ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC SƯ PHẠM

------------- 0 -------------

Phạm Thị Thu Trang

HÀM NHIỀU BIẾN VÀ CỰC TRỊ CỦA HÀM

Chuyên ngành: **Toán giải tích**

Mã số: **60.46.01**

LUẬN VĂN THẠC SĨ KHOA HỌC TOÁN HỌC

Thái Nguyên - 2009

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC SƯ PHẠM

------------ 0 -------------

Phạm Thị Thu Trang

HÀM NHIỀU BIẾN VÀ CỰC TRỊ CỦA HÀM

Chuyên ngành: **Toán giải tích**
Mã số: **60.46.01**

LUẬN VĂN THẠC SĨ TOÁN HỌC

Người hướng dẫn khoa học:
GS – TS Trần Vũ Thiệu

Thái Nguyên - 2009

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC SƯ PHẠM

------------ 0 -------------

Phạm Thị Thu Trang

HÀM NHIỀU BIẾN VÀ CỰC TRỊ CỦA HÀM

Chuyên ngành: **Toán giải tích**
Mã số: **60.46.01**

TÓM TẮT LUẬN VĂN THẠC SĨ TOÁN HỌC

Thái Nguyên - 2009

CÔNG TRÌNH ĐƯỢC HOÀN THÀNH TẠI
TRƯỜNG ĐẠI HỌC SƯ PHẠM –ĐẠI HỌC THÁI NGUYÊN

Người hướng dẫn khoa học : **GS.TS. Trần Vũ Thiệu**

Phản biện 1: PGS.TS ĐỖ VĂN LƯU

Phản biện 2 : GS.TSKH LÊ DŨNG MƯU

.

Luận văn đã được bảo vệ trước Hội đồng chấm luận văn họp tại:

Trường Đại học Sư phạm – Đại học Thái Nguyên

Ngày 8 tháng 11 năm 2009

Có thể tìm hiểu luận văn tại thư viện

TRƯỜNG ĐẠI HỌC SƯ PHẠM THÁI NGUYÊN

MỤC LỤC

LỜI NÓI ĐẦU

Toán học nói chung và toán giải tích nói riêng có những ứng dụng đa dạng trong nhiều ngành khoa học khác nhau, đặc biệt trong khoa học kinh tế. Các nghiên cứu và phân tích kinh tế về mặt định lượng thường được tiến hành thông qua các mô hình kinh tế toán (dùng toán học để mô tả, phân tích các mối quan hệ, các quá trình hay đối tượng kinh tế). Vì thế các nhà nghiên cứu kinh tế ngày càng có nhu cầu sử dụng nhiều hơn các công cụ toán học, đặc biệt là công cụ giải tích (như hàm số, đạo hàm, vi phân) và các phương pháp tối ưu hoá.

Đề tài luận văn đề cập tới những kiến thức toán giải tích và tối ưu hoá cơ bản cần dùng trong kinh tế. Việc tìm hiểu những kiến thức này là hoàn toàn cần thiết và hữu ích, giúp hiểu sâu hơn về các công cụ toán giải tích, tối ưu hoá và vận dụng tốt hơn trong thực tiễn giảng dạy toán cho các đối tượng kinh tế.

Mục tiêu của luận văn là tìm hiểu và trình bày khái quát những kiến thức toán học cơ bản cần dùng trong các nghiên cứu kinh tế, đặc biệt trong nghiên cứu lý thuyết kinh tế vi mô (micro-economic theory). Các nội dung đề cập tới trong luận văn được trình bày không quá hình thức mà gần gũi với tư duy kinh tế, với nhiều ví dụ minh hoạ cụ thể và có giải thích ý nghĩa kinh tế khi có thể, nhưng vẫn giữ tính chính xác, chặt chẽ về mặt toán học.

Nội dung luận văn được chia thành ba chương:

Chương 1 "**Kiến thức chuẩn bị**" giới thiệu tóm tắt một số khái niệm cơ bản về tập hợp và ánh xạ, quan hệ và hàm số: tập mở, tập đóng, tập compact trong R^n; cận trên (cận dưới) của tập hợp số thực; tính liên tục của ánh xạ, mối quan hệ giữa tính liên tục với ảnh ngược của các tập mở (đóng), ảnh liên tục của tập compact; định lý Weierstrass về tồn tại giá trị cực trị của hàm liên tục trên tập compact; tập lồi và tính chất, định lý Minkowski về tách các tập lồi ...

Chương 2 "**Hàm giá trị thực**" đề cập tới các hàm số thực thường gặp trong kinh tế và một số tập có liên quan mật thiết với hàm: đồ thị, tập mức, tập mức trên, tập mức dưới. Xét tính tăng (giảm), tính lồi (lõm), tính lồi chặt (lõm chặt), độ dốc, độ cong và mối liên hệ với các tập mức, với đạo hàm và vi phân của hàm số, hàm thuần nhất và tính chất ...

Chương 3 "**Bài toán tối ưu**" trình bày khái quát vấn đề cực trị của hàm số: cực trị địa phương và cực trị toàn cục, cực trị tự do và cực trị có điều kiện, điều kiện cần, điều kiện đủ của cực trị (cấp 1 và cấp 2). Tính duy nhất của điểm cực tiểu (cực đại) liên quan với tính lồi (lõm) chặt. của hàm. Cực trị với ràng buộc đẳng thức (phương pháp Lagrange), với ràng buộc không âm và tổng quát hơn là với ràng buộc bất đẳng thức (điều kiện Karush-Kuhn-Tucker (điều kiện KKT) ...

Do thời gian có hạn nên luận văn này mới chỉ dừng lại ở việc tìm hiểu, tập hợp tài liệu, sắp xếp và trình bày các nội dung chính theo chủ đề đặt ra. Trong quá trình viết luận văn cũng như trong xử lý văn bản chắc chắn không tránh khỏi có những sai sót nhất định. Tác giả luận văn rất mong nhận được sự góp ý của các thầy cô và các bạn đồng nghiệp để luận văn được hoàn thiện hơn.

Nhân dịp này, tác giả xin bày tỏ lòng biết ơn sâu sắc đến thầy hướng dẫn GS-TS Trần Vũ Thiệu đã tận tình giúp đỡ trong suốt quá trình làm luận văn.

Tác giả xin chân thành cảm ơn các thày, cô của Trường Đại học Sư phạm Thái Nguyên và Viện Toán học đã giảng dạy và tạo mọi điều kiện thuận lợi trong quá trình tác giả học tập và nghiên cứu.

Thái Nguyên, tháng 9/2009

Tác giả

Chương 1
KIẾN THỨC CHUẨN BỊ

Chương này đề cập tới một số khái niệm cơ bản về giải tích liên quan tới các hàm và cực trị của hàm. Nội dung của chương dựa chủ yếu trên các nguồn tài liệu [2], [3], [4].

1.1. TẬP LỒI TRONG $\mathbb{R}^n$ (Convex sets in $\mathbb{R}^n$)

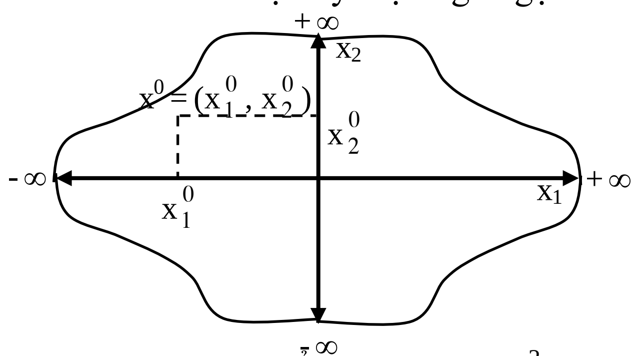
Tập số thực được biểu thị bởi ký hiệu đặc biệt $\mathbb{R}$ và được định nghĩa như sau

$$\mathbb{R} \equiv \{x \mid -\infty < x < +\infty\}.$$

Nếu ta xây dựng tích của hai tập hợp

$$\mathbb{R} \times \mathbb{R} \equiv \{(x_1, x_2) \mid x_1 \in \mathbb{R}, x_2 \in \mathbb{R}\}$$

thì một điểm bất kỳ thuộc tập này (cặp hai số thực bất kỳ) được đồng nhất với một điểm trong mặt phẳng Descarte vẽ ở Hình 1.1. Tập $\mathbb{R} \times \mathbb{R}$ đôi khi được gọi là "không gian Euclid hai chiều" và được ký hiệu ngắn gọn bởi $\mathbb{R}^2$.



Hình 1.1. Mặt phẳng Descarte $\mathbb{R}^2$

Tổng quát, véctơ n- chiều là một cặp có thứ tự của n số $(x_1, x_2, \dots, x_n)$ và được xem như một "**điểm**" trong không gian Euclid n - chiều hay "n - không gian". Cũng như trước, n - không gian được định nghĩa như tích của n tập hợp

$$\mathbb{R}^n \equiv \underbrace{\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}}_{\text{n lần}} \equiv \{(x_1, x_2, \dots, x_n) \mid x_i \in \mathbb{R}, i = 1, 2, \dots, n\}.$$

Ta sẽ thường ký hiệu các véctơ (hay điểm) trong $\mathbb{R}^n$ bằng chữ in đậm. Ví dụ, $\mathbf{x} \equiv \{x_1, x_2, \dots, x_n\}$. Đôi khi ta muốn thu hẹp sự chú ý vào tập con của $\mathbb{R}^n$, gọi là "**góc không âm**" và ký hiệu $\mathbb{R}^n_+$, trong đó

$$\mathbb{R}^n_+ \equiv \{(x_1, x_2, \ldots, x_n) \mid x_i \geq 0, i = 1, 2, \ldots, n\} \subset \mathbb{R}^n.$$

Ta qui ước viết $\mathbf{x} \geq 0$ để chỉ các véctơ trong $\mathbb{R}^n_+$ mà mỗi thành phần x_i của nó lớn hơn hay bằng 0 và dùng ký hiệu $\mathbf{x} > 0$ để chỉ các véctơ mà mọi thành phần của nó thực sự dương. Tổng quát, với bất kỳ $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, ta viết $\mathbf{x} \geq \mathbf{y} \Leftrightarrow x_i \geq y_i, \forall i = 1, \ldots, n$, và $\mathbf{x} > \mathbf{y} \Leftrightarrow x_i > y_i, \forall i = 1, \ldots, n$.

Định nghĩa 1.1. Tập hợp lồi trong $\mathbb{R}^n$

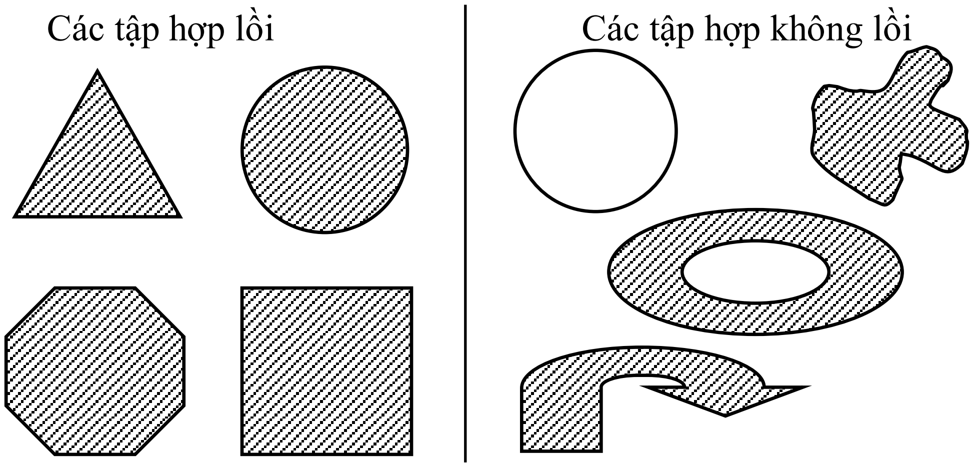
Tập $S \subset \mathbb{R}^n$ được gọi là **lồi** nếu với mọi $x^1 \in S$ và $x^2 \in S$ ta có

$$tx^1 + (1 - t)x^2 \in S.$$

đối với mọi t trong khoảng $0 \leq t \leq 1$.

Như vậy một tập hợp là lồi nếu nó chứa hai điểm bất kỳ thì nó chứa tất cả các điểm trung bình theo trọng số (tổng trọng số bằng 1) của hai điểm đó.

Các **ví dụ** về tập lồi và tập không lồi vẽ ở Hình 1.2. Các tập hợp lồi có hình dáng đẹp: không có hố, không nứt gẫy, không bị cong queo trên biên.



Hình 1.2. Các tập lồi và tập không lồi trong $\mathbb{R}^2$

Ta chú ý tới tính chất đơn giản nhưng quan trọng của các tập lồi.

Định lý 1.1. Giao của các tập lồi là lồi

Giả sử S *và* T *là các tập lồi trong* $\mathbb{R}^n$. *Khi đó,* $S \cap T$ *là một tập lồi.*

Chứng minh. Giả sử S và T là hai tập hợp lồi và $\mathbf{x}^1$, $\mathbf{x}^2$ là hai điểm bất kỳ thuộc $S \cap T$. Do $\mathbf{x}^1 \in S \cap T$ nên $\mathbf{x}^1 \in S$ và $\mathbf{x}^1 \in T$. Cũng cậy, do $\mathbf{x}^2 \in S \cap T$ nên $\mathbf{x}^2 \in S$ và $\mathbf{x}^2 \in T$. Cho $\mathbf{z} = t\mathbf{x}^1 + (1 - t)\mathbf{x}^2$ với $t \in [0, 1]$ là một tổ hợp lồi bất kỳ của $\mathbf{x}^1$ và $\mathbf{x}^2$. Do S là tập lồi nên $\mathbf{z} \in S$ và do T là tập lồi nên $\mathbf{z} \in T$. Vì $\mathbf{z} \in S$ và $\mathbf{z} \in T$ nên $\mathbf{z} \in S \cap T$. Do mọi tổ hợp lồi của hai điểm bất kỳ thuộc $S \cap T$ cũng thuộc $S \cap T$ nên $S \cap T$ là một tập hợp lồi.

1.2. QUAN HỆ VÀ HÀM SỐ(Relations and Functions)

• Ta đã thấy mỗi cặp có thứ tự (s, t) tuỳ ý đặt tương ứng phần tử $s \in S$ nào đó với phần tử $t \in T$. Các phần tử của S và T không nhất thiết là các số mà có thể là những đối tượng bất kỳ (người, vật hay đồ vật, …). Ta nói một **họ** hay một tập tuỳ ý các cặp có thứ tự là một **quan hệ nhị nguyên** (binary relation) của hai tập S và T. Như vậy, *quan hệ nhị nguyên là một tập hợp con của tích hai tập, trong đó phần tử đầu của mỗi cặp thuộc* S *và phần tử sau thuộc* T.

Thông thường, họ các cặp được thiết lập khi giữa hai phần tử của cặp có mối quan hệ ý nghĩa nào đó. Chẳng hạn, S là tập các thành phố {Hà Nội, Wasington, London, Paris, Marseilles, Huế} và T là tập các nước {Việt Nam, Hoa Kỳ, Anh, Pháp, Đức}. Cụm từ "**là thủ đô của**" xác định nên một quan hệ mà nó là tập con của tập tích S × T, bao gồm các cặp {(Hà Nội, Việt Nam), (Wasington, Hoa Kỳ), (London, Anh), (Paris, Pháp)}. Ta thường đặt một ký hiệu chung để chỉ quan hệ, thay cho bản thân quan hệ đó và cả cụm từ "là thủ đô của".

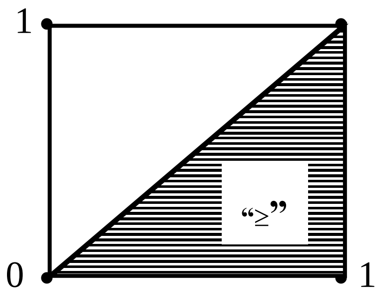
Ký hiệu R để chỉ cụm từ "có quan hệ ý nghĩa nào đó với". Ta nói R xác định một quan hệ và đọc xRy là "x có quan hệ với y". Để phân biệt giữa tập tất cả các cặp có quan hệ bởi cụm từ R với bản thân phát biểu R đó, ta đặt ký hiệu xác định quan hệ đó trong hai dấu nháy kép. Như vậy, định nghĩa tổng quát của một quan hệ được cho bởi

$$\text{“R”} \equiv \{(s, t) \mid s \in S, t \in T \text{ và } sR\,t\} \subset S \times T.$$

Hay gặp nhất là các quan hệ nhị nguyên xác định bởi một tập con của tích một tập hợp nào đó với chính nó. Chẳng hạn, S là tập các điểm thuộc khoảng đóng đơn vị S = [0, 1]. Với cụm từ có nghĩa R ≡ "lớn hơn hay bằng" thì quan hệ nhị nguyên

$$\text{"}\geq\text{"} \equiv \{(x, y) \mid x \in S, y \in S \text{ và } x \geq y\}$$

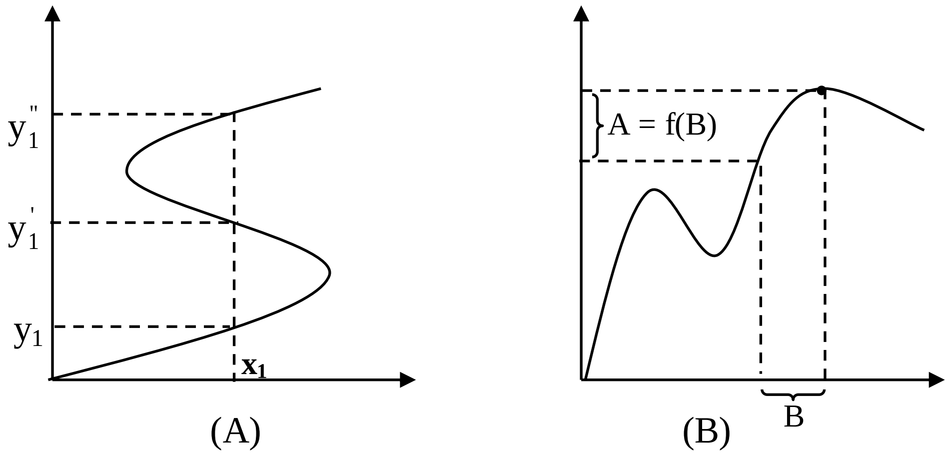
được minh hoạ ở Hình 1.3. Quan hệ này bao gồm mọi cặp có thứ tự các số giữa 0 và 1, trong đó số thứ nhất lớn hơn hay bằng số thứ hai. Khi quan hệ nhị nguyên là tập con của tích một tập S với chính nó thì ta nói đó là một **quan hệ trên S**.



$S = \{0, 1\}$
$S \times S = \{(x, y) \mid x \in S, y \in S\}$
$\text{"}\geq\text{"} = \{(x, y) \mid x \in S, y \in S, x \geq y\}$
$\text{"}\geq\text{"} \subset S \times S$

Hình 1.3. Quan hệ "≥" trên S = [0, 1]

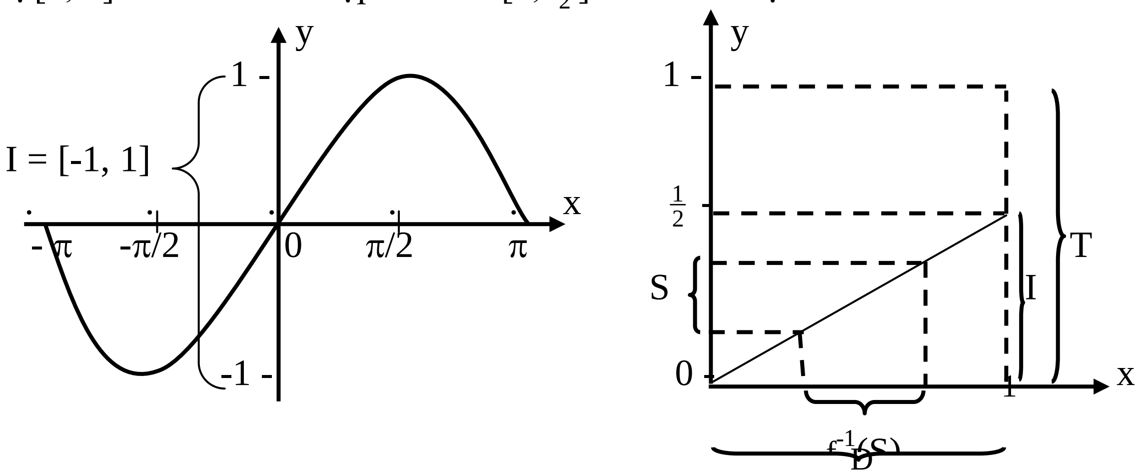
• **Hàm** (function) cũng là một quan hệ và là một kiểu quan hệ hết sức đặc biệt. Cụ thể, hàm là quan hệ đặt tương ứng mỗi phần tử của một tập với một phần tử duy nhất của một tập khác. Ta nói hàm f là một ánh xạ (mapping) từ một tập D vào một tập khác T và viết $f : D \to T$. Tập D các phần tử có ánh xạ từ đó gọi là **miền xác định** (domain) và tập T các phần tử được ánh xạ chuyển tới được gọi là **miền trị** (range). Nếu y là một điểm thuộc miền trị được ánh xạ chuyển tới từ một điểm x thuộc miền xác định thì ta viết y = f(x) và gọi y là ảnh (image) của x. Nếu tập điểm A trong miền trị được ánh xạ tới bởi tập điểm B trong miền xác định thì ta viết A = f(B). Để minh hoạ, ta xét Hình 1.4. Hình vẽ (A) **không** phải là một hàm, vì nhiều điểm trong miền trị được gắn với cùng một điểm trong miền xác định, x_1 chẳng hạn. Hình vẽ (B) mô tả một hàm, vì mỗi điểm thuộc miền xác định được gắn với một điểm duy nhất trong miền trị.



(A) (B)

Hình 1.4. Hàm và không phải hàm

Ảnh của D là tập điểm trong miền trị mà có một điểm thuộc miền xác định ánh xạ tới đó, tức là tập I ≡ f(D) = {y | y = f(x) với x nào đó ∈ D} ⊂ T. **Ảnh ngược** của tập điểm S ⊂ I được định nghĩa là tập $f^{-1}(S)$ ≡ {x | x ∈ D, f(x) ∈ S}. **Đồ thị** của hàm f hiểu theo nghĩa thông thường, đó là tập các cặp có thứ tự G ≡ {(**x**, y) | **x** ∈ D, y = f(**x**)}. Một số khái niệm về đồ thị được minh hoạ ở Hình 1.5. ở Hình 1.5 (A), D = ℝ, T = ℝ và nó mô tả đồ thị của hàm y = sin(x). Tuy nhiên, hàm sin(x) không bao giờ lấy giá trị nhỏ hơn - 1 và lớn hơn 1. Vì thế ảnh của D là tập con I = {-1, 1} của miền trị T. Hình 1.5 (B) là đồ thị của hàm f : [0,1] → [0, 1] cho bởi y = $\frac{1}{2}$x. ở đây ta giới hạn miền xác định và miền trị trong khoảng đơn vị [0, 1]. ảnh của D là tập con I = [0, $\frac{1}{2}$] của miền trị.



