$.

Khái niệm cực trị theo hướng này giúp ta làm rõ thêm bản chất của các điểm cực trị địa phương. Ta có định lý sau:

**Định lý 6.14.** *Cho hàm số $f : D \to \mathbb{R}$ khả vi bậc 1. Điểm $\mathbf{x}^* \in D^o$ là một điểm cực đại (cực tiểu) địa phương khi và chỉ khi với mọi vector $\mathbf{u} \in \mathbb{R}^n \setminus \{0\}$, hàm số theo $t$ sau*

$$\frac{\mathrm{d}f\left(\mathbf{x}^* + t\mathbf{u}\right)}{\mathrm{d}t} = D_{\mathbf{u}}f\left(\mathbf{x}^* + t\mathbf{u}\right) = \nabla f\left(\mathbf{x}^* + t\mathbf{u}\right)\cdot\mathbf{u},$$

*đổi dấu từ $\geq 0$ sang $\leq 0$ (từ $\leq 0$ sang $\geq 0$) khi đi qua $0$.*

Đối với các hàm khả vi bậc 2, sự đổi dấu thực chất là sự tăng/giảm qua 0 khi $t$ tiếp cận 0. Từ đó ta đến với điều kiện bậc 2 ở phần tiếp theo.

## 6.3 Định lý đạo hàm bậc 2

Xuất phát từ định lý 6.14 và định lý 5.29, ta có định lý sau:

**Định lý 6.15.** *Cho hàm số $f : D \to \mathbb{R}$ khả vi bậc 2. Điểm $\mathbf{x}^* \in D^o$ thỏa mãn $\nabla f\left(\mathbf{x}^*\right) = 0$ là một điểm cực tiểu địa phương nếu*

$$\mathbf{u}^T\mathbf{H}(f)\left(\mathbf{x}^*\right)\mathbf{u} > 0, \quad \forall\mathbf{u} \in \mathbb{R}^n \setminus \{0\};$$

*và là một điểm cực đại địa phương nếu*

$$\mathbf{u}^T\mathbf{H}(f)\left(\mathbf{x}^*\right)\mathbf{u} < 0, \quad \forall\mathbf{u} \in \mathbb{R}^n \setminus \{0\}.$$

*Nếu tồn tại cả $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n \setminus \{0\}$ sao cho*

$$\mathbf{u}^T\mathbf{H}(f)\left(\mathbf{x}^*\right)\mathbf{u} < 0 < \mathbf{v}^T\mathbf{H}(f)\left(\mathbf{x}^*\right)\mathbf{v},$$

*thì $\mathbf{x}^*$ là một điểm yên ngựa.*

Ta phát biểu không chứng minh định lý sau:

**Định lý 6.16.** *Cho ma trận đối xứng $M \in \mathbb{R}^{n\times n}$, khi đó*

- ⋆ $\mathbf{u}^T M\mathbf{u} > 0, \quad \forall \mathbf{u} \in \mathbb{R}^n \iff$ *mọi trị riêng của M đều dương. Ta nói M* **xác định dương**.
- ⋆ $\mathbf{u}^T M\mathbf{u} < 0, \quad \forall \mathbf{u} \in \mathbb{R}^n \iff$ *mọi trị riêng của M đều âm. Ta nói M* **xác định âm**.
- ⋆ $\exists \mathbf{u}, \mathbf{v} \in \mathbb{R} \setminus \{0\} : \mathbf{u}^T M\mathbf{u} < 0 < \mathbf{v}^T M\mathbf{v} \iff$ *M có cả trị riêng âm lẫn dương.*
- ⋆ $\mathbf{u}^T M\mathbf{u} \geq 0, \quad \forall \mathbf{u} \in \mathbb{R}^n \iff$ *mọi trị riêng của M đều không âm. Ta nói M* **nửa xác định dương**. *Nếu* $\exists \mathbf{u} \in \mathbb{R} \setminus \{0\} : \mathbf{u}^T M\mathbf{u} = 0$, *0 là một trị riêng của M.*
- ⋆ $\mathbf{u}^T M\mathbf{u} \leq 0, \quad \forall \mathbf{u} \in \mathbb{R}^n \iff$ *mọi trị riêng của M đều không dương. Ta nói M* **nửa xác định âm**. *Nếu* $\exists \mathbf{u} \in \mathbb{R} \setminus \{0\} : \mathbf{u}^T M\mathbf{u} = 0$, *0 là một trị riêng của M.*

Kết hợp 2 định lý trên, ta có định lý cuối cùng.

**Định lý 6.17** (Định lý đạo hàm bậc 2). *Cho hàm số $f : \mathbb{R}^n \to \mathbb{R}$ khả vi bậc 2 và điểm $\mathbf{x}^* \in D^o$ thỏa mãn $\nabla f(\mathbf{x}^*) = 0$.*

- ⋆ *Nếu ma trận $\mathbf{H}(f)(\mathbf{x}^*)$ xác định dương, $\mathbf{x}^*$ là một điểm cực tiểu địa phương.*
- ⋆ *Nếu ma trận $\mathbf{H}(f)(\mathbf{x}^*)$ xác định âm, $\mathbf{x}^*$ là một điểm cực đại địa phương.*
- ⋆ *Nếu ma trận $\mathbf{H}(f)(\mathbf{x}^*)$ có cả trị riêng âm lẫn dương, $\mathbf{x}^*$ là một điểm yên ngựa.*

**Ví dụ 6.18.** Ở ví dụ 6.11, $(0, 0)$ là điểm dừng của $f$ và Hessian của $f$ tại $(0, 0)$:

$$\mathbf{H}(f)(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix},$$

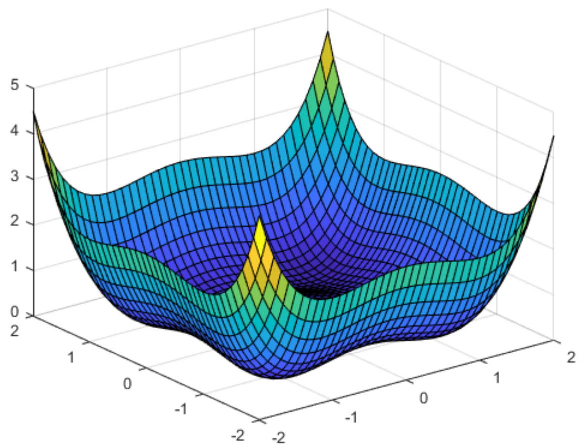
có hai trị riêng là 2 và $-2$ nên $(0, 0)$ là điểm yên ngựa của $f$

**Ví dụ 6.19.** Hàm số $f(x, y) = \frac{1}{4}\left(\left(1-x^2\right)^2 + \left(1-y^2\right)^2\right)$ có

$$\nabla f(x,y) = \left(-x(1-x^2), -y(1-y^2)\right),$$

do đó các điểm dừng của $f$ là:

$$(-1,-1), (-1,0), (-1,1), (0,-1), (0,0), (0,1), (1,-1), (1,0), (1,1).$$




**Hình 7:** Hình ảnh của hàm số xác định bởi $f(x, y) = \frac{1}{4}\left(\left(1-x^2\right)^2 + \left(1-y^2\right)^2\right)$.

Ma trận Hessian của $f$ xác định bởi:

$$\mathbf{H}(f)(x,y) = \begin{bmatrix} -1+2x^2 & 0 \\ 0 & -1+2y^2 \end{bmatrix}.$$

⋆ Bốn điểm $(-1,-1), (-1,1), (1,-1), (1,1)$ là các điểm cực tiểu bởi:

$$\mathbf{H}(f)(-1,-1) = \mathbf{H}(f)(-1,1) = \mathbf{H}(f)(1,-1) = \mathbf{H}(f)(1,1) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

⋆ $(0,0)$ là cực đại địa phương của $f$:

$$\mathbf{H}(f)(0,0) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

⋆ Bốn điểm $(0,-1), (0,1), (-1,0), (1,0)$ là các điểm yên ngựa của $f$ bởi:

$$\mathbf{H}(f)(0,-1) = \mathbf{H}(f)(0,1) = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix};$$

$$\mathbf{H}(f)(-1,0) = \mathbf{H}(f)(1,0) = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}.$$

**Nhận xét 6.20.** Định lý 6.17 không nhắc tới trường hợp ma trận Hessian là nửa xác định âm/dương với trị riêng 0. Lý do là vì những điểm với ma trận Hessian có một trong 2 tính chất này **có thể là điểm cực trị hoặc yên ngựa**.

Giả sử $\mathbf{H}(f)(\mathbf{x}^*)$ nửa xác định dương với trị riêng 0. Khi đó nếu vector $\mathbf{u}$ thỏa mãn

$$\mathbf{u}^T\mathbf{H}(f)(\mathbf{x}^*)\,\mathbf{u} > 0,$$

thì $\mathbf{x}^*$ sẽ là cực tiểu địa phương theo hướng $\mathbf{u}$. Tuy nhiên, với các vector $\mathbf{u}$ thỏa mãn

$$\mathbf{u}^T\mathbf{H}(f)(\mathbf{x}^*)\,\mathbf{u} = 0,$$

ta không thể suy ra gì (giống như phép thử đạo hàm bậc 2 với hàm 1 biến). Do đó về tổng thể ta không có kết luận gì khi ma trận Hessian nửa xác định âm/dương với trị riêng 0.

## Bài tập

**Bài tập 6.1.** Tìm miền xác định $D \subseteq \mathbb{R}^2$ của hàm hai biến $f$ và khảo sát cực trị của chúng:

⋆ $3x^2 - 9xy + \dfrac{x^3}{3} - 3y$.

⋆ $2xy - x^2$.

⋆ $x^2y - 2xy + 2y^2 - 15y$.

⋆ $x^3 - 6x^2 + 9x - 15y^3 + 5y^2 - 5$.

⋆ $\dfrac{1}{x} + \dfrac{1}{y} + 2xy$.

⋆ $e^{xy}$.

⋆ $x^2 - e^{y^2}$.

⋆ $(y-2)\ln(xy)$.

**Bài tập 6.2.** Tìm tất cả các điểm dừng (stationary point) của các hàm hai biến dưới đây (trên tập xác định). Dùng phép thử bậc hai để xác định các cực trị địa phương và điểm yên ngựa:

- $\star$ $x\ln(y) - y\ln(x)$.
- $\star$ $e^{x^2+y^2-z^2}$.
- $\star$ $e^{x^3+y^3+z^3}$.
- $\star$ $x\ln(x) + y\ln(y) + z\ln(z)$.

# Tài liệu tham khảo

[1] J. Lambers, University of Southern Mississippi. Lecture notes for MAT280.

[2] S. Angenent, J. Robbin, University of Wisconsin - Madison. Lecture notes for MATH221.

[4] Đỗ Minh Hải, dominhhai.github.io. [Giải Tích] Đạo hàm của hàm nhiều biến số

[5] Vũ Hữu Tiệp. Machine learning cơ bản.