(A) (B)

Hình 1.5. Miền hữu hiệu, miền trị và miền ảnh (image)

Hình 1.5 (A) cho thấy trong định nghĩa của hàm không ngăn cấm có nhiều phần tử trong miền xác định ánh xạ vào cùng một phần tử trong miền trị. Nếu mỗi điểm trong miền trị được gắn tối đa với một điểm trong miền xác định thì hàm được gọi là ánh xạ **một-một**. Thêm vào đó, nếu mỗi điểm trong miền trị đều là ảnh của một điểm nào đó trong miền xác định thì hàm được gọi là ánh xạ **lên**. Nếu hàm là ánh xạ 1 - 1 lên thì **hàm ngược** $f^{-1} : T \rightarrow D$ tồn tại, cũng là ánh xạ 1 - 1 lên.

1.3.TÔ PÔ TRONG $\mathbb{R}^n$

• Mục này đề cập tới một số khái niệm cơ bản về tôpô và thiết lập một số kết quả quan trọng về tập hợp và về ánh xạ liên tục từ một tập vào một tập khác. Mặc dù nhiều khái niệm đề cập tới ở đây có thể mở rộng cho các loại tập bất kỳ, song ta chỉ hạn chế xét các tập trong $\mathbb{R}^n$, tức là tập số thực hay tập véctơ thực.

Ta bắt đầu bằng khái niệm **metric** và **không gian metric** (metric space). Mêtric hiểu đơn giản là số đo khoảng cách (distance). Không gian metric chính là một tập, trong đó có định nghĩa khái niệm khoảng cách giữa các phần tử của tập đó. Đường thẳng số thực $\mathbb{R}$ là một không gian metric. Khoảng cách hay metric trong $\mathbb{R}$ chính là hàm giá trị tuyệt đối. Với hai điểm x^1, x^2 bất kỳ thuộc $\mathbb{R}$ khoảng cách giữa chúng, ký hiệu $d(x^1, x^2)$ được cho bởi

$$d(x^1, x^2) = |x^1 - x^2|.$$

Mặt phẳng Descarte $\mathbb{R}^2$ cũng là một không gian metric. Khoảng cách giữa hai điểm tuỳ ý $\mathbf{x}^1 = (x_1^1, x_2^1)$ và $\mathbf{x}^2 = (x_1^2, x_2^2)$ trong $\mathbb{R}^2$ được cho bởi

$$d(\mathbf{x}^1, \mathbf{x}^2) = \sqrt{(x_1^2 - x_1^1)^2 + (x_2^2 - x_2^1)^2}\ .$$

Tổng quát, với hai điểm bất kỳ $\mathbf{x}^1$ và $\mathbf{x}^2$ trong $\mathbb{R}^n$ ta định nghĩa

$$d(\mathbf{x}^1, \mathbf{x}^2) = \sqrt{(x_1^1 - x_1^2)^2 + (x_2^1 - x_2^2)^2 + \ldots + (x_n^1 - x_n^2)^2}\ .$$

Để cho gọn ta dùng ký hiệu $d(\mathbf{x}^1, \mathbf{x}^2) = \|\mathbf{x}^1 - \mathbf{x}^2\|$. Ta gọi đó là **chuẩn** (**metric**)

Euclid. Cũng là lẽ tự nhiên, ta gọi không gian metric $\mathbb{R}^n$ sử dụng chuẩn này để đo khoảng cách là **không gian Euclid** $\mathbb{R}^n$.

Khi có metric, ta có thể đưa ra khái niệm "**gần nhau**" của hai điểm. Ta lấy điểm bất kỳ $x^0 \in \mathbb{R}^n$ và gọi tập điểm có khoảng cách tới x^0 nhỏ hơn $\varepsilon > 0$ là một **ε-hình cầu mở** tâm x^0. Tập điểm có khoảng cách tới x^0 không quá $\varepsilon > 0$ là một **ε-hình cầu đóng** tâm x^0. Nói một cách chính xác, ta có

Định nghĩa 1.2. Hình cầu bán kính ε mở và đóng (open & closed ε-balls)

1. Hình cầu mở tâm tại điểm $x^0 \in \mathbb{R}^n$ và bán kính $\varepsilon > 0$ (ε là một số thực) là tập các điểm trong $\mathbb{R}^n$:

$$B_\varepsilon(x^0) \equiv \{x \in \mathbb{R}^n \mid \underbrace{d(x^0, x) < \varepsilon}_{\text{nhỏ hơn hẳn}}\}$$

2. Hình cầu đóng tâm tại điểm $x^0 \in \mathbb{R}^n$ và bán kính $\varepsilon > 0$ là tập các điểm trong $\mathbb{R}^n$:

$$\overline{B}_\varepsilon(x^0) \equiv \{x \in \mathbb{R}^n \mid \underbrace{d(x^0, x) \le \varepsilon}_{\text{nhỏ hơn hay bằng}}\}$$

Các khoảng mở và khoảng đóng trên đường thẳng số thực là các tập có những tính chất hoàn toàn khác nhau. Trong $\mathbb{R}$ ta có một cảm nhận trực quan khá tốt về sự khác nhau đó. Khái niệm ε-hình cầu cho phép ta hình thức hoá sự khác biệt này và tổng quát hoá nó để có thể áp dụng được cho những tập trong không gian số chiều cao hơn.

• Dưới đây ta sẽ dùng khái niệm ε-hình cầu để định nghĩa tập mở, tập đóng và thiết lập một số tính chất quan trọng của chúng.

Định nghĩa 1.3. Tập mở trong $\mathbb{R}^n$ (open set)

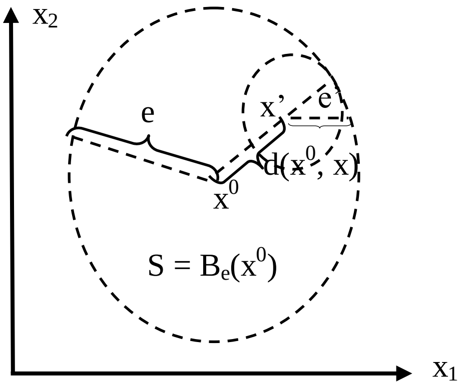
Ta nói tập $S \subset \mathbb{R}^n$ là **mở** nếu với mỗi $x \in S$ tồn tại $\varepsilon > 0$ sao cho hình cầu mở $B_\varepsilon(x) \subset S$. Nói nôm na, tập S là mở nếu ta có thể vẽ trong S một hình cầu mở, dù to hay nhỏ, bao quanh một điểm bất kỳ thuộc S.

Định lý 1.2. Về các tập mở trong $\mathbb{R}^n$

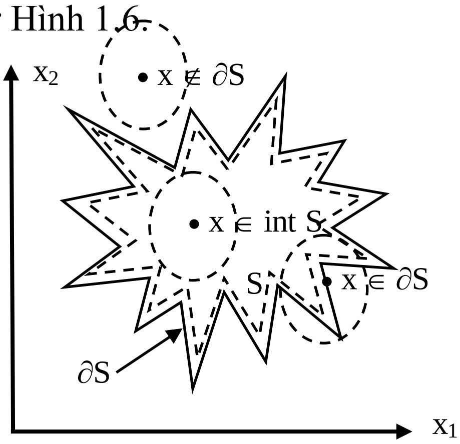
1. *Tập rỗng ∅ là một tập mở.*
2. *Toàn không gian $\mathbb{R}^n$ là một tập mở.*
3. *Hợp của hai (hay một số bất kỳ) tập mở là một tập mở*
4. *Giao của một số hữu hạn bất kỳ các tập mở là một tập mở.*

Chứng minh. (1) hiển nhiên, vì tập ∅ không chứa phần tử nào. (2) cũng là tự nhiên, vì $B_\varepsilon(x) \subset \mathbb{R}^n$ $\forall x \in \mathbb{R}^n$ và $\forall \varepsilon > 0$. Để chứng minh (3) giả sử A, B là các tập mở, ta chứng minh $A \cup B$ cũng là tập mở. Thật vậy, với $x \in A \cup B$ thì $x \in A$ hoặc $x \in B$. Nếu $x \in A$ thì do A mở nên tìm được $\varepsilon > 0$ sao cho $B_\varepsilon(x) \subset A$. Nếu $x \in B$ thì do B mở nên tìm được $\varepsilon' > 0$ sao cho $B_{\varepsilon'}(x) \subset B$. Trong mọi trường hợp, với bất kỳ $x \in A \cup B$ ta luôn tìm được một hình cầu mở tâm x nằm trọn trong $A \cup B$, vì thế $A \cup B$ là tập mở. Chứng minh (4) tương tự.

Các tập mở có những tính chất lý thú và hữu ích. Tập mở luôn có thể được mô tả chính xác bởi họ các tập mở khác nhau! Giả sử ta bắt đầu từ một tập mở nào đó. Vì tập là mở nên ta có thể "bọc" mỗi điểm của tập này bởi một hình cầu mở sao cho mọi điểm thuộc hình cầu đều nằm trong tập đã chọn. Bản thân mỗi hình cầu mở lại là một tập mở, như minh hoạ ở Hình 1.6.



Hình 1.6. Hình cầu mở là tập mở



Hình 1.7. Tập mở/ đóng trong $\mathbb{R}^2$

Bây giờ xét hợp của tất cả các hình cầu mở này. Theo Định lý 1.2, hợp đó là một tập mở. Có thể thấy rằng trên thực tế hai tập này là một. Tính chất này của các tập mở là rất quan trọng đủ để chứng tỏ vai trò của định lý sau.

Định lý 1.3. Mọi tập mở là họ của các hình cầu mở

Giả sử $S \subset \mathbb{R}^n$ *là một tập mở. Với mỗi* $x \in S$ chọn số $\varepsilon_x > 0$ *sao cho* $B_{\varepsilon_x}(x) \subset S$. *Khi đó*

$$S = \bigcup_{x \in S} B_{\varepsilon_x}(x).$$

• Ta dùng tập mở để định nghĩa tập đóng.

Định nghĩa 1.4. Tập đóng trong $\mathbb{R}^n$

Ta nói tập $S \subset R^n$ là **đóng** khi và chỉ khi phần bù $cS = (\mathbb{R}^n \setminus S)$ là tập mở.

Nói nôm na, một tập là mở nếu nó không chứa điểm nào trên "**biên**" của nó và là tập đóng nếu nó chứa mọi điểm trên biên của nó. Chính xác hơn, điểm x được gọi là **điểm biên** của tập S nếu mọi ε-hình cầu tâm x đều chứa những điểm thuộc S và những điểm không thuộc S. Tập các điểm biên của S được ký hiệu là ∂S. Tập S là mở nếu nó không chứa điểm biên nào của nó hay nếu $\partial S \cap S = \varnothing$. Tập S là đóng nếu nó chứa mọi điểm biên của nó hay nếu $\partial S \subset S$.

Cho một tập bất kỳ $S \subset \mathbb{R}^n$. Điểm $x \in S$ gọi là **điểm trong** của S nếu tìm được ε-hình cầu tâm x nằm trọn trong S: $B_\varepsilon(x) \subset S$. Tập tất cả các điểm trong của S gọi là **phần trong** của S và được ký hiệu là *int* S. Theo cách này ta thấy rằng tập S là mở nếu nó chỉ chứa các điểm trong, tức là nếu S = *int* S. Trái lại, tập S là đóng nếu nó chứa mọi điểm trong cùng với mọi điểm biên của nó, tức là nếu $S = \mathit{int}\, S \cup \partial S$.

Tập đóng có các tính chất tương tự như tính chất tập mở nêu trong Định lý 1.2.

Định lý 1.4. Về các tập đóng trong $\mathbb{R}^n$

1. *Tập rỗng ∅ là một tập đóng.*
2. *Toàn không gian* $\mathbb{R}^n$ *là một tập đóng.*
3. *Hợp của một số hữu hạn bất kỳ các tập đóng là một tập đóng.*
4. *Giao của hai (hay một số bất kỳ) tập đóng là một tập đóng.*

Chứng minh. Tập rỗng $\varnothing$ và toàn $|R^n$ là hai tập duy nhất vừa đóng vừa mở trong $\mathbb{R}^n$. Theo Định lý 1.2 hai tập này là mở. Do trong $\mathbb{R}^n$ tập này là phần bù của tập kia nên chúng là các tập đóng.

Để chứng minh (3) giả sử A, B là hai tập đóng. Ta chứng minh $A \cup B$ cũng là tập đóng. Thật vậy, do A, B đóng nên theo Định nghĩa 1.4 các phần bù cA, cB của chúng là các tập mở. Theo Định lý 1.2. tương giao $cA \cap cB$ là tập mở. Luật De Morgan và định nghĩa tập đóng cho thấy $c(cA \cap cB) = A \cup B$ là tập đóng. Chứng minh (4) tương tự.

Các tập đóng trên đường thẳng thực có một tính chất khá đặc thù và thực tế nó tỏ ra rất hữu ích: một tập đóng bất kỳ trong $\mathbb{R}$ có thể xem như tương giao (hữu hạn hay vô hạn) của hợp các khoảng đóng đơn giản.

Chính xác hơn, có thể chứng minh định lý sau.

Định lý 1.5. Các tập đóng trong $\mathbb{R}$ và các khoảng đóng

Giả sử S *là một tập đóng bất kỳ trong* $\mathbb{R}$. *Khi đó,*

$$S = \bigcap_{i\in I}\left((-\infty, a_i] \cup [b_i, +\infty)\right).$$

với các số thực $a_i < b_i$ *và tập chỉ số* I *nào đó.*

Định lý 1.5 cũng đúng cho các tập đóng gồm các số thực không âm. Ta có định lý sau đây.

Định lý 1.6. Các tập đóng trong $\mathbb{R}_+$ và các khoảng đóng

Giả sử S *là một tập đóng bất kỳ trong* $\mathbb{R}_+$. *Khi đó,*

$$S = \bigcap_{i\in I}\left([0, a_i] \cup [b_i, +\infty)\right).$$

với các số thực $0 \le a_i < b_i$ *và tập chỉ số* I *nào đó.*

• Một khái niệm quan trọng khác là **tập bị chặn**. Nói nôm na, tập là bị chặn nếu nó không "đi ra vô hạn". Sau đây là định nghĩa chính xác của khái niệm này.

Định nghĩa 1.5. Tập bị chặn trong $\mathbb{R}^n$ (bounded set)

Tập $S \subset \mathbb{R}^n$ được gọi là **bị chặn** nếu nó chứa được trong một hình cầu (mở hay đóng) bán kính ε nào đó. Nghĩa là, S bị chăn nếu $\exists x \in \mathbb{R}^n$ và số $\varepsilon > 0$ để $S \subset B_\varepsilon(x)$.

Theo định nghĩa này, một tập là bị chặn nếu ta có thể vẽ một ε-hình cầu bao quanh tập đó. Có một cách định nghĩa khác với nội dung trực quan hơn khi ta hạn chế ở hình cầu tâm tại gốc $0 \in \mathbb{R}^n$. Theo cách này có thể thấy tập S bị chặn khi và chỉ khi có một e > 0 hữu hạn sao cho mọi điểm trong S cách gốc không quá ε.

Có một số thuật ngữ liên quan tới các tập bị chặn trên đường thẳng thực $\mathbb{R}$. Giả sử $S \subset \mathbb{R}$ là một tập số thực khác rỗng bất kỳ. Một số thực l bất kỳ (không nhất thiết thuộc S) thoả mãn $l \le x$ với mọi $x \in S$ được gọi là một **cận dưới** (lower bound) của S. Chẳng hạn, nếu S = {3, 5, 7} thì số $0 \notin S$ là một cận dưới của S, số $3 \in S$ cũng là một cận dưới của S. Cũng vậy, một số thực u bất kỳ (không nhất thiết rhuộc S) sao cho $x \le u$ với mọi $x \in S$ được gọi là một **cận trên** (upper bound) của S. Trong ví dụ vừa xét $8 \notin S$ là một cận trên của S, số $7 \in S$ cũng là một cận trên của S. Tập $S \subset \mathbb{R}$ gọi là **bị chặn dưới** nếu nó có một cận dưới và **bị chặn trên** nếu nó có một cận trên. Khoảng (- ∞, 3) bị chặn trên nhưng không bị chặn dưới. Tập số bất kỳ vừa bị chặn trên vừa bị chặn dưới tất nhiên bị chặn theo Định nghĩa 1.5.

Ta vừa thấy một tập hợp số có thể có nhiều cận dưới hay cận trên. Số lớn nhất trong các cận dưới này gọi là **cận dưới lớn nhất** (greatest lower bound) hay cận dưới **đúng** của tập S. Số nhỏ nhất trong các cận trên gọi là **cận trên nhỏ nhất** (least upper bound) hay cận trên **đúng** của tập S. Có thể dùng tiên đề cơ bản của hệ thống số thực để chứng minh rằng một tập bị chặn bất kỳ trong $\mathbb{R}$ luôn có một cận dưới lớn nhất và một cận trên nhỏ nhất

Có thể chứng minh rằng một tập đóng bất kỳ trong $\mathbb{R}$ sẽ chứa cận dưới lớn nhất và cận trên nhỏ nhất của nó (nếu có).

Trái lại, một tập mở bất kỳ trong $\mathbb{R}$ sẽ không chứa cận dưới lớn nhất và cận trên nhỏ nhất của nó.

Định lý 1.7. Cận trên và cận dưới của tập hợp số thực

1. *Giả sử* $S \subset \mathbb{R}$ *là một tập mở bị chặn và giả sử* a *là cận dưới lớn nhất của* S *và* b *là cận trên nhỏ nhất của* S. *Khi đó,* $a \notin S$ *và* $b \notin S$.
2. *Giả sử* $S \subset \mathbb{R}$ *là một tập đóng bị chặn và giả sử* a *là cận dưới lớn nhất của* S *và* b *là cận trên nhỏ nhất của* S. *Khi đó,* $a \in S$ *và* $b \in S$.

Chứng minh. Ta chứng minh các kết luận đầu, phần sau chứng minh tương tự

Giả sử $S \subset \mathbb{R}$ là tập mở các số thực và a là cận dưới lớn nhất của S. Định lý khẳng định $a \notin S$. Nếu giả sử $a \in S$ ta sẽ tìm ra mâu thuẫn. Thật vậy, do giả thiết $a \in S$ và S là tập mở nên tìm được $\varepsilon > 0$ sao cho $B_\varepsilon(a) \subset S$. Từ đó điểm $a - \varepsilon/2 \in S$. Do $a - \varepsilon/2 < a$ và $a - \varepsilon/2 \in S$ nên điều này trái với a là cận dưới lớn nhất của S. Vì thế không thể có $a \in S$ mà phải có $a \notin S$..

Để chứng minh định lý cho trường hợp tập đóng, giả sử $S \subset \mathbb{R}$ là tập đóng, bị chặn và a là cận dưới lớn nhất của S. Theo định nghĩa cận dưới, $a \le x\ \forall x \in S$. Nếu $a = x$ với x nào đó $\in S$ thì $a \in S$ và chứng minh kết thúc.

Nếu $a < x\ \forall x \in S$ thì $a \notin S$, vì thế $a \in cS$ (phần bù của S). Do S đóng nên cS mở. Khi đó tìm được $\varepsilon > 0$ sao cho mọi điểm thuộc hình cầu mở $B_\varepsilon(a) = (a- \varepsilon, a+\varepsilon)$ chứa trong cS. Từ đó cho thấy mọi điểm thuộc khoảng mở $(a - \varepsilon, a + \varepsilon)$ đều thực sự nhỏ hơn mọi điểm trong S. Nói riêng, điểm $a + \varepsilon/2 \in (a - \varepsilon, a + \varepsilon)$ và $a + \varepsilon/2 < x\ \forall x \in S$, nghĩa là $a + \varepsilon/2$ là một cận dưới của S và $a < a + \varepsilon/2$, trái với a là cận dưới lớn nhất của S. Vậy ta phải có $a \in S$.

Một tập trong $\mathbb{R}^n$ vừa đóng, vừa bị chặn được gọi là một tập **compact.** Các tập này khá quen thuộc trong nhiều ứng dụng. Ta nhắc lại định nghĩa để dùng sau này.

Định nghĩa 1.6 (Heine - Borel). Tập compact trong $\mathbb{R}^n$

Tập $S \subset \mathbb{R}^n$ được gọi là **compact** khi và chỉ khi S đóng và bị chặn.

Khoảng mở trong $\mathbb{R}$ không phải là một tập compact. Nó có thể bị chặn nhưng không đóng. Cũng vậy, hình cầu mở trong $\mathbb{R}^n$ không compact. Tuy nhiên mọi khoảng đóng bị chặn trong $\mathbb{R}$, cũng như mọi hình cầu đóng trong $\mathbb{R}^n$ là một tập compact. Toàn bộ $\mathbb{R}^n$ không compact vì nó không bị chặn, mặc dù nó đóng. Tính compact thực ra là một tính chất tôpô. Tuy nhiên, Định lý Heine-Borel cho thấy đối với các tập trong $\mathbb{R}^n$ tính chất compact tương đương với tính đóng và bị chặn.

1.4. TÍNH LIÊN TỤC(Continuity)

Khái niệm **ánh xạ liên tục** (continuous mapping) hay **hàm liên tục** (continuous function) là một khái niệm quan trọng trong giải tích. Trong nhiều ứng dụng kinh tế hoặc ta muốn giả thiết các hàm đề cập tới là hàm liên tục hoặc muốn biết liệu chúng có liên tục khi mà ta không muốn đơn giản chỉ là giả thiết nó. Dù trường hợp nào đi nữa, tốt nhất là nên có hiểu biết rõ thế nào là hàm liên tục và các tính chất của hàm liên tục.

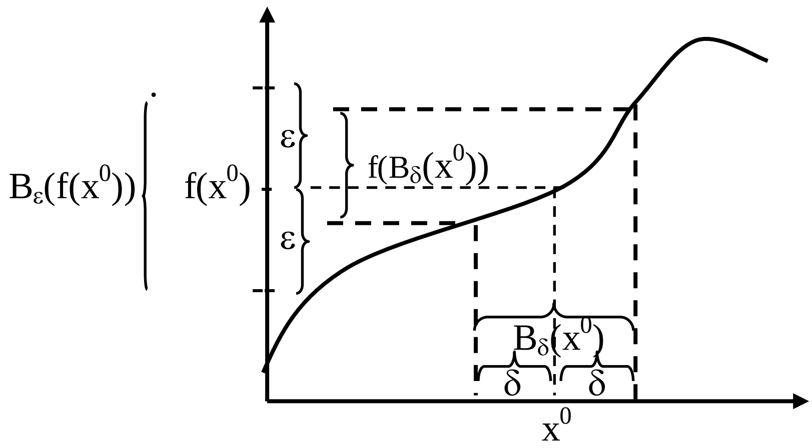
Về đại thể, một hàm gọi là **liên tục** nếu một "di chuyển nhỏ" trong miền xác định không gây ra "bước nhảy lớn" trong miền trị. Cụ thể hơn, một hàm gọi là liên tục tại điểm x^0 trong miền xác định nếu với mọi $\varepsilon > 0$ tìm được $\delta > 0$ sao cho mọi điểm trong miền xác định, cách x^0 không quá δ được ánh xạ f chuyển tới một điểm trong miền trị, cách $f(x^0)$ không quá ε. Định nghĩa sau đây cho cách hiểu chính xác về ánh xạ liên tục, áp dụng cho các ánh xạ từ tập D bất kỳ vào tập T bất kỳ khác, không nhất thiết trong không gian Euclid mà trong các không gian mêtric bất kỳ.

Định nghĩa 1.7. (Cauchy) Tính liên tục (Continuity)

Cho D là một tập, T là một tập khác và giả sử $f : D \to T$. Hàm f được gọi là liên tục tại điểm $x^0 \in D$ khi và chỉ khi với mọi $\varepsilon > 0$ tồn tại một $\delta > 0$ sao cho

$$f(B_\delta(x^0)) \subset B_\varepsilon(f(x^0)).$$

Hàm f được gọi là **liên tục** (trên D) nếu nó liên tục tại mọi điểm $x \in D$.



Hình 1.8. Tính liên tục của hàm (ánh xạ)

Định nghĩa liên tục nêu trên tập trung chủ yếu vào quan hệ giữa hai tập: tâp $f(B_\delta(x^0))$ (ảnh của tập mở trong miền xác định) và tập mở khác - tập $B_\varepsilon(f(x^0))$, cả hai tập này đều ở trong miền ảnh.

Hai định lý sau thiết lập sự tương đương giữa tính liên tục của ánh xạ với sự bảo toàn các tính chất tôpô cơ bản của các tập ảnh ngược.

Định lý 1.8. Tính liên tục và ảnh ngược của các tập mở

Giả sử $f : D \to T$ *là một ánh xạ và* $f^{-1} : T \to D$ *là ánh xạ ngược của* f *từ* T *tới* D. *Giả sư* $U \subset T$ *là một tập mở trong miền trị của* f. *Khi đó,* f *là liên tục khi và chỉ khi ảnh ngược* $f^{-1}(U) \subset D$ *là một tập mở trong miền xác định của* f.

Chứng minh. Cần. Giả sử f là ánh xạ liên tục thoả mãn Định nghĩa 1.7. Cho $U \subset T$ là tập mở bất kỳ trong miền trị của f. Xét tập ảnh ngược $f^{-1}(U)$ của U trong miền xác định của f. Ta chứng minh $f^{-1}(U)$ mở. Thật vậy, lấy điểm bất kỳ $x \in f^{-1}U) \subset D$. Theo định nghĩa của ảnh ngược $f(x) \in U$. Do U mở nên có $\varepsilon > 0$ sao cho hình cầu mở $B_\varepsilon(f(x)) \subset U$. Do f liên tục nên theo Định nghĩa 1.7 tìm được $\delta > 0$ sao cho $f(B_\delta(x)) \subset B_\varepsilon(f(x))$. Nhưng vì $B_\varepsilon(f(x)) \subset U$ nên $f(B_\delta(x)) \subset U$. Bằng cách áp dụng hàm f^{-1} vào cả hai vế của bao hàm thức này ta có $f^{-1}(f(B_\delta(x)) \subset f^{-1}(U)$ hay $B_\delta(x) \subset f^{-1}(U)$. Chứng tỏ $f^{-1}(U)$ là tập mở.

Đủ. Ta cần chứng minh nếu mỗi tập mở trong miền trị của f được f^{-1} biến thành tập mở trong miền xác định của f thì f là một ánh xạ liên tục. Lấy tập mở trong miền trị là hình cầu mở với bán kính $\varepsilon > 0$ bất kỳ, tâm tại điểm f(x) trong miền trị mà là ảnh của điểm x nào đó trong miền xác định của f. Như vậy, hình cầu $B_\varepsilon(f(x))$ là một tập mở trong miền trị của f và ta giả thiết rằng ảnh ngược của nó $f^{-f}(B_\varepsilon(f(x)))$ cũng là tập mở trong miền xác định của f. Ta chứng minh f liên tục. Thật vậy, theo giả thiết $f^{-1}(B_\varepsilon(f(x)))$ mở trong D nên tìm được một hình cầu mở quanh điểm $x \in f^{-1}(B_\varepsilon(f(x)))$ và nằm trọn trong $f^{-1}(B_\varepsilon(f(x)))$. Giả sử bán kính của hình cầu này là $\delta > 0$: $B_\delta(x) \subset f^{-1}(B_\varepsilon(f(x)))$. Từ đó $f(B_\delta(x)) \subset f(f^{-1}(B_\varepsilon(f(x))))$ hay $f(B_\delta(x)) \subset B_\varepsilon(f(x))$. Như vậy, f thoả mãn định nghĩa của hàm liên tục.

Ta cũng có định lý tương tự về hàm liên tục và ảnh ngược của các tập đóng.

Định lý 1.9. Tính liên tục và ảnh ngược của các tập đóng

Giả sử $f : D \to T$ *là một ánh xạ và* $f^{-1} : T \to D$ *là ánh xạ ngược của* f *từ* T *tới* D. *Giả sư* $U \subset T$ *là một tập đóng trong miền trị của* f. *Khi đó,* f *là liên tục khi và chỉ khi ảnh ngược* $f^{-1}(U) \subset D$ *là một tập đóng trong miền xác định của* f.

Chứng minh. Cho U là một tập đóng trong miền trị T. Ta tìm cách chứng minh ảnh ngược $f^{-1}(U)$ của U là một tập đóng trong miền xác định D $\Leftrightarrow$ f là liên tục. Thật vậy, U là tập đóng khi và chỉ khi phần bù của nó cU là tập mở trong T. Định lý 1.8 cho thấy f liên tục khi và chỉ khi ảnh ngược của tập mở $f^{-1}(cU)$ là tập mở trong D. Có thể chứng minh rằng ảnh ngược của phần bù của một tập bất kỳ trùng với phần bù của ảnh ngược của tập đó, vì thế $f^{-1}(cU) = c(f^{-1}(U))$. Như vậy, f liên tục khi và chỉ khi tập $c(f^{-1}(U))$ là mở trong D. Lấy phần bù một lần nữa ta thấy f liên tục khi và chỉ khi $f^{-1}(U) = c(c(f^{-1}(U)))$ là tập đóng trong D.

Hai định lý trên rất tổng quát và rất mạnh. Nếu biết được điều gì đó về ảnh ngược của các tập mở hay đóng trong miền trị thì có thể dùng các định lý này để phán đoán xem ánh xạ nói tới có liên tục hay không. Ngược lại, nếu biết được ánh xạ nói tới là liên tục thì có thể dùng các định lý này để chỉ ra những tính chất mà ảnh ngược của các tập mở hay đóng trong miền trị cần phải có.

Ta có thể chứng minh được rằng nếu $S \subset D$ là một tập compact và nếu f là một ánh xạ liên tục thì tập ảnh $f(S) \subset T$ cũng là một tập compact.

Định lý 1.10. ảnh liên tục của một tập compact là một tập compact

Giả sử $f : D \to \mathbb{R}$ *là một ánh xạ liên tục. Nếu tập con* $S \subset D$ *là một tập compact thì tập ảnh của nó* $f(S) \subset \mathbb{R}$ *cũng là một tập compact.*

Chứng minh. Xem chứng minh đầy đủ trong Nikaido (1972).

1.5. ĐỊNH LÍ TỒN TẠI(Existence Theorems)

Các định lý tồn tại chỉ rõ những điều kiện nếu được thoả mãn sẽ bảo đảm có các kết luận gì đó. Hai điểm cần lưu ý khi bàn về định lý tồn tại. Thứ nhất, điều kiện nêu ra trong các định lý này nói chung chỉ là điều kiện **đủ**, không nhất thiết là điều kiện cần. Nghĩa là khi các điều kiện của định lý được thoả mãn thì sự tồn tại của đối tượng đề cập tới được bảo đảm. Đồng thời trong những trường hợp không có các điều kiện này thì đối tượng đó vẫn có thể tồn tại. Thứ hai, các định lý này đảm bảo cho cái gì đó tồn tại, nhưng nói chung chúng không cho ta hình dung rõ nó như thế nào và tồn tại ở đâu.

Định lý thứ nhất là một kết quả cơ bản trong **lý thuyết tối ưu**. Nhiều bài toán kinh tế đòi hỏi tìm cực tiểu hay cực đại một hàm số xác định trên một tập nào đó của $\mathbb{R}^n$. Ta sẽ chủ yếu quan tâm tới bài toán tìm cực tiểu hay cực đại của các hàm biến đổi các véctơ trong $\mathbb{R}^n$ thành các số trong $\mathbb{R}$. Các hàm như thế được gọi là **hàm giá trị thực** và ta sẽ xét chi tiết lớp hàm này ở chương sau. Tuy nhiên, ở đây ta có thể dùng một số tính chất tôpô (đóng, mở, bị chặn …) để thiết lập một trong những định lý tồn tại thông dụng nhất với tên gọi **định lý Weierstrass**. Định lý đưa ra các điều kiện đủ cho sự tồn tại giá trị lớn nhất và giá trị nhỏ nhất của một hàm liên tục.

Định lý 1.11 (Weierstrass). Tồn tại giá trị cực trị (Extreme Values)

Giả sử $f : \mathbb{R}^n \to \mathbb{R}$ *là một hàm thực liên tục. Giả sử* S *là một tập compact trong* $\mathbb{R}^n$. *Khi đó, tìm được véctơ* $x^* \in S$ *và véctơ* $\overline{x} \in S$ *sao cho*

$$f(x^*) \leq f(x) \leq f(\bar{x}) \textit{ với mọi } x \in S.$$

Chứng minh. Do f liên tục và S compact nên theo Định lý 1.10, f(S) là một tập compact. Do f là hàm thực nên $f(S) \subset \mathbb{R}$. Do f(S) compact nên nó đóng và bị chặn. Theo Định lý 1.7, bất kỳ tập đóng và bị chặn trong tập số thực đều chứa cận dưới lớn nhất, gọi là a, và cận trên nhỏ nhất, gọi là b. Theo định nghĩa của tập ảnh, tìm được điểm $x^* \in S$ sao cho $f(x^*) = a \in f(S)$ và điểm $\bar{x} \in S$ sao cho $f(\bar{x}) = b \in f(S)$. Kết hợp với định nghĩa của cận dưới lớn nhất và cận trên nhỏ nhất ta có $f(x^*) \leq f(x)$ và $f(x) \leq f(\bar{x})$ với mọi $x \in S$.

• **Định lý tách** (Separation Theorems). Nói nôm na, các định lý tách cho những điều kiện đủ để một siêu phẳng (Hyperplane) có thể "đi xuyên qua" hai tập hợp lồi và chúng có nhiều ứng dụng trong toán học và trong lý thuyết kinh tế. Trước khi nêu định lý ta hãy làm quen với một số thuật ngữ.

Định nghĩa 1.8. Siêu phẳng H trong $\mathbb{R}^n$ là tập hợp các điểm $\mathbf{x} \in \mathbb{R}^n$ thoả mãn phương trình $\langle \mathbf{a}, \mathbf{x} \rangle = \alpha$, trong đó véctơ $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{a} \neq 0$, $\alpha \in \mathbb{R}$ là một số thực.

Trong $\mathbb{R}^2$ siêu phẳng là một đường thẳng có dạng $a_1x_1 + a_2x_2 = \alpha$ với a_1 hoặc $a_2 \neq 0$ hay $x_2 = \alpha/a_2 - a_1x_1/a_2$ (giả sử $a_2 \neq 0$). Dễ nhận ra đó là một đường thẳng có độ dốc $-a_1/a_2$ và cắt trục tung tại điểm α/a_2. Trong $\mathbb{R}^3$ siêu phẳng là một mặt phẳng. Trong không gian số chiều cao hơn siêu phẳng là một tập afin $(n - 1)$ chiều.

Định nghĩa 1.9. Ta nói siêu phẳng H **tách** hai tập S và T trong $\mathbb{R}^n$ nếu

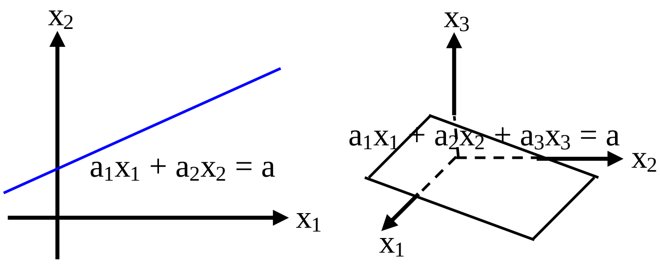
$$\langle \mathbf{a}, \mathbf{x} \rangle \leq \alpha \text{ với mọi } \mathbf{x} \in S \text{ và } \langle \mathbf{a}, \mathbf{y} \rangle \geq \alpha \text{ với mọi } \mathbf{y} \in T,$$

nghĩa là siêu phẳng H tách hai tập S và T nếu mọi điểm thuộc S nằm ở một phía của H, còn mọi điểm thuộc T nằm ở phía kia của H. Nếu H có ít nhất một điểm chung với biên của một trong hai tập thì ta nói H **tựa** (support) vào tập hợp đó và gọi H là **siêu phẳng tựa** (supporting hyperplane) của nó.

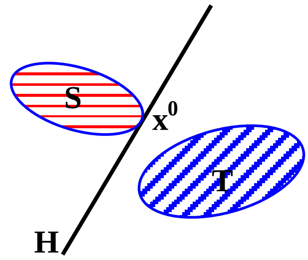
Định lý sau nêu một điều kiện đủ để có thể tách hai tập hợp lồi trong $\mathbb{R}^n$.

Định lý 1.12. Định lý tách Minkowski (Minkowski's Separation Theorem)

Cho S *và* T *là hai tập lồi, khác rỗng, rời nhau trong* $\mathbb{R}^n$. *Khi đó, tìm được véctơ* $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{a} \neq 0$ *và số* $\alpha \in \mathbb{R}$ *sao cho* $<\mathbf{a}, \mathbf{x}> \leq \alpha \ \forall \mathbf{x} \in S$ *và* $<\mathbf{a}, \mathbf{y}> \geq \alpha \ \forall \mathbf{y} \in T$.



Hình 1.9. Siêu phẳng trong $\mathbb{R}^2$ & $\mathbb{R}^3$



Hình 1.10. Siêu phẳng tách

Chương 2
HÀM GIÁ TRỊ THỰC

Chương này đề cập tới các hàm số thực thường gặp trong kinh tế và trong tính toán tối ưu. Khảo sát hàm số thông qua các tập có liên quan (đồ thị, tập mức), phân tích một số hàm thông dụng: hàm lồi, hàm lõm, hàm thuần nhất và cuối cùng xét tính vi phân của hàm số. Nội dung của chương dựa chủ yếu trên các tài liệu [1], [2], [3], [4].

2.1. HÀM SỐ THỰC VÀ CÁC HÀM CÓ LIÊN QUAN

Hàm giá trị thực rất hay gặp trong lý thuyết kinh tế vi mô. Hàm chi phí sản xuât, hàm lợi ích tiêu dùng, hàm cung cầu vật tư, hàng hoá … là những hàm quen thuộc nhất. Nói một cách hình thức

Định nghĩa 2.1. Hàm giá trị thực (Real Valued Functions)

$f : D \to T$ *là hàm giá trị thực nếu* D *là một tập bất kỳ và* $T \subset \mathbb{R}$.

(D là miền xác định, T là miền giá trị của hàm và $\mathbb{R}$ tập hợp các số thực).

Nói nôm na, f là hàm giá trị thực nếu nó biến đổi các phần tử trong miền xác định của nó vào đường thẳng thực. Nếu miền xác định là một tập con trong $\mathbb{R}^n$ thì hàm thực biến đổi véctơ trong $\mathbb{R}^n$ thành một số thực trong $\mathbb{R}$.

Các hàm

$$y = ax_1 + bx_2,\ y = \sqrt{z^2 + w^2} \text{ hay } y = \sum_{i=1}^{n} a_i x_i^2$$

là những ví dụ về hàm thực, vì trong mỗi trường hợp vế phải là một số thực.

Nói một cách đơn giản, hàm có giá trị thực khi "đầu ra" của nó là một số chứ không phải là véctơ, bất kể "đầu vào" của hàm là gì.

Lớp hàm thực rất rộng. Chương này sẽ giới thiệu một số loại hàm thực đặc biệt và khám phá những tính chất quan trọng của chúng. Trong suốt chương, ta hạn chế sự chú ý tới các hàm thực có miền xác định là các tập lồi.

Giả thiết 2.1. Hàm giá trị thực trên tập hợp lồi (over Convex Sets)

Cho $f : D \to T$ *là hàm thực, trong đó* $D \subset \mathbb{R}^n$ *là một tập lồi và* $T \subset \mathbb{R}$, nghĩa là nếu $x^1 \in D$, $x^2 \in D$ và $x^t = tx^1 + (1 - t)x^2$ với $t \in [0, 1]$ thì $x^t \in D$.

Các hàm thực trong nhiều ứng dụng kinh tế tiêu biểu thường có xu hướng tăng hoặc giảm một cách đều đặn trên miền xác định của chúng. Ta gọi đó là các hàm tăng hoặc hàm giảm. Ta nêu ra định nghĩa chặt chẽ cho các thuật ngữ này để dùng về sau.

Định nghĩa 2.2. Hàm tăng (Increasing Functions)

Hàm $f : D \to T$ được gọi là tăng hay **hàm tăng** nếu $f(x^0) \geq f(x^1) \Leftrightarrow x^0 \geq x^1$ ***và*** $x^0 \neq x^1$. Ta nói hàm f **tăng chặt** nếu $f(x^0) > f(x^1) \Leftrightarrow x^0 \geq x^1$ và $x^0 \neq x^1$.

Theo định nghĩa này, hàm được gọi là **tăng** nếu mỗi khi tăng một hay một số thành phần của véctơ $\mathbf{x} = (x_1, \dots , x_n)$ không làm giảm giá trị của hàm. Ta nói hàm là **tăng chặt** nếu mỗi khi tăng một hay nhiều thành phần của $\mathbf{x}$ làm tăng thực sự giá trị của hàm. Hàm giảm được định nghĩa tương tự.

Định nghĩa 2.3. Hàm giảm (Decreasing Functions)

Hàm $f : D \to T$ được gọi là giảm hay **hàm giảm** nếu $f(x^0) \leq f(x^1) \Leftrightarrow x^0 \geq x^1$ và $x^0 \neq x^1$. Ta nói hàm f **giảm chặt** nếu $f(x^0) < f(x^1) \Leftrightarrow x^0 \geq x^1$ và $x^0 \neq x^1$.

• CÁC TẬP CÓ LIÊN QUAN VỚI HÀM(Relateed Sets)

Như đã biết, đồ thị của hàm là một tập có liên quan mật thiết với hàm, đôi khi đồ thị cho một cách hiểu đơn giản và trực quan về hàm. Có một số tập khác có liên quan với hàm trở thành công cụ quen thuộc trong lý thuyết kinh tế, chúng có biểu diễn hình học đơn giản và thường cho dạng tương đương để khảo sát và thao tác với các hàm, đặc biệt đối với các hàm thực có miền trị là tập con của đường thẳng số thực $\mathbb{R}$.

Sau đây ta sẽ nêu định nghĩa về các tập có liên quan và nêu mối quan hệ giữa chúng với nhau nói chung và với hàm nói riêng, Tiếp đó, ta sẽ xét một số hàm thực đặc biệt và các tính chất đặc trưng của các tập có liên quan.

Khái niệm **tập mức** (hay **đường mức**) chắc chắn đã rất quen thuộc với nhiều người, mặc dàu nó có các tên gọi khác nhau. Nhiều đối tượng quen thuộc trong kinh tế vi mô như đường "đẳng mức", đường "đẳng lượng", đường "đẳng lợi nhuận" … đều là đường mức của các hàm thực. Đại thể, tập mức là tập hợp các phần tử thuộc miền xác định của hàm mà chúng được biến đổi thành cùng một giá trị số hay "mức" trong miền trị. Như vậy, hai phần tử bất kỳ trong cùng một tập mức sẽ cho cùng một giá trị số trong miền trị, khi đưa các phần tử đó vào hàm. Một cách hình thức:

Định nghĩa 2.4. Tập mức (Level Sets)

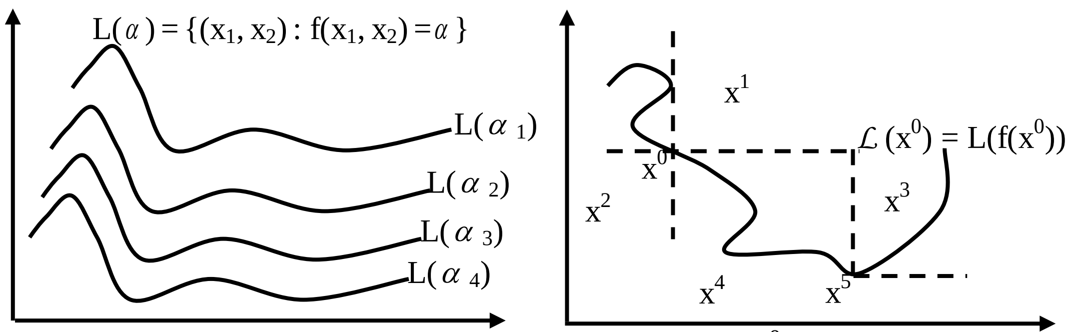
$L(\alpha)$ được gọi là **tập mức** của hàm thực $f : D \to T \Leftrightarrow L(\alpha) = \{x \mid x \in D, f(x) = \alpha\}$ với $\alpha \in T \subset \mathbb{R}$.

Có thể thấy hai tập mức khác nhau của hàm không bao giờ cắt nhau, vì nếu trái lại sẽ có một phần tử trong miền xác định được đặt tương ứng với hai giá trị (hai mức) khác nhau, trái với định nghĩa của hàm.

Ta cũng xác định tập mức theo mức $\alpha = f(x^0)$ với x^0 thuộc miền xác định:

Định nghĩa 2.5. Tập mức đối với điểm $x^0 \in D$

$\mathcal{L}(x^0)$ được gọi là **tập mức** đối với điểm $x^0 \in D \Leftrightarrow \mathcal{L}(x^0) = \{x \mid x \in D, f(x) = f(x^0)\}$ (Nhận xét $\mathcal{L}(x^0) = L(f(x^0))$ và hai ký hiệu khác nhau: $\mathcal{L} \neq L$).



Hình 2.1. Tập mức $L(\alpha)$ và $\mathcal{L}(x^0)$

Định nghĩa 2.6. Tập mức dưới / mức trên (Inferior & Superior Sets)

1. $I(\alpha) \equiv \{x \mid x \in D, f(x) \le \alpha\}$ được gọi là **tập mức dưới** của mức α.

2. $S(\alpha) \equiv \{x \mid x \in D, f(x) \ge \alpha\}$ được gọi là **tập mức trên** của mức α.

3. $I'(\alpha) \equiv \{x \mid x \in D, f(x) < \alpha\}$ gọi là **tập mức dưới chặt** của mức α.

4. $S'(\alpha) \equiv \{x \mid x \in D, f(x) > \alpha\}$ gọi là **tập mức trên chặt** của mức α.

Tập mức **dưới** bao gồm tất cả các điểm của D có giá trị hàm bằng hoặc nhỏ hơn giá trị α, còn tập mức **dưới chặt** chỉ gồm các điểm của D có giá trị hàm nhỏ hơn hẳn giá trị α. Tập mức **trên** bao gồm tất cả các điểm thuộc D có giá trị hàm bằng hoặc lớn hơn giá trị α, còn tập mức **trên chặt** chỉ gồm các điểm thuộc D có giá trị hàm lớn hơn hẳn giá trị α.

Định lý sau cho thấy rõ mối quan hệ giữa các tập mức này.

Định lý 2.1. Tập mức, tập mức trên & tập mức dưới (sup./ inf. level sets)

Với mọi $f : D \to I$ và $\alpha \in I$ ta có các hệ thức

1. $L(\alpha) \subset I(\alpha)$.

2. $L(\alpha) \subset S(\alpha)$.

3. $L(\alpha) = I(\alpha) \cap S(\alpha)$

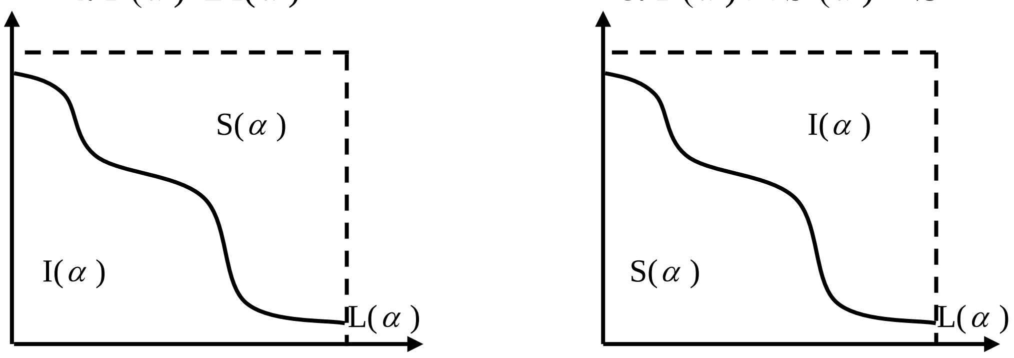
4. $I'(\alpha) \subset I(\alpha)$

5. $S'(\alpha) \subset S(\alpha)$.

6. $I'(\alpha) \cap L(\alpha) = \varnothing$

7. $S'(\alpha) \cap L(\alpha) = \varnothing$

8. $I'(\alpha) \cap S'(\alpha) = \varnothing$



a) Hàm tăng b) Hàm giảm

Hình 2.2. Tập mức, tập mức dưới/ tập mức trên của hàm tăng/ hàm giảm

Nhận xét khi f(x) là hàm **tăng**, $S(\alpha)$ nằm phía trên tập mức $L(\alpha)$, còn $I(\alpha)$ nằm phía dưới tập mức $L(\alpha)$. Ngược lại, khi hàm **giảm**, $S(\alpha)$ nằm phía dưới tập mức $L(\alpha)$, còn $I(\alpha)$ nằm phía trên tập mức $L(\alpha)$ (xem Hình 2.2).

2.2. CÁC HÀM THÔNG DỤNG

2.2.1. HÀM LỒI VÀ HÀM TỰA LỒI(Convex and Quasi-convex Functions)

Định nghĩa 2.7. Hàm lồi (Convex Functions)

$f : D \to \mathbb{R}$ được gọi là hàm **lồi** $\Leftrightarrow$ với mọi $\mathbf{x}^1$, $\mathbf{x}^2$ thuộc D ta có

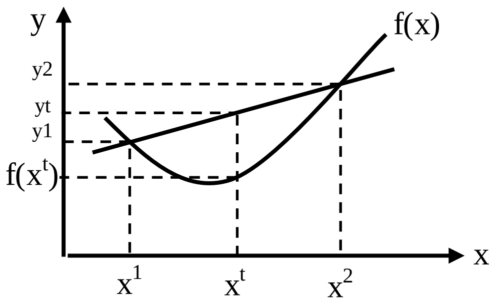
$$f(tx^1 + (1-t)x^2) \le tf(\mathbf{x}^1) + (1-t)f(\mathbf{x}^2) \text{ với mọi } t \in [0, 1].$$

Định nghĩa 2.8. Hàm lồi chặt (Strictly Convex Functions)

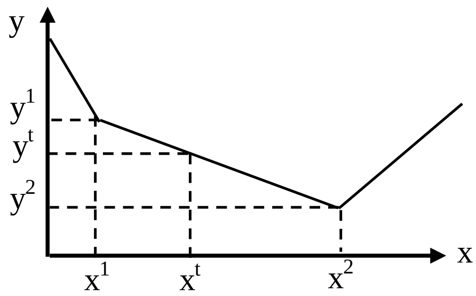
$f : D \to \mathbb{R}$ được gọi là hàm **lồi chặt** $\Leftrightarrow$ với mọi $\mathbf{x}^1 \ne \mathbf{x}^2$ thuộc D ta có

$$f(tx^1 + (1-t)x^2) < tf(\mathbf{x}^1) + (1-t)f(\mathbf{x}^2) \text{ với mọi } t \in (0, 1).$$

Định nghĩa của hàm lồi đòi hỏi giá trị của hàm tại một tổ hợp lồi nào đó của hai điểm bất kỳ $\mathbf{x}^1$, $\mathbf{x}^2$ không lớn hơn giá trị nhận được khi lấy cùng tổ hợp lồi như thế của hai giá trị $f(\mathbf{x}^1)$, $f(\mathbf{x}^2)$. Về hình học, f lồi nếu điểm $(x^t, tf(\mathbf{x}^1) + (1-t)f(\mathbf{x}^2))$ ở trên dây cung nối hai điểm $(\mathbf{x}^1, f(\mathbf{x}^1))$, $(\mathbf{x}^2, f(\mathbf{x}^2))$ không thấp hơn điểm $(\mathbf{x}^t, f(\mathbf{x}^t))$ trên đồ thị của f. Đồ thị của hàm lồi không khi nào nằm cao hơn dây cung nối hai điểm bất kỳ của nó và tập các điểm nằm về phía trên đồ thị của một hàm lồi luôn là một tập lồi (Hình 2.3).



Hình 2.3. Hàm lồi (chặt)



Hình 2.4. Hàm lồi (không chặt)

Định lý 2.2. Toàn bộ các điểm thuộc đồ thị và các điểm nằm ở phía trên đồ thị của một hàm lồi luôn tạo nên một tập hợp lồi

Cho $D \subset \mathbb{R}^n$ *là một tập hợp lồi.* Ký hiệu $A \equiv \{(\mathbf{x}, \alpha) \mid \mathbf{x} \in D, f(\mathbf{x}) \le \alpha\}$ *là tập hợp các điểm "thuộc và ở phía trên" đồ thị của* $f : D \to \mathbb{R}$. *Khi đó*

$$f \text{ là hàm lồi} \Leftrightarrow A \text{ là tập hợp lồi.}$$

Ta xét lớp hàm rộng hơn các hàm lồi và hàm lồi chặt.

Định nghĩa 2.9. Hàm tựa lồi (Quasi-convex Functions) f : D → ℝ được gọi là **hàm tựa lồi** ⇔ với mọi $\mathbf{x}^1$, $\mathbf{x}^2$ thuộc D ta có

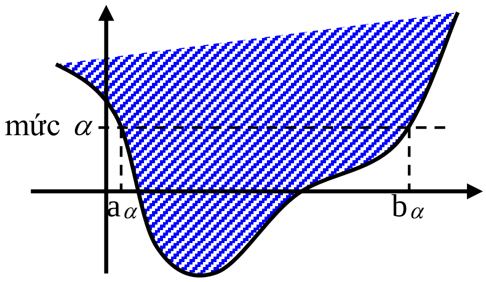
$$f(t\mathbf{x}^1 + (1-t)\mathbf{x}^2) \le \max[f(x^1), f(\mathbf{x}^2)]\ \forall t \in [0, 1].$$

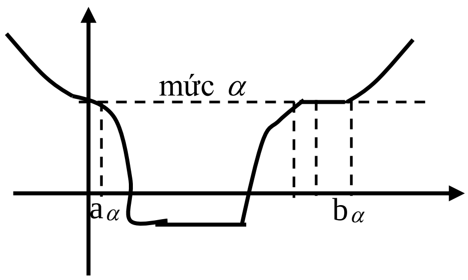
Định nghĩa 2.10. Hàm tựa lồi chặt (Strictly Quasi-convex Functions)

f : D → ℝ được gọi là **hàm tựa lồi chặt** ⇔ với mọi $\mathbf{x}^1 \neq \mathbf{x}^2$ thuộc D ta có

$$f(t\mathbf{x}^1 + (1-t)\mathbf{x}^2) < \max[f(\mathbf{x}^1), f(\mathbf{x}^2)]\ \forall t \in (0, 1).$$

Trong các định nghĩa vừa nêu, phép toán max[a, b] là số lớn nhất của a và b. Nếu a > b thì max[a, b] = a. Nếu a = b thì max[a, b] = a hay b.





Hình 2.5. Hàm tựa lồi (chặt)   Hình 2.6. Hàm tựa lồi (không chặt)

Định lý 2.3. Tựa lồi và tập mức dưới (Quasi-convexity & the Inferior Sets)

f : D → ℝ *là hàm tựa lồi* ⇔ I(α) *là tập lồi với mọi* $\alpha \in$ ℝ.

Tập mức dưới của hàm tựa lồi chặt không chứa đoạn thẳng ở biên của nó.

Định lý 2.4. Tính lồi kéo theo tính tựa lồi

Hàm lồi luôn là hàm tựa lồi. Hàm lồi chặt luôn là hàm tựa lồi chặt.

Chứng minh. Ta nêu ra chứng minh kiến thiết cho trường hợp hàm lồi, trường hợp lồi chặt chứng minh tương tự.

Giả sử f : D → ℝ hàm lồi. Lấy bất kỳ $\mathbf{x}^1$, $\mathbf{x}^2 \in$ D. Không giảm tổng quát ta xem như $f(\mathbf{x}^1) \le f(\mathbf{x}^2)$. Từ định nghĩa hàm lồi, với $\mathbf{x}^t \equiv t\mathbf{x}^1 + (1-t)\mathbf{x}^2$ ta có

$$f(\mathbf{x}^t) \le tf(\mathbf{x}^1) + (1-t)f(\mathbf{x}^2) \text{ với mọi } t \in [0, 1] \text{ hay}$$

$$f(\mathbf{x}^t) \le f(\mathbf{x}^2) + t(f(\mathbf{x}^1) - tf(\mathbf{x}^2)) \text{ với mọi } t \in [0, 1].$$

Do $t \ge 0$ và $f(\mathbf{x}^1) \le f(\mathbf{x}^2)$ nên $t(f(\mathbf{x}^1) - tf(\mathbf{x}^2)) \le 0$. Từ đó $f(\mathbf{x}^t) \le f(\mathbf{x}^2)$. Theo trên $f(\mathbf{x}^2) = \max\{f(\mathbf{x}^t), f(\mathbf{x}^2)\}$. Vì thế, $f(\mathbf{x}^t) \le \max\{f(\mathbf{x}^t), f(\mathbf{x}^2)\}\ \forall t \in [0, 1]$, nghĩa là f thoả mãn định nghĩa của hàm tựa lồi.

2.2.2. HÀM LÕM VÀ HÀM TỰA LÕM (Concave & Quasi-Concave Functions)

Định nghĩa 2.11. Hàm lõm (Concave Functions)

f : D → T được gọi là **hàm lõm** ⇔ với mọi $\mathbf{x}^1$ và $\mathbf{x}^2$ thuộc D ta có

$$f(t\mathbf{x}^1 + \mathbf{(1 - 1)x}^2) \geq t.f(\mathbf{x}^1) + (1 - t).f(\mathbf{x}^2)]\ \forall t \in [0, 1].$$

Hàm lõm phản ánh qui luật "tiết kiệm do qui mô mang lại": khối lượng sản xuất càng lớn chi phí sản xuất trên một đơn vị sản phẩm càng hạ.

Về trực giác, ta thấy: Đồ thị của một hàm lõm không khi nào nằm thấp hơn dây cung nối hai điểm bất kỳ của đồ thị và tập các điểm nằm về phía dưới đồ thị của một hàm lõm luôn là một tập lồi.

Định lý 2.5. Tập các điểm thuộc đồ thị và các điểm nằm ở phía dưới đồ thị của một hàm lõm luôn tạo nên một tập hợp lồi

Cho D ⊂ $\mathbb{R}^n$ *là một tập hợp lồi*. Ký hiệu B ≡ {(x, α) | x ∈ D, α ≤ f(x)} *là tập hợp các điểm* "*thuộc và ở phía dưới*" *đồ thị của* f : D → ℝ. *Khi đó*

f *là hàm lõm* ⇔ B *là tập hợp lồi*.

Chứng minh. Cần chỉ rõ: f lõm ⇒ B lồi và B lồi ⇒ f lõm.

Định nghĩa 2.12. Hàm lõm chặt (Strictly Concave Functions)

f : D → ℝ được gọi là **hàm lõm chặt** ⇔ với mọi $\mathbf{x}^1 \neq \mathbf{x}^2$ thuộc D ta có

$$f(t\mathbf{x}^1 + \mathbf{(1 - t)x}^2) > t.f(\mathbf{x}^1) + (1 - t).f(\mathbf{x}^2)]\ \forall t \in (0, 1).$$

Định nghĩa 2.13. Hàm tựa lõm (Quasi-concave functions)

f : D → ℝ được gọi là **hàm tựa lõm** ⇔ với mọi $\mathbf{x}^1$ và $\mathbf{x}^2$ thuộc D ta có

$$f(t\mathbf{x}^1 + \mathbf{(1 - t)x}^2) \geq \min\,[f(\mathbf{x}^1), f(\mathbf{x}^2)]\ \forall t \in [0, 1].$$

Trong định nghĩa trên, phép toán min[a, b] là số nhỏ nhất của a và b. Nếu a > b thì min[a, b] = b. Nếu a = b thì min[a, b] = a hoặc b.

Định nghĩa 2.14. Hàm tựa lõm chặt (Strictly Quasi-concave Functions)

f : D → ℝ được gọi là **tựa lõm chặt** ⇔ với mọi $\mathbf{x}^1 \neq \mathbf{x}^2$ thuộc D ta có

$f(t\mathbf{x}^1 + \mathbf{(1 - t)x}^2) > \min\,[f(\mathbf{x}^1), f(\mathbf{x}^2)]\ \forall t \in (0, 1).$

Định lý 2.6. Tựa lõm và tập mức trên (Quasi-concavity & the superior sets)

f : D → ℝ *là hàm tựa lõm* ⇔ S(x) *là tập lồi với mọi* x ∈ D.

Tập mức trên của hàm tựa lõm chặt không chứa đoạn thẳng ở biên của nó

Định lý 2.7. Tính lõm kéo theo tính tựa lõm

Hàm lõm luôn là hàm tựa lõm. Hàm lõm chặt luôn là hàm tựa lõm chặt.

Chứng minh. Tương tự chứng minh Định lý 2.4.

Định lý sau cho thấy mối liên hệ chặt chẽ giữa hàm lồi (lồi chặt) với hàm lõm (lõm chặt) cũng như giữa hàm tựa lồi (tựa lồi chặt) với hàm tựa lõm (tựa lõm chặt).

Định lý 2.8. Hàm lồi, hàm lõm và hàm tựa lồi, hàm tựa lõm

1. f(**x**) *là hàm lồi* (*lồi chặt*) ⇔ - f(**x**) *là hàm lõm* (*lõm chặt*).

2. f(**x**) *là hàm tựa lồi* (tựa *lồi chặt*) ⇔ - f(**x**) *là hàm tựa lõm* (tựa *lõm chặt*).

Chứng minh hiển nhiên, do tính (tựa) lồi là 'đảo dấu' của tính (tựa) lõm.

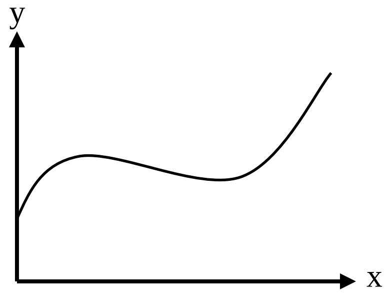
Mối quan hệ đã xét giữa các hàm lồi và lõm được tóm tắt như sau

1.	f lồi	⇔	phía trên đồ thị là tập lồi
2.	f lõm	⇔	phía dưới đồ thị là tập lồi
3.	f tựa lồi	⇔	các tập mức dưới là lồi
4.	f tựa lõm	⇔	các tập mức trên là lồi
5.	f lồi (lồi chặt)	⇔	- f lõm (lõm chặt)
6.	f tựa lồi (tựa lồi chặt)	⇔	- f tựa lõm (tựa lõm chặt)
7.	f lồi	⇒	f tựa lồi
8.	f lõm	⇒	f tựa lõm

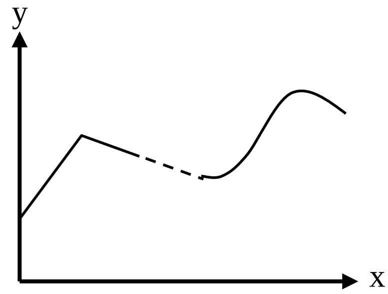
2.3. VI PHÂN CỦA HÀM SỐ

2.3.1. HÀM MỘT BIẾN (Functions of a Single Variable)

• Nói nôm na, hàm y = f(x) khả vi nếu nó liên tục và trơn (không có điểm gẫy hay xoắn). Tính khả vi là đòi hỏi cao hơn tính liên tục. Đó cũng là yêu cầu để có thể dùng các công cụ giải tích quen thuộc.



Hình 2.7. Hàm khả vi



Hình 2.8. Hàm không khả vi

Khi nói tới đạo hàm của hàm tại giá trị x, ta hiểu đó là độ dốc hay tốc độ thay đổi tức thời của giá trị f(x), Vì thế đôi khi ta viết

$$\frac{dy}{dx} = f'(x) \qquad (2.1)$$

để chỉ ra rằng f'(x) cho biết y thay đổi (tức thời) một lượng dy khi x thay đổi một lượng dx. Nếu đạo hàm cấp một là hàm khả vi thì ta lại có thể lấy đạo hàm của nó và nhận được đạo hàm cấp hai của hàm ban đầu

$$\frac{d^2y}{dx^2} = f''(x) \qquad (2.2)$$

Nếu hàm có các đạo hàm liên tục f', f'', … , $f^{(n)}$ thì hàm được gọi là **khả vi liên tục n lần** hay hàm thuộc lớp C^n.

• Vi phân là một khái niệm liên quan chặt chẽ với đạo hàm, nhưng khác biệt với đạo hàm. Vi phân của hàm f được ký hiệu là dy hay df(x) và được xem như số đo độ gia tăng tức thời của giá trị hàm tại điểm x theo một thay đổi "nhỏ" dx của x. Nếu y = f(x) thì độ gia tăng dy theo thay đổi dx sẽ là

$$dy = f'(x)dx. \qquad (2.3)$$

Vi phân cũng là một hàm và ta có thể lấy vi phân của nó. Ta gọi đó là vi phân cấp hai và có thể xem như để đo tại mỗi điểm x "mức độ thay đổi của sự gia tăng" giá trị của hàm theo sự gia tăng của x. Vi phân cấp hai, ký hiệu là d^2y hay $d^2f(x)$, nhận được bằng cách lấy vi phân của vi phân cấp một

$$d^2y = d(dy) = d(f'(x)dx) = (f''(x)dx)dx = f''(x)dx^2. \qquad (2.4)$$

Vi phân cấp một và cấp hai bao gồm đạo hàm cấp một và cấp hai của hàm. Các đạo hàm này cho thông tin quan trọng về hành vi tổng quát của hàm được xét. Đạo hàm cấp một cho biết giá trị hàm tăng hay giảm khi tăng x, còn đạo hàm cấp hai cho biết "độ cong" của hàm. Vi phân cấp một và hai cũng cho cùng thông tin tương tự.

Định lý sau cho thông tin về độ dốc, độ cong rút ra từ vi phân cấp 1 và 2.

Định lý 2.9. Độ dốc, độ cong và vi phân (Slope, Curvature & Differentials)

Với hàm 2 lần khả vi liên tục f(x) *trong lân cận điểm* x *và* $\forall dx \neq 0$, *ta có*

Vi phân cấp một:

$dy \geq 0 \Leftrightarrow f'(x) \geq 0 \Leftrightarrow$ f *tăng địa phương*

$dy \leq 0 \Leftrightarrow f'(x) \leq 0 \Leftrightarrow$ f *giảm địa phương*

$dy = 0 \Leftrightarrow f'(x) = 0 \Leftrightarrow$ f *hằng địa phương*

Vi phân cấp hai:

$d^2y \geq 0 \Leftrightarrow f''(x) \geq 0 \Leftrightarrow$ f *lồi địa phương*

$d^2y \leq 0 \Leftrightarrow f''(x) \leq 0 \Leftrightarrow$ f *lõm địa phương*

$d^2y = 0 \Leftrightarrow f''(x) = 0 \Leftrightarrow$ f *tuyến tính địa phương*

2.3.2. HÀM NHIỀU BIẾN (Functions of Several Variables)

Ta sẽ thường xuyên làm việc với hàm thực nhiều biến số. Có thể dễ dàng mở rộng các ý tưởng vừa nêu cho những hàm này.

Định nghĩa 2.15. Đạo hàm riêng (Partial Derivatives)

Cho $y = f(x_1, \dots, x_n)$. Khi đó đạo hàm riêng của f đối với x_j xác định bởi

$$\frac{\partial f(x)}{\partial x_j} \equiv \lim_{h\to 0} \frac{f(x_1,\dots,x_j+h,\dots,x_n) - f(x_1,\dots,x_j,\dots x_n)}{h}.$$

Đôi khi ta còn dùng một số ký hiệu khác để chỉ đạo hàm riêng, trong đó thông dụng nhất là $\partial y/\partial x_j$ hay $f'_j(\mathbf{x})$.

Lưu ý một số điểm quan trọng về đạo hàm riêng. Trước hết, có tất cả n đạo hàm riêng, mỗi đạo hàm theo từng biến x_j. Thứ hai, mỗi đạo hàm riêng cũng là một hàm. Cuối cùng, các đạo hàm riêng được xác định tại mỗi điểm thuộc miền xác định và cho biết sự thay đổi giá trị hàm theo sự thay đổi của biến x_j khi giữ nguyên giá trị các biến khác. Xét ví dụ sau đây về hàm 2 biến.

Ví dụ 2.1. Cho $f(x_1, x_2) = x_1^2 + 3x_1x_2 - x_2^2$. Đây là hàm của hai biến, vì thế có hai đạo hàm riêng. Lấy đạo hàm theo biến x_1 ta được

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = 2x_1 + 3x_2.$$

Lấy đạo hàm theo biến x_2 ta được

$$\frac{\partial f(x_1, x_2)}{\partial x_2} = 3x_1 - 2x_2.$$

Nhận xét là mỗi đạo hàm riêng ở đây lại là hàm của x_1, x_2. Các đạo hàm riêng này có giá trị khác nhau tại các điểm (x_1, x_2) khác nhau: Tại điểm (1, 2), $f_1^{'}(1, 2) = 8$, $f_2^{'}(1, 2) = -1$. Tại điểm (2, 1), $f_1^{'}(2, 1), = 7$, $f_2^{'}(2, 1) = 4$. □

Với hàm nhiều biến $y = f(\mathbf{x})$, để xét xem giá trị y thay đổi thế nào khi các biến x_j đồng thời thay đổi, mỗi biến một lượng "nhỏ" dx_j, ta dùng **vi phân toàn phần** cấp một của hàm.

$$dy = \frac{\partial f(x)}{\partial x_1} dx_1 + \ldots + \frac{\partial f(x)}{\partial x_n} dx_n = f_1' dx_1 + \ldots + f_n' dx_n = \sum_{j=1}^{n} f_j^{'}(x) dx_j$$

Dùng ký hiệu véctơ $\nabla f(\mathbf{x}) \equiv (f_1', \ldots, f_n')^T$ và $d\mathbf{x} = (dx_1, \ldots, dx_n)^T$. Ta thấy

$$dy = \nabla f(\mathbf{x}).d\mathbf{x}. \tag{2.5}$$

Lập ma trận các đạo hàm riêng cấp hai, gọi là ma trận Hess của f tại x:

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

Sau đây là một định lý quan trọng về các đạo hàm riêng cấp hai.

Định lý 2.10. Định lý Young (Young's Theorem)

Với hàm hai lần khả vi liên tục f(x) *ta có*

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial^2 f(x)}{\partial x_j \partial x_i} \ \forall i \text{ và } j.$$

Định lý Young cho thấy ma trận Hess là đối xứng. Tuy không nêu chứng minh định lý, nhưng ta có thể dễ dàng kiểm tra nó bằng việc xét một ví dụ.

Ví dụ 2.2. Xét hàm hai biến $f(x_1, x_2) = x_1x_2^2 + x_1x_2$. Các đạo hàm riêng cấp một của hàm này là

$$\frac{\partial f}{\partial x_1} \equiv f_1'(\mathbf{x}) = x_2^2 + x_2 \text{ và } \frac{\partial f}{\partial x_2} \equiv f_2'(\mathbf{x}) = 2x_1x_2 + x_1.$$

Lấy đạo hàm của f_1' theo x_2 ta được

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} \equiv f_{12}''(\mathbf{x}) = 2x_2 + 1.$$

Lấy đạo hàm của f_2' theo x_1 ta được

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} \equiv f_{21}''(\mathbf{x}) = 2x_2 + 1.$$

Rõ ràng $f_{12}'' = f_{21}''$ với mọi $\mathbf{x}$, như đã khẳng định trong Định lý Young.

Lấy vi phân (2.5) ta có

$$d^2y = \nabla(\nabla f(\mathbf{x}).d\mathbf{x}).d\mathbf{x} = \mathbf{dx}^{\mathbf{T}}.\mathbf{H}(\mathbf{x}).\mathbf{dx}. \qquad (2.6)$$

Biểu thức (2.6) là một dạng toàn phương của $dx_1, \dots, dx_n$ và là mở rộng của f"(x)dx^2 trong trường hợp một biến. Dấu của dạng thức này cho ta biết về độ cong của hàm. Định lý sau là một mở rộng của phần hai trong Định lý 2.9.

Định lý 2.11. Độ cong theo nhiều biến (Curvature in Several Variables)

Giả sử $f : D \to \mathbb{R}$ *hai lần khả vi liên tục và* $\mathbf{x} \in D$. *Khi đó*

$$d^2y \geq 0 \Leftrightarrow f \textit{ lồi tại } \mathbf{x} \Leftrightarrow d\mathbf{x}^{\mathbf{T}}.\mathbf{H}(\mathbf{x}).d\mathbf{x} \geq 0 \ \forall d\mathbf{x}$$

$$d^2y \leq 0 \Leftrightarrow f \textit{ lõm tại } \mathbf{x} \Leftrightarrow d\mathbf{x}^{T}.\mathbf{H}(\mathbf{x}).d\mathbf{x} \leq 0 \ \forall d\mathbf{x}$$

$d^2y > 0 \Rightarrow$ f *lồi chặt tại* $\mathbf{x} \Leftarrow d\mathbf{x}^T.\mathbf{H}(\mathbf{x}).d\mathbf{x} > 0\ \forall d\mathbf{x} \neq \mathbf{0}$.

$d^2y < 0 \Rightarrow$ f *lõm chặt tại* $\mathbf{x} \Leftarrow d\mathbf{x}^T.\mathbf{H}(\mathbf{x}).d\mathbf{x} < 0\ \forall d\mathbf{x} \neq \mathbf{0}$

Các quan hệ này là toàn cục nếu chúng đúng với mọi $\mathbf{x} \in D$.

Ta không nêu chứng minh định lý này ở đây, mặc dàu các kết luận nêu trong định lý đã được biết rõ. Hai phần cuối của định lý chỉ có kết luận một chiều. Bằng ví dụ cho thấy không thể thay dấu $\Rightarrow$ hay $\Leftarrow$ bởi dấu $\Leftrightarrow$.

Chú ý là trong trường hợp một biến điều kiện cần và đủ để hàm là lồi (lõm) trong một miền nào đó là đạo hàm cấp một của nó không giảm (không tăng). Trong trường hợp nhiều biến, ta chỉ có điều kiện cần, nhưng không đủ, cho tính lồi hay tính lõm tuỳ thuộc dấu của tất cả các đạo hàm riêng cấp hai.

Định lý 2.12. Tính lồi, tính lõm và đạo hàm riêng cấp hai

(Convexity, Concavity and Second-Order Partial Derivatives)

Giả sử $y = f(\mathbf{x})$ *là hàm hai lần khả vi liên tục*

1. *Nếu* $f(\mathbf{x})$ *lồi thì* $f''_{jj}(\mathbf{x}) \geq 0$, $j = 1, \ldots, n$.

2. *Nếu* $f(\mathbf{x})$ *lõm thì* $f''_{jj}(\mathbf{x}) \leq 0$, $j = 1, \ldots, n$.

3. *Nếu* $f(\mathbf{x})$ *lồi chặt hay lõm chặt thì các bất đẳng thức trên được thay tương ứng bằng các bất đẳng thức thực sự* > *hay* <.

Chứng minh. Ta nêu chứng minh cho trường hợp hàm lồi bằng phản chứng (với hàm lõm chứng minh tương tự).

Giả sử f là hàm lồi và $f''_{jj} < 0$ với j nào đó. Do f lồi nên theo Định lý 2.11 ta có $d^2y \geq 0$ với mọi $d\mathbf{x}$. Nói riêng, $d^2y \geq 0$ với $d\mathbf{x} = (0, \ldots, dx_j, \ldots, 0)$, $dx_j \neq 0$. Nhưng khi đó $d^2y = d\mathbf{x}^T\mathbf{H}d\mathbf{x} = (f''_{jj})(d\mathbf{x_j})^2$. Do $d\mathbf{x}_j \neq 0$ và $f''_{jj} < 0$ nên $d^2y < 0$. Nhưng theo Định lý 2.11, hàm f lõm, ta gặp mâu thuẫn.

2.3.3. HÀM THUẦN NHẤT(Homogeneous Functions)

Hàm thực thuần nhất cũng rất hay gặp trong các ứng dụng kinh tế vi mô. Trong mục này ta xét vắn tắt các hàm loại này và sử dụng các công cụ giải tích để thiết lập một số tính chất quan trọng của chúng.

Định nghĩa 2.16. Hàm thuần nhất (Homogeneous Functions)

Hàm thực f(**x**) gọi là

1. thuần nhất bậc k nếu và chỉ nếu

$$f(t\mathbf{x}) \equiv t^k f(\mathbf{x}) \ \forall t > 0$$

2. thuần nhất bậc 1 (hay thuần nhất tuyến tính) nếu và chỉ nếu

$$f(t\mathbf{x}) \equiv t f(\mathbf{x}) \ \forall t > 0$$

3. thuần nhất bậc 0 nếu và chỉ nếu

$$f(t\mathbf{x}) \equiv f(\mathbf{x}) \ \forall t > 0$$

Tính thuần nhất là một đặc trưng toàn cục (đúng với mọi x). Hàm thuần nhất biểu thị hành vi rất đều đặn, khi mọi biến tăng theo cùng một tỉ lệ. Chẳng hạn, khi hàm là thuần nhất bậc 1 thì khi tăng gấp đôi (gấp ba) mọi biến, giá trị của hàm cũng tăng lên gấp đôi (gấp ba). Với hàm thuần nhất bậc 0, khi các biến thay đổi theo cùng một tỉ lệ thì giá trị của hàm không hề thay đổi.

Ví dụ 2.3. Dạng Cobb-Douglas:

$$f(x_1, x_2) \equiv Ax_1^{\alpha} x_2^{\beta}, \ A > 0, \ \alpha > 0, \ \beta > 0.$$

Đây là hàm thuần nhất bậc $\alpha + \beta > 0$. Thật vậy,

$$f(tx_1, tx_2) \equiv A(tx_1)^{\alpha}(tx_2)^{\beta} \equiv t^{\alpha}.t^{\beta}Ax_1^{\alpha} x_2^{\beta} \equiv t^{\alpha+\beta}f(tx_1, tx_2).$$

Nếu các hệ số thoả mãn $\alpha + \beta = 1$ thì đó là hàm thuần nhất bậc 1.

Các đạo hàm riêng của một hàm thuần nhất cũng là một hàm thuần nhất.

Định lý 2.13. Đạo hàm riêng của hàm thuần nhất

(Partial Derevatives of Homogeneous Functions)

Nếu f(x) *là hàm thuần nhất bậc* k *thì các đạo hàm riêng của nó là hàm thuần nhất bậc* k - 1.

Chứng minh. Giả sử f(**x**) là hàm thuần nhất bậc k. Khi đó

$$f(t\mathbf{x}) \equiv t^k f(\mathbf{x}) \ \forall t > 0. \qquad \text{(P.1)}$$

Lấy đạo hàm vế trái theo x_j ta có

$$\frac{\partial}{\partial x_j}(f(t\mathbf{x})) \equiv \frac{\partial f(tx)}{\partial x_j} = \frac{\partial f(tx)}{\partial x_j}\frac{\partial (tx)}{\partial x_j} = \frac{\partial f(tx)}{\partial x_j}t, \qquad (P.2)$$

Lấy đạo hàm vế phải theo x_j ta được

$$\frac{\partial}{\partial x_j}(t^k f(\mathbf{x})) = t^k \frac{\partial f(x)}{\partial x_j}. \qquad (P.3)$$

Do (P.1) là một đồng nhất thức nên (P.2) phải bằng (P.3), nghĩa là

$$\frac{\partial f(tx)}{\partial x_j}t = t^k \frac{\partial f(x)}{\partial x_j}.$$

Chia cả hai vế cho t ta nhận được

$$\frac{\partial f(tx)}{\partial x_j} = t^{k-1}\frac{\partial f(x)}{\partial x_j} \text{ với } j = 1, \dots, n \text{ và } t > 0.$$

Nhiều ứng dụng thường gặp khi hàm là thuần nhất bậc 1. ở đây ta ghi lại kết quả này như một hệ quả trực tiếp.

Hệ quả 2.1. Hàm thuần nhất tuyến tính (Linear Homogeneous Functions)

Nếu f(x) *là hàm thuần nhất bậc* 1 *thì*

$$\frac{\partial f(tx)}{\partial x_j} = \frac{\partial f(x)}{\partial x_j} \text{ với mọi } j = 1, \dots, n \text{ và } \forall t > 0.$$

Hệ quả này nói rằng nếu hàm là tuyến tính thuần nhất thì khi tăng mọi biến theo cùng một tỉ lệ, tất cả n đạo hàm riêng của hàm sẽ không thay đổi. Ta hãy kiểm tra lại tính chất này đối với hàm Cobb-Douglas.

Ví dụ 2.4. Giả sử $f(x_1, x_2) \equiv Ax_1^{\alpha}x_2^{\beta}$ và $\alpha + \beta = 1$, vì thế hàm là tuyến tính thuần nhất

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = \alpha A x_1^{\alpha-1}x_2^{\beta}.$$

Nhân x_1, x_2 với t và lấy đạo hàm riêng tại (tx_1, tx_2) ta nhận được

$$\frac{\partial f(tx_1, tx_2)}{\partial x_1} = \alpha A(tx_1)^{\alpha-1}(tx_2)^{\beta} = t^{\alpha+\beta-1}\alpha A x_1^{\alpha-1}x_2^{\beta} = \frac{\partial f(x_1, x_2)}{\partial x_1}.$$

do $\alpha + \beta = 1$ và $t^{\alpha+\beta-1} = t^0 = 1$. Đó là điều cần chứng minh.

Tính chất cuối cùng của hàm thuần nhất được nêu chi tiết trong định lý Euler, đôi khi gọi là định lý cộng (Adding-up Theorem): Hàm thuần nhất có thể viết được theo các đạo hàm riêng của nó. Ta cũng nhận được kết quả quan trọng đối với hàm tuyến tính thuần nhất.

Định lý 2.14. Định lý Euler (Euler's Theorem)

1. *Nếu* $f(\mathbf{x})$ *là hàm thuần nhất bậc* k *thì*

$$kf(\mathbf{x}) = \sum_{j=1}^{n} \frac{\partial f(x)}{\partial x_j} x_j$$

2. *Nếu* $f(\mathbf{x})$ *là hàm thuần nhất bậc* 1 *thì*

$$f(\mathbf{x}) = \sum_{j=1}^{n} \frac{\partial f(x)}{\partial x_j} x_j .$$

Chứng minh. Giả thiết $f(\mathbf{x})$ là hàm thuần nhất bậc k. Theo định nghĩa

$$t^k f(\mathbf{x}) \equiv f(t\mathbf{x}) \ \forall t > 0.$$

Cách chứng minh là xem đồng nhất thức này như một hàm của t, rồi lấy vi phân hai vế của nó theo t. Trước hết lấy vi phân vế trái ta được

$$kt^{k-1} f(\mathbf{x}) \qquad \text{(P.1)}$$

Khi lấy vi phân vế phải đối với t ta cần nhớ rằng f phụ thuộc n biến và t tác động vào tất cả n biến này. Ta cần xem hàm f ở dạng $f(g_1(t), \dots , g_n(t))$, trong đó $g_j(t) \equiv tx_j$. áp dụng qui tắc lấy đạo hàm của hàm hợp ta được

$$\sum_{j=1}^{n} \frac{\partial f(tx_1,\dots,tx_n)}{\partial x_j} \frac{\partial (tx_j)}{\partial t}$$

Nhưng $\partial(tx_j)/\partial t = x_j$, vì thế biểu thức này trở thành

$$\sum_{j=1}^{n} \frac{\partial f(tx)}{\partial x_j} x_j . \qquad \text{(P.2)}$$

Phép lấy vi phân bảo toàn đẳng thức, vì thế (P.1) và (P.2) bằng nhau:

$$kt^{k-1}f(\mathbf{x}) = \sum_{j=1}^{n} \frac{\partial f(tx)}{\partial x_j} x_j .$$

Bất đẳng thức này đúng với mọi $t > 0$. Đặt $t = 1$ ta được

$$kf(\mathbf{x}) = \sum_{j=1}^{n} \frac{\partial f(tx)}{\partial x_j} x_j .$$

Đó là điều ta muốn chứng minh. Phần hai là trường hợp riêng khi $k = 1$.

✍ Tóm lại, chương này đã trình bày khái quát về hàm thực nhiều biến số và một số tập liên quan mật thiết với hàm (đồ thị, các tập mức), đồng thời phân tích các hàm thường gặp trong nghiên cứu kinh tế và tối ưu hoá (hàm lồi, lõm, hàm tựa lồi, tựa lõm, hàm thuần nhất …) cùng với các tính chất đặc trưng của chúng. Cuối cùng, xét tính khả vi của hàm số và liên hệ giữa tính khả vi với tính lồi, tính lõm hay tính thuần nhất của hàm.

Chương 3
BÀI TOÁN TỐI ƯU

Chương này đề cập tới cách tiếp cận giải tích cho các bài toán tối ưu, một dạng bài toán thường gặp trong nhiều nghiên cứu và phân tích kinh tế. Xét các bài toán không ràng buộc và có ràng buộc. Giới thiệu khái quát các điều kiện tối ưu cần và đủ và trình bày phương pháp Lagrange thông dụng. Nội dung chính của chương dựa chủ yếu trên các tài liệu [1], [3] và [5].

3.1. CỰC TRỊ CỦA HÀM SỐ

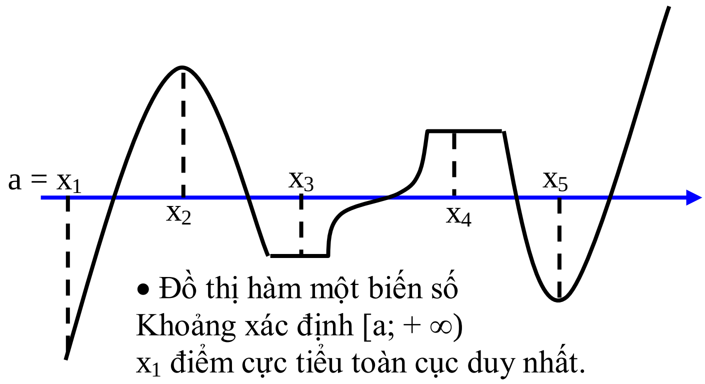
Xét hàm của một hay nhiều biến số $y = f(x)$ và giả thiết hàm này khả vi (hàm trơn) bậc một hoặc bậc hai tuỳ theo yêu cầu.

Ta nói hàm f đạt **cực tiểu địa phương** tại điểm x^* nếu $f(x^*) \le f(x)$ với mọi x trong một lân cận nào đó của x^* (chẳng hạn $\| x - x^* \| < \varepsilon$). Ta nói hàm f đạt **cực tiểu toàn cục** tại điểm x^* nếu $f(x^*) \le f(x)$ với mọi x trong miền xác định của hàm. Hàm f đạt **cực tiểu địa phương chặt** tại điểm x^* nếu $f(x^*) < f(x)$ với mọi x trong lân cận của $x^*, x \ne x^*$. Hàm f đạt **cực tiểu toàn cục duy nhất** tại điểm x^* nếu $f(x^*) < f(x)$ với mọi x trong miền xác định, $x \ne x^*$.

Tương tự, ta nói hàm f đạt **cực đại địa phương** (**cực đại địa phương chặt**) tại điểm $\tilde{x}$ nếu $f(\tilde{x}) \ge f(x)$ ($f(\tilde{x}) > f(x)$) với mọi x trong một lân cận nào đó của $\tilde{x}$ (chẳng hạn $\| x - \tilde{x} \| < \varepsilon$). Ta nói hàm f đạt **cực đại toàn cục** (**cực đại toàn cục duy nhất**) tại điểm $\tilde{x}$ nếu $f(\tilde{x}) \ge f(x)$ ($f(\tilde{x}) > f(x)$) với mọi x trong miền xác định của hàm, $x \ne$ [redacted].

Nếu điểm cực tiểu x^* (điểm cực đại $\tilde{x}$) là một điểm trong của miền xác định thì ta nói đó là điểm cực tiểu (cực đại) **bên trong** (interior minima/ maxima). Còn nếu đó là một điểm biên của miền xác định thì ta nói đó là điểm cực tiểu (cực đại) **trên biên** (boundary minima/ maxima).

Hình 3.1 giúp ta hình dung rõ hơn các khái niệm này



• Đồ thị hàm một biến số
Khoảng xác định [a; + ∞)
x_1 điểm cực tiểu toàn cục duy nhất.
(Không có cực đại toàn cục)
x_2 điểm cực đại địa phương chặt
x_3 điểm cực tiểu địa phương (không duy nhất)
x_4 điểm cực đại địa phương (không duy nhất)
x_5 điểm cực tiểu địa phương chặt

Hình 3.1. Cực tiểu (cực đại) địa phương (toàn cục)

Trong lý thuyết kinh tế, người ta ít khi cần tới tính toán điểm tối ưu (cực tiểu hay cực đại) mà thường chỉ muốn mô tả đặc trưng của những điểm này để nêu ra điều kiện phải thoả mãn tại điểm tối ưu (gọi là điều kiện cần của tối ưu) rồi sau đó làm việc với các điều kiện này hơn là với các con số cụ thể.

3.2. TỐI ƯU KHÔNG RÀNG BUỘC(Uncontrained Optimization)

Định lý 3.1. Điều kiện cần của tối ưu địa phương - trường hợp 1 biến.

Giả sử f(x) *là hàm một biến, khả vi. Khi đó,* f(x) *đạt*

a) *cực tiểu địa phương tại* $x^* \Rightarrow f'(x^*) = 0$ (điều kiện cần cấp 1)

$\Rightarrow f''(x^*) \geq 0$ (điều kiện cần cấp 2)

b) *cực đại địa phương tại* $\tilde{x} \Rightarrow f'(\tilde{x}) = 0$ (điều kiện cần cấp 1)

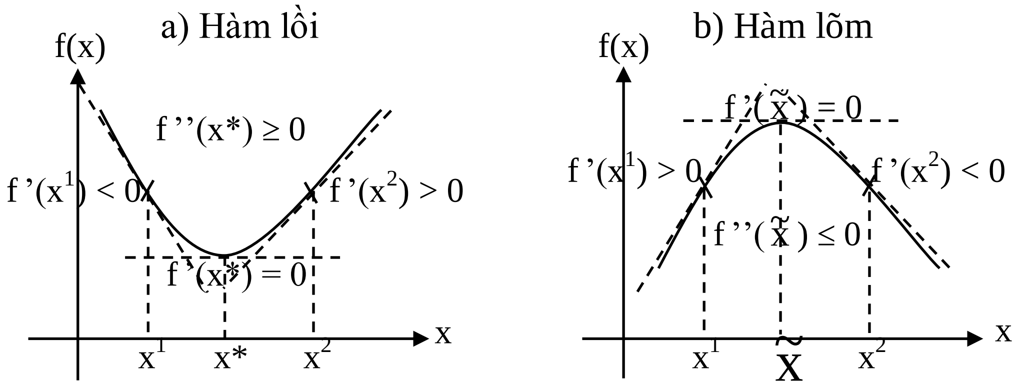
$\Rightarrow f'(\tilde{x}) \leq 0$ (điều kiện cần cấp 2).

Với hàm một hay nhiều biến, cực tiểu địa phương của hàm lồi (lồi chặt) luôn trùng với cực tiểu toàn cục của hàm đó và cực đại địa phương của hàm lõm (lõm chặt) luôn trùng với cực đại toàn cục của hàm đó.

Định lý 3.2. Định lý tối ưu địa phương & toàn cục (không ràng buộc)

a) *Giả sử* f(**x**) *là hàm lồi. Khi đó,* f(**x**) *đạt cực tiểu địa phương tại điểm* **x*** ⇔ f(**x**) *đạt cực tiểu toàn cục tại* **x***.

b) *Giả sử* f(**x**) *là hàm lõm. Khi đó,* f(**x**) *đạt cực đại địa phương tại điểm* $\tilde{x} \Leftrightarrow$ f(**x**) *đạt cực đại toàn cục tại* $\tilde{x}$.



Hình 3.2. **a)** $f'(x^*) = 0$, $f'(x)$ tăng dần; **b)** $f'(\tilde{x}) = 0$, $f'(x)$ giảm dần

Chứng minh. a) Điều kiện cần là hiển nhiên, vì mỗi điểm cực tiểu toàn cục cũng là điểm cực tiểu địa phương. Ta chứng minh điều kiện đủ bằng phản chứng: giả sử x* là điểm cực tiểu địa phương của f nhưng x* không là điểm cực tiểu toàn cục, dựa vào tính lồi của f ta sẽ chỉ ra mâu thuẫn.

Thật vậy, giả sử D là miền xác định của f. Do x* là cực tiểu địa phương của f nên tìm được e > 0 sao cho $f(x^*) \le f(x)$ với mọi $x \in D$ thoả mãn $\|x - x^*\| < e$. Nếu x* không là cực tiểu toàn cục của f trên D thì tìm được $\bar{x} \in D$ sao cho $f(\bar{x}) < f(x^*)$ hay $f(\bar{x}) - f(x^*) < 0$. Đặt $x^t = (1 - t)x^* + t\bar{x}$, $0 \le t \le 1$. Khi đó, $x^t \in D$ (giả thiết D lồi) và $\|x^t - x^*\| = t\|\bar{x} - x^*\| < e$ với t > 0 đủ nhỏ. Do f là hàm lồi và x* là điểm cực tiểu địa phương của f nên với t > 0 đủ nhỏ ta có

$$f(x^t) \le (1 - t)f(x^*) + tf(\bar{x}) = f(x^*) + t[f(\bar{x}) - f(x^*)] < f(x^*),$$

nghĩa là $f(x^t) < f(x^*)$, trái với x* là điểm cực tiểu địa phương. Vậy nếu x* là điểm cực tiểu địa phương của f thì x* phải là điểm cực tiểu toàn cục của f.

b) Chứng minh tương tự.

Định lý 3.2 cho thấy với tính lồi hoặc lõm, bất kỳ điểm tối ưu địa phương nào cũng là điểm tối ưu toàn cục, vì thế chỉ có duy nhất một giá trị nhỏ nhất và một giá trị lớn nhất của hàm. Tuy nhiên, giá trị nhỏ nhất (lớn nhất) có thể đạt được tại nhiều điểm thuộc miền xác định. Nếu ta muốn giá trị nhỏ nhất (lớn nhất) của hàm đạt tại duy nhất một điểm thì ta phải giả thiết thêm hàm là lồi chặt hay lõm chặt.

Định lý 3.3. Tính lồi / lõm chặt và tính duy nhất của tối ưu toàn cục

a) *Giả sử* $f(\mathbf{x})$ *là hàm lồi chặt. Nếu* $\mathbf{x}^*$ *đạt cực tiểu của* $f(\mathbf{x})$ *thì* $\mathbf{x}^*$ *là điểm cực tiểu toàn cục duy nhất và* $f(\mathbf{x}^*) < f(\mathbf{x})\ \forall \mathbf{x} \in D$.

b) *Giả sử* $f(\mathbf{x})$ *là hàm lõm chặt. Nếu* $\tilde{x}$ *đạt cực đại của* $f(\mathbf{x})$ *thì* $\tilde{x}$ *là điểm cực đại toàn cục duy nhất và* $f(\tilde{x}) > f(\mathbf{x})\ \forall \mathbf{x} \in D$.

Chứng minh. Ta chứng minh định lý cho hàm lồi chặt bằng phản chứng. Trường hợp còn lại chứng minh tương tự. Giả sử $\mathbf{x}^*$ là cực tiểu toàn cục của f, nhưng $\mathbf{x}^*$ không duy nhất. Khi đó, tìm được điểm $\mathbf{x} \neq \mathbf{x}^*$ sao cho $f(\mathbf{x}) = f(\mathbf{x}^*)$. Đặt $\mathbf{x}^t = t\mathbf{x} + (1 - t)\mathbf{x}^*$ với $0 < t < 1$, tính lồi chặt của hàm f kéo theo

$$f(\mathbf{x}^t) < tf(\mathbf{x}) + (1 - t)f(\mathbf{x}^*) \text{ với mọi } t \in (0, 1).$$

Do $f(\mathbf{x}) = f(\mathbf{x}^*)$ nên bất đẳng thức này cho thấy

$$f(\mathbf{x}^t) < tf(\mathbf{x}^*) + (1 - t)f(\mathbf{x}^*) = f(\mathbf{x}^*) \Rightarrow f(\mathbf{x}^t) < f(\mathbf{x}^*),$$

trái với $\mathbf{x}^*$ là cực tiểu toàn cục của f. Vậy điểm cực tiểu toàn cục của một hàm lồi chặt phải duy nhất.

3.2.1. ĐIỀU KIỆN CẤP MỘT(First – Order Conditions)

Cho $f : |R^n \to |R$. Nếu x^* là điểm trong tối ưu thì f không thể tăng hay giảm đối với mọi thay đổi nhỏ $dx_i \neq 0$ của bất kỳ biến i nào, $i = 1, \dots, n$. Như vậy, ta có

$$dy = \nabla f(\mathbf{x}^*).d\mathbf{x} = 0\ \forall d\mathbf{x} = (dx_1, \dots, dx_n)^T \neq \mathbf{0}.$$

Từ đó suy ra $\nabla f(\mathbf{x}^*) = 0$. Hệ thức này đặc trưng cho điểm trong tối ưu của hàm nhiều biến. Đó cũng là điều kiện cần cấp một cho điểm trong tối ưu địa phương.

Định lý 3.4. Điều kiện cần cấp một cho điểm trong tối ưu địa phương của hàm thực nhiều biến

Giả sử $f(\mathbf{x})$ *là hàm khả vi. Nếu* $f(\mathbf{x})$ *đạt cực tiểu (cực đại) địa phương tại điểm trong* $\mathbf{x}^*$ *thì* $\mathbf{x}^*$ *nghiệm đúng hệ phương trình:*

$$\frac{\partial f(x^*)}{\partial x_1} = 0,\ \frac{\partial f(x^*)}{\partial x_2} = 0,\ \dots,\ \frac{\partial f(x^*)}{\partial x_n} = 0.$$

Chứng minh. Mặc dù ở trên đã chỉ ra tính đúng đắn của định lý này, song ở đây ta sẽ đưa ra một chứng minh khác. Giả sử f(**x**) đạt cực trị địa phương tại điểm trong **x*** và tìm cách chỉ ra rằng ∇f(**x***) = 0. Ta sẽ nêu ra chứng minh kiến thiết nhờ dùng các qui tắc quen thuộc trong giải tích đối với hàm một biến. Để bắt đầu ta chọn véctơ gia số bất kỳ d**x** ≠ **0**. Khi đó, với mọi t ta lập hàm một biến

$$g(t) = f(\mathbf{x}^* + td\mathbf{x}) = f(x_1^* + dx_1, \dots, x_n^* + tdx_n). \qquad \text{(P.1)}$$

Khi t ≠ 0, **x*** + td**x** là véctơ khác **x***, vì thế g(t) trùng với giá trị nào đó của f khác f(**x***). Khi t = 0, **x*** + td**x** trùng với **x***, vì thế g(0) trùng với giá trị f tại **x***. Do g(t) trùng với giá trị f nào đó với mọi t và trùng với f(x*) khi t = 0 nên g(0) đạt cực trị địa phương tại t = 0 (vì đã giả thiết f đạt cực trị tại **x***). Theo Định lý 3.1 g'(0) = 0. Lấy đạo hàm của (P.1) theo t ta có

$$g'(t) = \sum_{i=1}^{n} \frac{\partial f(x^* + tdx)}{\partial x_i} dx_i,$$

với mọi t. Nếu ta tính tại t = 0 và áp dụng điều kiện g'(0) = 0 cực trị địa phương của g tại 0 kéo theo

$$g'(0) = \sum_{i=1}^{n} \frac{\partial f(x^*)}{\partial x_i} dx_i = \nabla f(\mathbf{x}^*)d\mathbf{x} = 0.$$

Do d**x** là véctơ khác 0 tuỳ ý nên đẳng thức trên kéo theo ∇f(**x***) = 0 ∀dx ≠ 0.

Ví dụ 3.1. Tính điểm dừng của hàm 2 biến $f = 2x_1^2 + 3x_2^2 - 2x_1x_2 - 10x_2$.

Giải. Lấy đạo hàm riêng của f theo biến x_1, x_2 và cho chúng bằng 0:

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = 4x_1 - 2x_2 = 0,\ \frac{\partial f(x_1, x_2)}{\partial x_2} = 6x_2 - 2x_1 - 10 = 0.$$

Giải hệ phương trình này ta được nghiệm

$$x_1^* = 1,\ x_2^* = 2.$$

Vậy hàm đã cho có điểm dừng tại điểm **x*** = (1, 2). Tuy nhiên ta chưa biết liệu đó có phải là điểm cực tiểu hay cực đại không? Muốn thế, ta cần xét điều kiện cấp 2.

3.2.2. ĐIỀU KIỆN CẤP HAI(Second – Order Conditions)

Về đại thể, điều kiện cấp hai trong trường hợp nhiều biến cũng giống như trường hợp một biến. Khi tìm thấy điểm tại đó $\nabla f(\mathbf{x}) = 0$, ta biết đó là điểm cực tiểu nếu tại đó hàm là lồi địa phương và ta biết đó là điểm cực đại nếu tại đó hàm là lõm địa phương. Định lý 3.3 cho thấy rằng độ cong phụ thuộc sự thay đổi của y là tăng hay giảm hoặc phụ thuộc vào dấu của vi phân cấp hai $d^2y = d\mathbf{x}^T\mathbf{H}(\mathbf{x})d\mathbf{x}$. Hàm là lồi (địa phương) quanh **x** nếu dạng toàn phương này không âm và là lõm (địa phương) nếu nó không dương ở gần **x**. Như vậy, trực giác gợi ý điều kiện cần cấp hai sau đây cho điểm trong tối ưu địa phương.

Định lý 3.5. Điều kiện cần cấp hai cho điểm trong tối ưu địa phương của hàm thực nhiều biến

Giả sử y = f(**x**) *hai lần khả vi.*

a) *Nếu* f(**x**) *đạt cực tiểu địa phương bên trong tại* **x*** *thì*

$$d^2y = d\mathbf{x}^T\mathbf{H}(\mathbf{x}^*)d\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}^*(x^*)dx_i dx_j \geq 0 \ \forall dx.$$

b) *Nếu* f(**x**) *đạt cực đại địa phương bên trong tại* $\tilde{x}$ *thì*

$$d^2y = d\mathbf{x}^T\mathbf{H}(\tilde{x})d\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}^*(\tilde{x})dx_i dx_j \leq 0 \forall dx.$$

Chứng minh. Ta có thể thiết lập trực tiếp từ chứng minh Định lý 3.4. Nhớ rằng ta đã xây dựng hàm một biến

$$g(t) = f(\mathbf{x} + td\mathbf{x})$$

với $d\mathbf{x} \neq \mathbf{0}$ và **x** là điểm dừng của f. Ta nhận xét là nếu f có điểm dừng tại **x** thì g có điểm dừng tại t = 0. Hơn nữa, với mọi t

$$g'(t) = \sum_{i=1}^{n} \frac{\partial f(x + tdx)}{\partial x_i} dx_i .$$

Lấy đạo hàm một lần nữa theo t ta nhận được đạo hàm cấp hai

$$g''(t) = \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f(x + tdx)}{\partial x_i \partial x_j} dx_i dx_j . \qquad \text{(P.1)}$$

Bây giờ giả sử f đạt cực tiểu tại **x** = **x***. Theo Định lý 3.1, g''(0) ≥ 0. Đánh giá (P.1) tại **x*** và t = 0 ta được

$$g''(0) = \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f(x^*)}{\partial x_i \partial x_j} dx_i dx_j \ \geq 0.$$

Hoàn toàn tương tự, nếu f đạt cực đại tại **x** = $\tilde{x}$ thì g''(0) ≤ 0. Vì thế

$$g''(0) = \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f(\tilde{x})}{\partial x_i \partial x_j} dx_i dx_j \ \leq 0.$$

Định lý đã được chứng minh đầy đủ. □

✍ Các Định lý 3.4 và 3.5 quan trọng và hữu ích. Ta có thể dùng các định lý này để mô tả tính cách một điểm trong tối ưu mỗi khi ta biết hay giả sử nó tồn tại. Điều kiện cần nói rằng "nếu **x*** đạt cực tiểu của f(**x**) thì

$$f'_i(\mathbf{x}^*) = 0,\ i = 1, \ldots, n, \text{ và}$$

$$d^2y = d\mathbf{x}^T\mathbf{H}(\mathbf{x}^*)d\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} f''_{ij}(x^*)dx_i dx_j \ \geq 0\text{''}.$$

Tuy nhiên, các điều kiện này không giúp ta tìm ra điểm cực tiểu (hay cực đại) của một hàm cụ thể nào đó. Muốn thế, ta cần tới các điều kiện **đủ**.

Định lý 3.6. Điều kiện đủ cho tính lồi chặt / lõm chặt của hàm thực

Giả sử f(**x**) 2 *lần khả vi,* D_i(**x**) - *tử thức chính thứ* i *của ma trận Hess* H(x).

a) *Nếu* $D_i(x) > 0$, i = 1, … , n, *thì* f(**x**) *lồi chặt tại* **x**,

b) *Nếu* $(-1)^n D_n(x) > 0$ *thì* f(**x**) *lõm chặt tại* **x**.

Nếu điều kiện 1 *hay* 2 *nêu trên đúng với mọi* **x** *thuộc miền xác định thì hàm là lồi chặt* (*hay lõm chặt*) *toàn cục.*

Nói cách khác, hàm f là **lồi chặt** tại x nếu mọi tử thức chính của ma trận Hess của f tại x có dấu dương. Hàm f là **lõm chặt** tại x nếu các tử thức chính của ma trận Hess của f tại x đan dấu, bắt đầu từ dấu âm.

Chứng minh xem [3] tr. 83 – 84.

Bây giờ ta sẽ phát biểu điều kiện đủ cấp một và cấp hai cho điểm trong tối ưu địa phương. Các điều kiện này được suy ra trực tiếp từ những điều kiện đã thiết lập, vì thế chúng không cần phải chứng minh. Ta chỉ đơn giản ghép các kết quả đã có lại với nhau và viết ra để tiện theo dõi, trích dẫn về sau.

Định lý 3.7. Điều kiện đủ cho điểm trong tối ưu địa phương của hàm thực nhiều biến.

Giả sử hàm $f(\mathbf{x})$ *hai lần khả vi*:

a) *Nếu* $f_i^{'}(\mathbf{x}^*) = 0$ *và* $D_n(\mathbf{x}^*) > 0$, $i = 1, \ldots, n$, *thì* $f(\mathbf{x})$ *đạt cực tiểu địa phương tại* $\mathbf{x}^*$,

b) *Nếu* $f_i^{'}(\tilde{x}) = 0$ *và* $(-1)^n D_i(\tilde{x}) > 0$, $i = 1, \ldots, n$, *thì* $f(\mathbf{x})$ *đạt cực đại địa phương tại* $\tilde{x}$.

Ví dụ 3.2. Ta hãy kiểm tra xem điểm dừng tìm được trong ví dụ 3.1 là điểm cực tiểu hay cực đại? Ta có $f(x_1, x_2) = 2x_1^2 + 3x_2^2 - 2x_1x_2 - 10x_2$ và

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = 4x_1 - 2x_2, \quad \frac{\partial f(x_1, x_2)}{\partial x_2} = 6x_2 - 2x_1 - 10.$$

Điểm dừng $\mathbf{x}^* = (x_1^*, x_2^*) = (1, 2)$. Tính các đạo hàm riêng cấp hai

$$\frac{\partial^2 f}{\partial x_1^2} = 4, \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = -2, \quad \frac{\partial^2 f}{\partial x_2 \partial x_1} = -2, \quad \frac{\partial^2 f}{\partial x_2^2} = 6$$

và lập ma trận Hess

$$\mathbf{H}(\mathbf{x}) = \begin{pmatrix} 4 & -2 \\ -2 & 6 \end{pmatrix}.$$

Kiểm tra các tử thức chính của ma trận này ta thấy

$$D_1(\mathbf{x}) = 4 > 0,$$

$$D_2(\mathbf{x}) = \begin{vmatrix} 4 & -2 \\ -2 & 6 \end{vmatrix} = 24 - 4 = 20 > 0.$$

Hai tử thức này đều dương. Định lý 3.7 cho thấy điểm dừng x* = (1, 2) là điểm cực tiểu địa phương.

✍ Ma trận Hess trong ví dụ này hoàn toàn không phụ thuộc **x**. Vì thế, các tử thức chính cũng dương như vậy, bất kể ta tính các đạo hàm riêng cấp hai tại điểm nào. Định lý 3.6 cho thấy đó là điều kiện đủ đảm bảo cho hàm nói tới là lồi chặt toàn cục. Ta hãy thử hình dung đồ thị của một hàm như thế trong không gian ba chiều. Nếu đồ thị có điểm thấp thì nó chỉ có thể có duy nhất một điểm thấp và đó là điểm thấp nhất, vì theo Định lý 3.2 mọi cực tiểu địa phương đều là cực tiểu toàn cục và theo Định lý 3.3 điểm cực tiểu toàn cục là duy nhất. Điều này gợi ý những điều kiện đủ sau đây cho cực trị toàn cục của hàm lồi chặt hay hàm lõm chặt.

Định lý 3.8. Điều kiện đủ cho tối ưu toàn cục duy nhất

Giả sử f(**x**) *khả vi*:

a) *Nếu* f(**x**) *là hàm lồi chặt toàn cục và* $f_i^{'}(\mathbf{x}^*) = 0$, i = 1, … , n, *thì* **x*** *là điểm cực tiểu toàn cục duy nhất của* f(**x**).

b) *Nếu* f(**x**) *là hàm lõm chặt toàn cục và* $f_i^{'}(\tilde{x}) = 0$, i = 1, … , n, *thì* $\tilde{x}$ *là điểm cực đại toàn cục duy nhất của* f(**x**).

Chứng minh xem [3], tr. 86 – 87.

3.3. TỐI ƯU CÓ RÀNG BUỘC(Contrained Optimization)

Trên đây ta chỉ mới mô tả đặc trưng cho những điểm tại đó hàm đạt cực trị địa phương khi ta có thể tự do lựa chọn các biến **x**. Trong kinh tế ta thường không có sự tự do đó. Sự thiếu thốn lan tràn khắp đời sống kinh tế và là đặc trưng vượt trội của hầu hết các bài toán kinh tế thực tế. Thậm chí một số người định nghĩa kinh tế là sự nghiên cứu hành vi trong điều kiện thiếu thốn. Sự thiếu thốn thường được biểu hiện ở các **ràng buộc** đối với giá trị được phép của các biến kinh tế. Khi đó, nhiệm vụ của nhà doanh nghiệp là tìm các giá trị tốt nhất có thể cho các biến với các ràng buộc đặt ra cho họ. Đó là loại bài toán ta sẽ gặp thường xuyên. Ta cần sửa đổi kỹ thuật tối ưu hoá và các thuật ngữ đặc trưng cho tối ưu trong những trường hợp như thế một cách tương ứng.

Ta sẽ bàn tới ba loại ràng buộc chính: ràng buộc **đẳng thức**, ràng buộc **không âm** và tổng quát hơn là ràng buộc **bất đẳng thức**. Ta sẽ nêu ra các phương pháp giải bài toán đối với từng loại ràng buộc này. Ta chỉ xét đại diện bài toán cực tiểu và ghi chú những thay đổi (nếu có) đối với bài toán cực đại.

3.3.1. RÀNG BUỘC ĐẲNG THỨC(Equality Contraints)

Trước hết xét bài toán tối ưu với ràng buộc đẳng thức của hai biến

$$\min_{x_1, x_2} f(x_1, x_2) \text{ với điều kiện } g(x_1, x_2) = 0. \qquad (3.1)$$

ở đây $f(x_1, x_2)$ gọi là **hàm mục tiêu** hay hàm chi phí. x_1, x_2 là **biến lựa chọn** và thường viết dưới toán tử "min" để nhắc ta cần tìm giá trị của x_1, x_2. Còn $g(x_1, x_2)$ gọi là **hàm ràng buộc. Nó** qui định những giá trị nào của các biến được xem là **chấp nhận được** hay **được phép** khi giải bài toán. Tập hợp tất cả x_1, x_2 thoả mãn ràng buộc đôi khi gọi là **tập ràng buộc** hay **tập chấp nhận được**.

Một cách giải đơn giản bài toán này là dùng phép thế. Nếu hàm ràng buộc cho phép giải một biến x_i theo biến còn lại thì ta có thể đưa bài toán có ràng buộc của hai biến về bài toán không ràng buộc của một biến. Chẳng hạn, giả sử từ $g(x_1, x_2) = 0$ có thể viết tách biệt x_2 ở một vế

$$x_2 = \tilde{g}(x_1) \qquad (3.2)$$

Thế trực tiếp biểu thức này vào hàm mục tiêu ta nhận được bài toán

$$\min_{x_1} f(x_1, \tilde{g}(x_1)). \qquad (3.3)$$

Điều kiện cần cấp 1 đòi hỏi ta cho đạo hàm toàn phần df/dx_1 bằng 0 và giải phương trình nhận được để tìm nghiệm cực tiểu x_1^*. Khi lấy vi phân toàn phần (3.3) ta cần nhớ rằng f có hai đạo hàm riêng và phải áp dụng qui tắc lấy vi phân hàm số hợp

$$\frac{df}{dx_1} = \frac{\partial f(x_1^*, \tilde{g}(x_1^*))}{\partial x_1} + \frac{\partial f(x_1^*, \tilde{g}(x_1^*))}{\partial x_2}\frac{d\tilde{g}(x_1^*)}{dx_1} = 0.$$

Sau khi tìm được x_1^* ta thay vào (3.2) để tìm $x_2^* = \tilde{g}(x_1^*)$. Khi đó, cặp (x_1^*, x_2^*) là nghiệm cực tiểu của bài toán có ràng buộc, miễn là điều kiện cấp hai tương ứng được thoả mãn.

Đáng tiếc là nhiều bài toán cần giải có hơn hai biến lựa chọn và bao gồm nhiều ràng buộc, hơn nữa trong nhiều trường hợp hệ ràng buộc lại rất phức tạp, không cho phép ta giải dễ dàng một biến theo các biến khác. Vì thế, phương pháp thế không phải khi nào cũng thích hợp: trong một số trường hợp phương pháp thế có thể thực hiện được, song trong nhiều trường hợp khác phương pháp thế lại không thể áp dụng được. Tuy nhiên, có một cách khác tốt hơn cho phép xử lý hiệu quả một lớp bài toán rộng hơn nhiều.

3.3.1.1.PHƯƠNG PHÁP LAGRANGE (Lagrange's Method)

Phương pháp Lagrange là một phương pháp mạnh hay được sử dụng để giải các bài toán tối ưu có ràng buộc trong kinh tế.

- Xét bài toán tối ưu hai biến và một ràng buộc đẳng thức:

$$\min_{x_1, x_2} f(x_1, x_2) \text{ với điều kiện } g(x_1, x_2) = 0.$$

Thêm biến mới λ và lập hàm Lagrange theo ba biến x_1, x_2 và λ

$$L(x_1, x_2, \lambda) \equiv f(x_1, x_2) + \lambda g(x_1, x_2).$$

Tìm cực tiểu không ràng buộc của hàm $L(.)$ bằng cách lấy đạo hàm của L theo các biến x_1, x_2, λ và cho các đạo hàm đó bằng 0. Cách làm này cho ta

$$\frac{\partial \mathcal{L}}{\partial x_1} = \frac{\partial f(x_1^*, x_2^*)}{\partial x_1} + \lambda\frac{\partial g(x_1^*, x_2^*)}{\partial x_1} = 0 \qquad (3.4)$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = \frac{\partial f(x_1^*, x_2^*)}{\partial x_2} + \lambda\frac{\partial g(x_1^*, x_2^*)}{\partial x_2} = 0 \qquad (3.5)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = g(x_1^*, x_2^*) = 0. \qquad (3.6)$$

Có ba phương trình theo ba biến x_1, x_2 và λ. Phương pháp Lagrange khẳng định rằng nghiệm x_1^*, x_2^*, λ của hệ ba phương trình này là một điểm dừng của hàm $f(x_1, x_2)$ với ràng buộc $g(x_1, x_2) = 0$.

Với nghiệm tìm được ta thấy (x_1^*, x_2^*) thoả mãn ràng buộc của bài toán. Ta sẽ chứng tỏ rằng (x_1^*, x_2^*) đạt cực trị của $f(x_1, x_2)$ với ràng buộc $g(x_1, x_2) = 0$. Thật vậy, lấy vi phân toàn phần của hàm $L(.)$ ta được

$$dL = \frac{\partial \mathcal{L}}{\partial x_1} dx_1 + \frac{\partial \mathcal{L}}{\partial x_2} dx_2 + \frac{\partial \mathcal{L}}{\partial \lambda} d\lambda.$$

Do giả thiết x_1^*, x_2^* và λ thoả mãn điều kiện cấp một (3.4) – (3.6) tại điểm tối ưu của L nên dL tính tại điểm này phải bằng 0.

$$dL = \frac{\partial f(x_1^*, x_2^*)}{\partial x_1} dx_1 + \frac{\partial f(x_1^*, x_2^*)}{\partial x_2} dx_2 + g(x_1^*, x_2^*) d\lambda$$

$$+ \lambda \left[\frac{\partial g(x_1^*, x_2^*)}{\partial x_1} dx_1 + \frac{\partial g(x_1^*, x_2^*)}{\partial x_2} dx_2 \right] = 0 \qquad (3.7)$$

với mọi dx_1, dx_2 và $d\lambda$. Ta sẽ chứng tỏ (3.7) kéo theo $df = 0$ với mọi dx_1, dx_2 được phép, tức là đảm bảo thoả mãn ràng buộc $g(x_1, x_2) = 0$. Do $g(x_1^* + dx_1, x_2^* + dx_2) = 0$ với mọi dx_1, dx_2 nên vi phân toàn phần $dg = 0$, tức là

$$dg = \left[\frac{\partial g(x_1^*, x_2^*)}{\partial x_1} dx_1 + \frac{\partial g(x_1^*, x_2^*)}{\partial x_2} dx_2 \right] = 0. \qquad (3.8)$$

Nhớ rằng $g(x_1^*, x_2^*) = 0$ nên từ (3.7) và (3.8) suy ra

$$dL = \frac{\partial f(x_1^*, x_2^*)}{\partial x_1} dx_1 + \frac{\partial f(x_1^*, x_2^*)}{\partial x_2} dx_2 = 0 \text{ hay } df(x_1^*, x_2^*) = 0$$

với mọi dx_1, dx_2 thoả mãn ràng buộc. Điều này có nghĩa là hàm f có giá trị cực tiểu (cực đại) như nó vốn có, với điều kiện các biến phải thoả mãn ràng buộc và mọi di chuyển khỏi điểm (x_1^*, x_2^*) cũng là di chuyển dọc theo ràng buộc. Như vậy, điều kiện cấp một (3.4) – (3.6) đặc trưng cho “điểm dừng” có ràng buộc của hàm mục tiêu. Tuy nhiên, chỉ với điều kiện cấp một ta chưa thể biết các điểm dừng này là điểm cực tiểu (cực đại) có ràng buộc. Để phân biệt rõ cực tiểu hay cực đại đòi hỏi có hiểu biết thêm về "độ cong" của các hàm mục tiêu và ràng buộc tại điểm dừng đang xét. Ta sẽ bàn tới vấn đề này sau. Bây giờ xét một ví dụ đơn giản.

Ví dụ 3.3. Xét bài toán với ràng buộc đẳng thức và áp dụng phương pháp Lagrange để giải. Giả sử bài toán cần giải có dạng

$$\min_{x_1,x_2}(ax_1^2 + bx_2^2) \text{ với điều kiện } x_1 + x_2 - 1 = 0, \tag{E.1}$$

trong đó a > 0 và b > 0. Trước hết ta xây dựng hàm Lagrange

$$L(x_1, x_2, \lambda) \equiv (ax_1^2 + bx_2^2) + \lambda(x_1 + x_2 - 1).$$

Cho mọi đạo hàm riêng bậc nhất của hàm này bằng 0

$$\frac{\partial \mathcal{L}}{\partial x_1} = 2ax_1 + \lambda = 0 \tag{E.2}$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = 2bx_2 + \lambda = 0 \tag{E.3}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = x_1 + x_2 - 1 = 0. \tag{E.4}$$

Giải hệ phương trình này đối với x_1, x_2 và λ ta được

$$x_1 = \frac{b}{a+b},\ x_2 = \frac{a}{a+b},\ \lambda = -\frac{2ab}{a+b}. \tag{E.5}$$

Nhân tử Lagrange λ chỉ là "phụ". x_1, x_2 trong (E.5) là điểm "dừng" có khả năng là nghiệm của bài toán (E.1). Giá trị hàm mục tiêu tương ứng bằng

$$y^* = a\left(\frac{b}{a+b}\right)^2 + b\left(\frac{a}{a+b}\right)^2 = \frac{ab^2 + ba^2}{(a+b)^2} = \frac{ab}{a+b}. \tag{E.6}$$

Nhớ rằng nếu chỉ dựa vào điều kiện cấp 1 thì ta chưa thể biết được giá trị hàm mục tiêu tại điểm

$$\left(\frac{b}{a+b}, \frac{a}{a+b}\right)$$

là lớn nhất hay nhỏ nhất khi có ràng buộc.

• Bây giờ ta mô tả phương pháp Lagrange cho các hàm có số biến tuỳ ý, trong các bài toán có số ràng buộc tuỳ ý, miễn là số ràng buộc ít hơn số biến.

Xét bài toán tìm cực tiểu của hàm n biến với m ràng buộc (m < n) dạng

$$\min_{x_1,\dots,x_n} f(x_1, \dots, x_n) \text{ với } g_1(x_1, \dots, x_n) = 0, \dots, g_m(x_1, \dots, x_n) = 0. \qquad (3.10)$$

Để giải bài toán ta lập hàm Lagrange bằng cách nhân ràng buộc g_j với nhân tử Lagrange λ_j rồi cộng vào hàm mục tiêu f. Với $\mathbf{x} = (x_1, \dots, x_n)$ và $\Lambda = (\lambda_1, \dots, \lambda_m)$ ta nhận được hàm của m + n biến

$$L(x, \Lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x). \qquad (3.11)$$

Điều kiện cấp một yêu cầu mọi đạo hàm riêng cấp một của L bằng 0 tại điểm tối ưu. Do L có n + m biến nên có tất cả n + m phương trình để xác định n + m biến $\mathbf{x}^*$ và Λ. Cụ thể là

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_i} &= \frac{\partial f(x^*)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x^*)}{\partial x_i} = 0, i = 1, \dots, n, \\ \frac{\partial \mathcal{L}}{\partial \lambda_j} &= g_j(x^*) = 0, j = 1, \dots, m. \end{aligned} \qquad (3.12)$$

Về nguyên tắc có thể giải hệ phương trình này theo n + m biến x* và Λ. Khi đó, véctơ x* có thể là nghiệm tối ưu của bài toán có ràng buộc (3.10).

Phương pháp Lagrange rất hữu ích. Trên thực tế nó là một thuật toán để tìm nghiệm tối ưu có ràng buộc cho một lớp rộng các bài toán thực tiễn. Định lý sau nêu ra những điều kiện cho phép tìm được x* và Λ.

Định lý 3.9. Định lý Lagrange (Lagrange's Theorem)

Giả sử f($\mathbf{x}$) *và* $g_j(\mathbf{x})$, j = 1, … , m, *là những hàm thực hai lần khả vi liên tục trên miền* $D \subset \mathbb{R}^n$. *Giả sử* x* *là một điểm trong của* D *và* x* *là điểm tối ưu của hàm* f($\mathbf{x}$) *với các ràng buộc* $g_j(\mathbf{x}) = 0$, j = 1, … , m. *Nếu* m < n *và nếu các véctơ* *gradiên* $\nabla g_j(\mathbf{x}^*)$, j = 1, … , m, *độc lập tuyến tính thì tồn tại duy nhất* m *số* λ_j, j = 1, … , m, *sao cho hàm Lagrange L có điểm tối ưu theo* **x** *tại* **x*** *và*

$$\frac{\partial \mathcal{L}(x^*,\Lambda)}{\partial x_i} = \frac{\partial f(x^*)}{\partial x_i} + \sum_{j=1}^{m}\lambda_j \frac{\partial g_j(x^*)}{\partial x_i} = 0,\ i = 1, \dots, n.$$

3.3.1.2. Ý NGHĨA HÌNH HỌC(Geometrical Interpretation)

Trở lại xét bài toán (3.1). Về hình học, ta có thể biểu diễn hàm mục tiêu bằng các tập mức của nó $L(\alpha) \equiv \{(x_1, x_2) \mid f(x_1, x_2) = \alpha\}$ với số α nào đó thuộc miền trị. Tất cả các điểm trên tập mức này cần thoả mãn phương trình

$$f(x_1, x_2) = \alpha$$

Nếu ta thay đổi x_1, x_2 và không rời khỏi tập mức này thì các vi phân dx_1, dx_2 cần giữ cho giá trị hàm f không đổi ở mức a, tức là cần thoả mãn:

$$\frac{\partial f(x_1,x_2)}{\partial x_1}dx_1 + \frac{\partial f(x_1,x_2)}{\partial x_2}dx_2 = 0, \tag{3.13}$$

đẳng thức này nhận được bằng cách lấy vi phân toàn phần hai vế của phương trình xác định tập mức và nhớ rằng vi phân toàn phần của hằng số a bằng 0. Đẳng thức trên cần được thoả mãn tại một điểm bất kỳ trên tập mức bất kỳ của hàm mục tiêu.

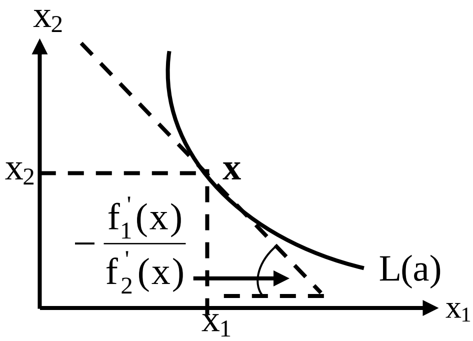
Ta có thể tính độ dốc của một đường mức bất kỳ tại điểm dừng nào đó. Giải (3.13) theo dx_2/dx_1, độ dốc của tập mức qua (x_1, x_2) sẽ là

$$\left.\frac{dx_2}{dx_1}\right|_{\text{doc theo } L(y^0)} = -\frac{f_1'(x_1,x_2)}{f_2'(x_1,x_2)}. \tag{3.14}$$

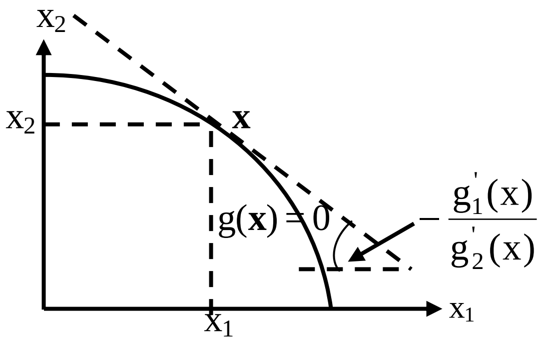
Ký hiệu $|_{\text{dọc theo} \dots}$ được dùng để ta nhớ loại thay đổi đặc biệt đang xét của dx_1, dx_2. Vậy, như vẽ ở Hình 3.3, độ dốc của tập mức qua điểm bất kỳ (x_1, x_2) cho bởi (số đối) tỉ số giữa các đạo hàm riêng cấp một của f tại điểm (x_1, x_2).

Cùng vậy, giả sử ràng buộc g(x) = 0 có dạng như ở Hình 3.2, vẽ trên cùng mặt phẳng với tập mức. Ta có thể xem ràng buộc như một loại tập mức. Đó là tập tất cả các điểm (x_1, x_2) thoả mãn $g(x_1, x_2) = 0$.

\



Hình 3.3. Độ dốc của tập mức



Hình 3.4. Độ dốc của ràng buộc

Tương tự, với mọi điểm (x_1, x_2) thoả mãn ràng buộc thì hệ thức

$$\frac{\partial g(x_1,x_2)}{\partial x_1}dx_1 + \frac{\partial g(x_1,x_2)}{\partial x_2}dx_2 = 0,$$

cần được thoả mãn đối với mọi thay đổi dx_1, dx_2 dọc theo ràng buộc. Độ dốc của ràng buộc tại điểm (x_1, x_2) sẽ là

$$\left.\frac{dx_2}{dx_1}\right|_{\text{doc theo } g(.)=0} = -\frac{g_1'(x_1,x_2)}{g_2'(x_1,x_2)}. \qquad (3.15)$$

Mặt khác, ta có thể viết lại các điều kiện (3.4) – (3.6) dưới dạng

$$\frac{\partial f(x_1^*,x_2^*)}{\partial x_1} = -\lambda\frac{\partial g(x_1^*,x_2^*)}{\partial x_1}$$

$$\frac{\partial f(x_1^*,x_2^*)}{\partial x_2} = -\lambda\frac{\partial g(x_1^*,x_2^*)}{\partial x_2}$$

$$g(x_1^*, x_2^*) = 0.$$

Khử λ từ hai hệ thức đầu ta nhận được điều kiện xác định 2 biến x_1^*, x_2^*.

$$\frac{f_1'(x_1^*,x_2^*)}{f_2'(x_{1*}^*,x_2^*)} = \frac{g_1'(x_1^*,x_2^*)}{g_2'(x_{1*}^*,x_2^*)} \qquad (3.16)$$

$$g(x_1^*, x_2^*) = 0. \qquad (3.17)$$

Hai điều kiện này nói gì? Vế trái (3.16) bằng (- 1) lần độ dốc của tập mức đối với hàm mục tiêu qua điểm (x_1^*, x_2^*). Vế phải (3.16) bằng (- 1) lần độ dốc của tập mức đối với hàm ràng buộc. Điều kiện (3.16) cho thấy nghiệm (x_1^*, x_2^*)

là điểm tại đó độ dốc của tập mức đối với hàm mục tiêu bằng độ dốc của tập mức đối với hàm ràng buộc. Điều kiện (3.17) cho thấy ta cần phải ở trên tập mức của phương trình ràng buộc. Điểm trên tập ràng buộc có độ dốc của tập mức mục tiêu bằng độ dốc của tập mức ràng buộc, theo định nghĩa, là điểm **tiếp xúc** (point of tangency) giữa ràng buộc và tập mức mục tiêu.

3.3.1.3. ĐIỀU KIỆN CẤP HAI(Second – Order Conditions)

• Trước hết xét bài toán tối ưu hai biến, một ràng buộc.

Xem x_1 như biến tự do và x_2 như hàm của x_1. Ràng buộc có dạng: $g(x_1, x_2(x_1)) = 0$. Lấy vi phần toàn phần theo dx_2/dx_1 ta được hệ thức quen thuộc:

$$\frac{dx_2}{dx_1} = -\frac{g_1'}{g_2'} \tag{3.18}$$

đối với độ dốc của hệ thức ràng buộc trong mặt phẳng (x_1, x_2). Đặt $y = f(x_1, x_2(x_1))$ là giá trị hàm mục tiêu có ràng buộc, ta xem y như hàm của một biến x_1. Lấy vi phân đối với x_1 ta được $dy/dx_1 = f_1' + f_2'(dx_2/dx_1)$. Chú ý đến (3.18) ta được

$$\frac{dy}{dx_1} = f_1' - f_2'\frac{g_1'}{g_2'} \tag{3.19}$$

Lấy vi phân lần nữa và nhớ ràng x_2 là hàm của x_1 ta được vi phân cấp hai:

$$\begin{aligned}\frac{d^2y}{dx_1^2} = f''_{11} + f''_{12}\frac{dx_2}{dx_1} - \left[f''_{21} + f''_{22}\frac{dx_2}{dx_1}\right]\frac{g_1'}{g_2'} \\ - f_2'\left[\frac{g_2'(g''_{11} + g''_{12}\frac{dx_2}{dx_1}) - g_1'(g''_{21} + g''_{22}\frac{dx_2}{dx_1})}{(g_2')^2}\right]\end{aligned} \tag{3.20}$$

Điều kiện cần cấp hai cho cực tiểu hàm một biến đòi hỏi đạo hàm cấp hai này lớn hơn hay bằng 0 tại điểm thoả mãn các điều kiện cấp một. Điều kiện đủ đòi hỏi bất đẳng thức này được thoả mãn chặt tại điểm đó. Các điều kiện cấp một (3.4) – (3.6) đòi hỏi $f_1' = -\lambda g_1'$ và $f_2' = -\lambda g_2'$. Định lý Young cho thấy $f''_{12} = f''_{21}$ và $g''_{12} = g''_{21}$. Chú ý tới (3.18) ta viết lại (3.20) dưới dạng

$$\frac{d^2y}{dx_1^2} = \frac{1}{(g_2')^2}[(f_{11} + \lambda g_{11})(g_2)^2 - 2(f_{12} + \lambda g_{12})g_1g_2 + (f_{22} + \lambda g_{22})(g_1)^2]. \quad (3.21)$$

Từ $L = f + \lambda g$ ta suy ra các đạo hàm riêng của hàm Lagrange đối với x_i là

$$L_i = f_i + \lambda g_i$$

Các đạo hàm riêng cấp hai của hàm Lagrange sẽ là

$$L_{11} = f_{11} + \lambda g_{11}$$

$$L_{12} = f_{12} + \lambda g_{12} \quad (3.22)$$

$$L_{22} = f_{22} + \lambda g_{22}$$

Lập ma trận đối xứng $\overline{H} = \begin{pmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} & g_1 \\ \mathcal{L}_{21} & \mathcal{L}_{22} & g_2 \\ g_1 & g_2 & 0 \end{pmatrix}$.

Ma trận này được gọi là ma trận Hessian **biên** (bordered Hessian) của hàm Lagrange vì nó chứa các đạo hàm riêng cấp hai của hàm L bao quanh bởi các đạo hàm riêng cấp một của hàm ràng buộc và số 0. Tính định thức của ma trận này (khai triển theo cột cuối) ta được

$$\overline{D} \equiv \begin{vmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} & g_1 \\ \mathcal{L}_{21} & \mathcal{L}_{22} & g_2 \\ g_1 & g_2 & 0 \end{vmatrix} = [L_{11}(g_2)^2 - 2L_{12}g_1g_2 + L_{22}(g_1)^2]. \quad (3.23)$$

Kết hợp (3.21), (3.22) và (3.23) ta thấy đạo hàm cấp hai của hàm mục tiêu có ràng buộc có thể viết lại theo định thức của ma trận Hessian biên của hàm Lagrange như sau

$$\frac{d^2y}{dx_1^2} = \frac{-1}{(g_2')^2}\overline{D}. \quad (3.24)$$

Như vậy, độ cong của hàm mục tiêu dọc theo ràng buộc đặc trưng bởi dấu của đạo hàm cấp hai d^2y/dx_1^2 có thể suy ra trực tiếp từ dấu định thức của ma trận Hessian biên của hàm Lagrange (giả thiết $g_2 \neq 0$). Đến đây ta có thể phát biểu điều kiện **đủ** cho bài toán hai biến, một ràng buộc.

Định lý 3.10. Điều kiện đủ đối với bài toán tối ưu hai biến, một ràng buộc (Sufficient Conditions for the Two-Variables, One-Constraint Optimization Problem)

Nếu (x_1^*, x_2^*, λ) *nghiệm đúng điều kiện cấp* ***1*** (3.4) – (3.6) *và nếu* $\overline{D} < 0$ (> 0) *trong* (3.23) *tại điểm* (x_1^*, x_2^*, λ) *thì* (x_1^*, x_2^*) *là cực tiểu* (*cực đại*) *của hàm* $f(x_1, x_2)$ *với ràng buộc* $g(x_1, x_2) = 0$.

Ví dụ 3.4. Hãy xét xem điểm dừng nhận được từ Ví dụ 3.1 (tr. ?) là cực tiểu hay cực đại. Dễ thấy rằng $L_{11} = 2a$, $L_{12} = L_{21} = 0$, $L_{22} = 2b$. Từ phương trình ràng buộc $g_1' = 1$ và $g_2' = 1$. Xây dựng ma trận Hessian biên và tính định thức của nó

$$\overline{D} = \begin{vmatrix} 2a & 0 & 1 \\ 0 & 2b & 1 \\ 1 & 1 & 0 \end{vmatrix} = -2(a + b) < 0.$$

Vì ở đây $\overline{D} < 0$ với mọi x_1, x_2 và λ, nên cũng có $\overline{D} < 0$ đối với nghiệm (E.5) của điều kiện cấp 1 trong Ví dụ 3.1. Do đó, giá trị hàm mục tiêu (E.6) là giá trị cực tiểu có ràng buộc. □

- Trường hợp nhiều biến, nhiều ràng buộc: ma trận Hessian biên có dạng

$$\overline{H} = \begin{pmatrix} \mathcal{L}_{11} & \cdots & \mathcal{L}_{1n} & g_1^1 & \cdots & g_1^m \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}_{n1} & \cdots & \mathcal{L}_{nn} & g_n^1 & \cdots & g_n^m \\ g_1^1 & \cdots & g_n^1 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ g_1^m & \cdots & g_n^m & 0 & \cdots & 0 \end{pmatrix}, \qquad (3.25)$$

trong đó L_{kj} là các đạo hàm riêng cấp hai của hàm Lagrange L theo x_k, x_j và g_j^i là đạo hàm riêng của hàm g_i theo x_j (i = 1, 2, … , m; j, k = 1, 2, … , n).

Các tử thức chính là định thức của các ma trận con nhận được từ ma trận H nhờ di chuyển từ trên xuống và từ trái sang phải theo đường chéo chính. Các tử thức chính ta quan tâm ở đây bắt đầu từ định thức cấp 3 và kết thúc bởi định thức cấp (n + m).

$$\overline{D}_3 = \begin{vmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} & g_1^1 \\ \mathcal{L}_{21} & \mathcal{L}_{22} & g_2^1 \\ g_1^1 & g_2^1 & 0 \end{vmatrix}, \ \overline{D}_4 = \begin{vmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} & \mathcal{L}_{13} & g_1^1 & g_1^2 \\ \mathcal{L}_{21} & \mathcal{L}_{22} & \mathcal{L}_{23} & g_2^1 & g_2^2 \\ \mathcal{L}_{31} & \mathcal{L}_{32} & \mathcal{L}_{33} & g_3^1 & g_3^2 \\ g_1^1 & g_2^1 & g_3^1 & 0 & 0 \\ g_1^2 & g_2^2 & g_3^2 & 0 & 0 \end{vmatrix}, \ \dots, |\overline{H}|. \quad (3.26)$$

Ta có thể tóm tắt các điều kiện **đủ** tối ưu trong trường hợp tổng quát ở định lý sau.

Định lý 3.11. Điều kiện đủ cho tối ưu với ràng buộc đẳng thức

(Sufficient Conditions for Optima with Equality Constraints)

Giả sử f(**x**) *là hàm mục tiêu và* m *ràng buộc là* $g_i(\mathbf{x}) = 0$, i = 1, … , m. *Giả sử hàm Lagrange cho bởi* (3.11). *Giả sử* ($\mathbf{x}^*$, Λ) *nghiệm đúng điều kiện cấp một trong* (3.12). *Khi đó*:

a) $\mathbf{x}^*$ *đạt cực tiểu có ràng buộc của* f(x) *nếu các tử thức chính trong* (3.26) *đều âm* $D_3 < 0$, $D_4 < 0$, … *khi tính tại* ($\mathbf{x}^*$, Λ).

b) $\mathbf{x}^*$ *đạt cực đại có ràng buộc của* f(x) *nếu các tử thức chính trong* (3.26) *luân phiên đổi dấu bắt đầu từ dấu dương* $D_3 > 0$, $D_4 < 0$, … *khi tính tại* ($\mathbf{x}^*$, Λ).

3.3.2. RÀNG BUỘC KHÔNG ÂM (Non-negativity Constraints)

Trong các ứng dụng kinh tế ta thường gặp bài toán cực tiểu (cực đại) với ràng buộc bất đẳng thức, thay thế hoặc bổ sung cho ràng buộc đẳng thức. Ràng buộc bất đẳng thức đơn giản nhất là ràng buộc không âm $\mathbf{x} \geq \mathbf{0}$ (các biến kinh tế lấy giá trị không âm). Để giúp hiểu rõ các bài toán phức tạp hơn, ta xét trường hợp một biến (Có ba trường hợp xảy ra như đã chỉ ra ở Hình 3.5).

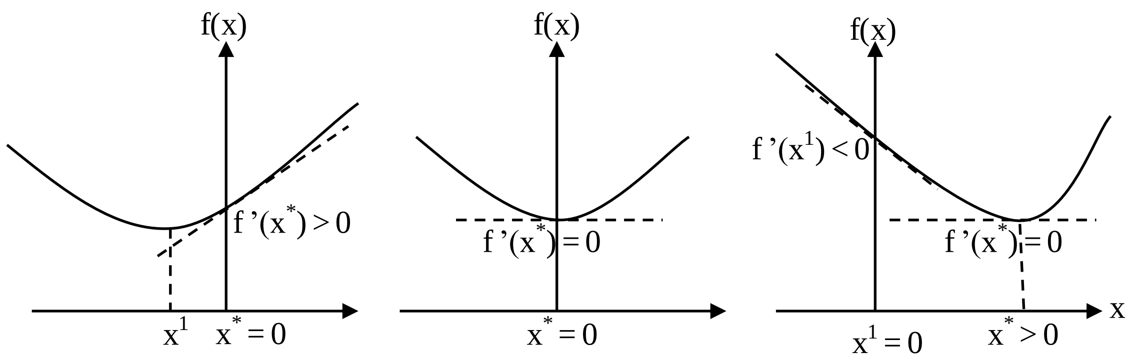
$$\min_x f(x) \text{ với điều kiện } x \geq 0. \quad (3.27)$$

Nghiệm x* của bài toán (3.27) cần thoả mãn ba điều kiện sau:

+ Điều kiện 1. $f'(\mathbf{x}^*) \geq 0$

+ Điều kiện 2. $\mathbf{x}^*.f'(\mathbf{x}^*) = 0$ (3.28)

+ Điều kiện 3. $\mathbf{x}^* \geq \mathbf{0}$.



Hình 3.5. Cực tiểu với ràng buộc không âm

Ví dụ 3.5. Xét bài toán $\min_x \{x^2 + 4x - 2\}$ với điều kiện x ≥ 0.

Lấy vi phân ta được f'(x) = 2x + 4. Từ 3.28 x* cần thoả mãn

1. - 2x* - 4 ≥ 0

2. x*[2x* + 4] = 0

3. x* ≥ 0.

Từ các điều kiện này cho thấy x* = 0 là nghiệm cực tiểu.

• Điều kiện cho cực đại của f(**x**) với **x** ≥ 0 cũng dễ dàng được nêu ra: Nếu $\tilde{x}$ là điểm cực đại của bài toán với ràng buộc không âm **x** ≥ **0** thì

+ Điều kiện 1. $f'(\tilde{x}) \le 0$

+ Điều kiện 2. $\tilde{x}[f'(\tilde{x})] = 0$ (3.29)

+ Điều kiện 3. $\tilde{x} \ge \mathbf{0}$.

Trong trường hợp nhiều biến, ba điều kiện trên cần đúng với từng biến riêng biệt và các đạo hàm riêng của hàm thay cho đạo hàm duy nhất. Định lý sau là một mở rộng trực tiếp của trường hợp một biến.

Định lý 3.12. Điều kiện cần tối ưu của hàm với ràng buộc không âm (Necessary Conditions for Optima of Real Valued Functions Subject to Non-negativity Constraints)

Giả sử f(**x**) *là hàm khả vi liên tục. Khi đó,*

1. *Nếu* **x*** *đạt cực tiểu của* f(**x**) *với điều kiện* **x** ≥ 0 *thì* **x*** *thoả mãn*

(1) $\dfrac{\partial f(x^*)}{\partial x_i} \geq 0,\ i = 1, \ldots, n$

(2) $x_i^*[\dfrac{\partial f(x^*)}{\partial x_i}] = 0,\ i = 1, \ldots, n$

(3) $x_i^* \geq 0,\ i = 1, \ldots, n.$

2. *Nếu* $\tilde{x}$ *đạt cực đại của* f(**x**) *với điều kiện* **x** ≥ 0 *thì* $\tilde{x}$ *thoả mãn*

(1) $\dfrac{\partial f(\tilde{x})}{\partial x_i} \leq 0,\ i = 1, \ldots, n$

(2) $\tilde{x}_i[\dfrac{\partial f(\tilde{x})}{\partial x_i}] = 0,\ i = 1, \ldots, n$

(3) $\tilde{x}_i \geq 0,\ i = 1, \ldots, n.$

3.3.3. ĐIỀU KIỆN KARASH-KUHN-TUCKER(KKT Conditions)

Cho đến nay trên thực tế ta chưa sử dụng đến phương pháp Lagrange, bởi vì ràng buộc bất đẳng thức được xét là đơn giản. Bây giờ ta xét bài toán với ràng buộc bất đẳng thức phức tạp hơn.

$$\min_{x_1, x_2} f(x_1, x_2) \text{ với điều kiện } g(x_1, x_2) \leq 0,\ x_1 \geq 0,\ x_2 \geq 0. \qquad (3.30)$$

Bài toán này được gọi là bài toán **quy hoạch phi tuyến** (non-linear programming problem). Thêm biến mới z ≥ 0 để đưa bài toán (3.30) về dạng:

$$\min_{x_1, x_2, z} f(x_1, x_2) \text{ với điều kiện } g(x_1, x_2) + z = 0,\ x_1 \geq 0,\ x_2 \geq 0,\ z \geq 0. \quad (3.31)$$

Định lý 3.1 cho thấy cực tiểu theo **x** của f với ràng buộc đẳng thức trùng với cực tiểu không điều kiện theo **x** của hàm Lagrange tương ứng khi không có ràng buộc về dấu. Định lý 3.4 cho biết phải thay đổi thế nào điều kiện cấp một đối với cực tiểu không ràng buộc của hàm Lagrange để tính đến các ràng buộc không âm này. Để có thể vận dụng các định lý trên, trước hết ta xây dựng hàm Lagrange cho bài toán (3.31):

$$L(x_1, x_2, z, \lambda) \equiv f(x_1, x_2) + \lambda[g(x_1, x_2) + z].$$

Ta sẽ mô tả đặc trưng cho điểm cực tiểu của hàm Lagrange L với các ràng buộc $x_1 \geq 0$, $x_2 \geq 0$, $z \geq 0$. Điều kiện cấp một đối với x_1, x_2 và z có dạng

$$L_1 \equiv f_1 + \lambda g_1 \geq 0 \qquad \text{(i)}$$

$$x_1 L_1 \equiv x_1[f_1 + \lambda g_1] = 0 \qquad \text{(ii)}$$

$$L_2 \equiv f_2 + \lambda g_2 \geq 0 \qquad \text{(iii)}$$

$$x_2 L_2 \equiv x_2[f_2 + \lambda g_2] = 0 \qquad \text{(iv)}$$

$$L_z \equiv \lambda \geq 0 \qquad \text{(v)}$$

$$z L_z \equiv \lambda z = 0 \qquad \text{(vi)}$$

$$x_1 \geq 0,\ x_2 \geq 0,\ z \geq 0. \qquad \text{(vii)}$$

Điều kiện cấp một đối với λ, khuyết điều kiện không âm, đơn giản chỉ là

$$L_\lambda \equiv g(x_1, x_2) + z = 0. \qquad \text{(viii)}$$

Từ các điều kiện (v) – (viii) suy ra

$$g(x_1, x_2) \leq 0$$

$$\lambda g(x_1, x_2) = 0$$

$$\lambda \geq 0.$$

Kết hợp các điều kiện này với (i) – (iv) ta nhận được các điều kiện gọi là **điều kiện Karush - Kuhn - Tucker** (hay đơn giản là điều kiện KKT):

$$f_1 + \lambda g_1 \geq 0 \qquad \text{(i)}$$

$$x_1[f_1 + \lambda g_1] = 0 \qquad \text{(ii)}$$

$$f_2 + \lambda g_2 \geq 0 \qquad \text{(iii)}$$

$$x_2[f_2 + \lambda g_2] = 0 \qquad \text{(iv)}$$

$$g(x_1, x_2) \leq 0 \qquad \text{(v')}$$

$$\lambda g(x_1, x_2) = 0 \qquad \text{(vi')}$$

$$x_1 \geq 0,\ x_2 \geq 0,\ \lambda \geq 0. \qquad \text{(vii')}$$

Ta bàn kỹ các điều kiện này: Các điều kiện từ (i) – (iv) cùng với hai điều kiện đầu của (vii') là những điều kiện cấp một thông thường cho cực tiểu của L với ràng buộc không âm trên x_1, x_2. Các điều kiện (v'), (vi') và điều kiện cuối của (vii') không thể xem như điều kiện cho cực tiểu của L theo λ, bởi vì bất đẳng thức trong (v') "đảo ngược dấu". Song nếu nhìn lại Định lý 3.4 ta biết sẽ nhận được gì nếu ta định tìm cực đại của L theo λ với ràng buộc không âm. Đúng thế, các điều kiện (v'), (vi') và điều kiện cuối của (vii') chính là điều kiện $L_\lambda \le 0$, $\lambda L_\lambda = 0$ và $\lambda \ge 0$ mà ta sẽ nhận được nếu ta định tìm cực đại của L theo λ với điều kiện không âm $\lambda \ge 0$.

Bây giờ có thể thấy rằng (i) – (vii') là các điều kiện cần hay điều kiện cho điểm cực tiểu của hàm Lagrange theo các biến x_i và cho cực đại của hàm Lagrange theo nhân tử λ. Nếu tại một điểm nào đó (x_1^*, x_2^*, λ), L đạt cực tiểu theo x_1 và x_2, đồng thời L đạt cực đại theo λ, thì (x_1^*, x_2^*, λ) được gọi là **điểm yên ngựa** (saddle point) của hàm Lagrange. Tất nhiên điều này được mở rộng cho trường hợp bài toán có nhiều biến và nhiều ràng buộc, miễn là các hàm ràng buộc cần thoả mãn một số điều kiện nhất định nào đó (gọi là **điều kiện chính qui**).

Ta cũng có thể đưa ra các điều kiện tương tự, nhưng không đồng nhất, đối với bài toán cực đại với ràng buộc bất đẳng thức và ràng buộc không âm. Với bài toán cực đại ta qui ước viết (các) bất đẳng thức ràng buộc ở dạng $g(.) \ge 0$, chứ không phải ở dạng $g(.) \le 0$ như đã làm khi xét bài toán cực tiểu. Lúc này ta có thể áp dụng cùng những lập luận và cùng phương pháp để tìm thấy rằng điểm yên ngựa của hàm Lagrange trùng với nghiệm của bài toán cực đại có ràng buộc. Tuy nhiên, lúc này điểm yên ngựa bao gồm cực đại của hàm Lagrange theo các biến quyết định và cực tiểu theo các nhân tử Lagrange.

Ta tổng kết các kết quả này trong định lý sau.

Định lý 3.13. Điều kiện cần tối ưu KKT của hàm thực với ràng buộc bất đẳng thức và ràng buộc không âm (KKT Necessary Conditions for Optima of Real Valued Functions Subject to Inequality and Non-negativity Constraints)

Giả sử f(x) *hai lần khả vi liên tục.*

1. *Xét bài toán cực tiểu*:

$$\min_{x} f(\mathbf{x}) \text{ với điều kiện } g_i(\mathbf{x}) \le 0,\ i = 1, \dots, m, \text{ và } x \ge 0 \qquad \text{(T.1)}$$

với hàm Lagrange

$$L(x, \Lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) \qquad \text{(T.2)}$$

Nếu **x*** *là nghiệm của* (T.1) *và nếu các véctơ gradiên của những ràng buộc chặt tại* **x*** *độc lập tuyến tính thì tồn tại* m *số* $\lambda_j^* \ge 0$, j = 1, … , m *sao cho* $(\mathbf{x}^*, \Lambda^*)$ *là điểm yên ngựa của hàm Lagrange và thoả mãn điều kiện Karush - Kuhn - Tucker*:

$$L_i(\mathbf{x}^*, \Lambda^*) \ge 0 \text{ và } x_j^* L_i(\mathbf{x}^*, \Lambda^*) = 0,\ j = 1, \dots, n$$

$$L_{\lambda_i}(\mathbf{x}^*, \Lambda^*) \le 0 \text{ và } \lambda_i^* L_{\lambda_i}(\mathbf{x}^*, \Lambda^*) = 0,\ i = 1, \dots, m.$$

2. *Xét bài toán cực đại*:

$$\max_{x} f(\mathbf{x}) \text{ với điều kiện } g_i(\mathbf{x}) \ge 0,\ i = 1, \dots, m, \text{ và } \mathbf{x} \ge \mathbf{0} \qquad \text{(T.3)}$$

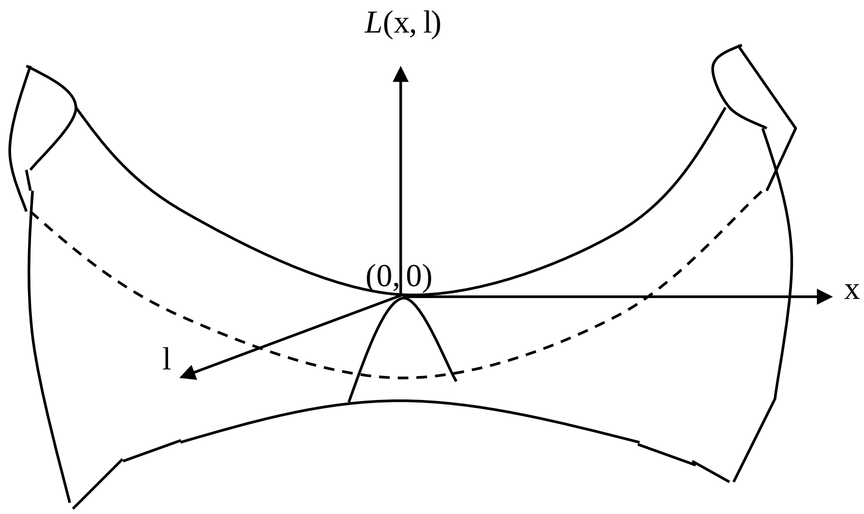
với hàm Lagrange tương ứng (T.2). *Nếu* $\tilde{x}$ *là nghiệm của* (T.3) *và nếu các véctơ gradiên của những ràng buộc chặt tại* $\tilde{x}$ *độc lập tuyến tính thì tồn tại* m *số* $\tilde{\lambda}_i \ge 0$, i = 1, … , m *sao cho* $(\tilde{x}, \tilde{\Lambda})$ *là điểm yên ngựa của hàm Lagrange và thoả mãn điều kiện Karush - Kuhn - Tucker*:

$$L_i(\tilde{x}, \tilde{\Lambda}) \le 0 \text{ và } \tilde{x}_i L_j(\tilde{x}, \tilde{\Lambda}) = 0,\ j = 1, \dots, n$$

$$L_{\lambda_i}(\tilde{x}, \tilde{\Lambda}) \ge 0 \text{ và } \tilde{\lambda}_i L_{\lambda_i}(\tilde{x}, \tilde{\Lambda}) = 0,\ i = 1, \dots, m.$$

Chứng minh đầy đủ định lý này có thể tìm trong Luenberger (1973).

Điều kiện Karush - Kuhn - Tucker nêu trong Định lý 3.13 là các điều kiện **cần**. Chúng cho biết những điều kiện gì cần được thoả mãn khi ta biết (hay giả thiết) có một nghiệm tối ưu cho bài toán quy hoạch phi tuyến. Vì thế, chúng không cho ta một "thuật toán" để giải thực sự bài toán. Muốn thế cần có thêm các điều kiện **đủ** và chúng thực sự khó ở trường hợp tổng quát này. Tuy nhiên, trong một số trường hợp riêng khi hàm mục tiêu lồi (lõm) hay tựa lồi (lõm) thì các điều kiện đủ như thế đã được bàn tới. Với các nhà kinh tế thì các điều kiện cần nêu trong Định lý 3.13 là tạm đủ.



Hình 3.6. Điểm yên ngựa của hàm Lagrange

✍ Tóm lại, trong chương này chúng tôi đã trình bày khái quát về vấn đề tìm cực trị của các hàm số một hay nhiều biến số và giới thiệu tương đối đày đủ các khái niệm và kiến thức tối ưu cơ bản, chủ yếu dưới hình thức phi toán, các kết quả chính được ghi thành các định lý, phần lớn chúng được giải thích và minh hoạ thông qua nhiều ví dụ số và hình vẽ cụ thể.

Đáng chú ý là các điều kiện cần và điều kiện đủ, cấp 1 và cấp 2, cho điểm cực tiểu hay cực đại của hàm số khi có hay không có ràng buộc. Phương pháp quen biết tìm cực trị là phương pháp nhân tử Lagrange và tổng quát hơn là phương pháp dùng điều kiện cần KKT, kết hợp với các điều kiện đủ tối ưu.

KẾT LUẬN

Hàm (rộng hơn là ánh xạ) là một trong những khái niệm cơ bản của giải tích toán học. Nói riêng, hàm thực nhiều biến được sử dụng rộng rãi trong nhiều lĩnh vực khác nhau của khoa học và kỹ thuật. Nhiều tính chất đáng quí của hàm được khai thác triệt để và là giả thiết không thể thiếu trong nhiều nghiên cứu: tính liên tục, tính khả vi và tính chất cực trị của hàm.

Luận văn này nhằm tập trung tìm hiểu những kiến thức giải tích và tối ưu hoá cơ bản liên quan đến các hàm nhiều biến số, cần dùng trong phân tích và nghiên cứu kinh tế về mặt định lượng (bổ sung cho các nghiên cứu định tính).

Chương 1 giới thiệu tóm tắt một số kiến thức giải tích cơ bản về tập hợp và ánh xạ: tập mở, tập đóng, tập compact trong R^n, cận trên (cận dưới) của tập hợp số thực, tập lồi và tính chất; tính liên tục của ánh xạ, quan hệ giữa tính liên tục với ảnh ngược của các tập mở (đóng), ảnh liên tục của tập compact ...

Chương 2 đề cập tới các hàm số thường gặp trong kinh tế và trong các tính toán tối ưu: hàm lồi, hàm lõm, hàm thuần nhất. Khảo sát tính tăng (giảm), tính lồi (lõm), độ dốc, độ cong của hàm qua các tập liên quan mật thiết với hàm (đồ thị, tập mức, tập mức trên, dưới), qua đạo hàm và vi phân của hàm.

Chương 3 trình bày khái quát về cực trị của hàm số nhiều biến số và các kiến thức tối ưu cơ bản: điều kiện cần (điều kiện đủ) đối với điểm cực trị trong các bài toán tối ưu có hay không có ràng buộc, phương pháp Lagrange cho tối ưu với ràng buộc đẳng thức và mở rộng cho tối ưu với ràng buộc bất đẳng thức.

Tác giả đã cố gắng sắp xếp và trình bày vấn đề theo cách hiểu rõ ràng và trực quan nhất có thể, đưa ra các ví dụ và hình vẽ để minh hoạ cho nhiều khái niệm và sự kiện được đề cập tới trong luận văn.

Hy vọng luận văn này sẽ là một tài liệu tham khảo bổ ích cho các đối tượng không chuyên sâu về toán muôn tìm hiểu và vận dụng công cụ giải tích, đặc biệt là các phương pháp tối ưu trong chuyên môn của mình.

TÀI LIỆU THAM KHẢO

Tiếng Việt

[1] N. T. B. Kim (2008), *Giáo trình các phương pháp tối ưu* (*Lý thuyết và thuật toán*), Nxb Bách khoa - Hà Nội.

[2] Đ. V. Lưu và P. H. Khải (2000), *Giải tích lồi*, Nxb Khoa học và Kỹ thuật Hà Nội.

Tiếng Anh

[3] G. A. Jehle (1995), *Advanced Microeconomic Theory*, Part I, Prentice Hall, Englewood Cliffs, New Jersey.

[4] W. F. Trench (2003), *Introduction to Real Analysis*, Free Edition, Library of Congress Cataloging-in-Publication Data.

[5] D. G. Luenberger and Y. Ye (2008), *Linear and Nonlinear Programming*, 3rd Edition, Springer.