

ĐẠI HỌC BÁCH KHOA HÀ NỘI

KHOA TOÁN - TIN



**Phương pháp hướng giảm gradient cho bài toán
tối ưu đa mục tiêu không ràng buộc:
lí thuyết và thử nghiệm**

ĐỒ ÁN II

Chuyên ngành: Toán Tin

Chuyên sâu: Toán ứng dụng

Giảng viên:

TS. Phạm Thị Hoài

Sinh viên thực hiện:

Hoàng Kim Khánh

Lớp:

Toán Tin 01 K66

Mã số sinh viên:

20216839

HÀ NỘI - 2024

NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN

1. Mục tiêu và nội dung của đề án

- Mục tiêu: Tìm hiểu về phương pháp hướng giảm gradient giải bài toán tối ưu đa mục tiêu không ràng buộc.
- Nội dung: Trình bày phương pháp hướng giảm gradient giải bài toán tối ưu đa mục tiêu không ràng buộc trong các trường hợp không lồi, lồi và lồi mạnh. Tính toán thử nghiệm một số bài toán ứng với các trường hợp.

2. Kết quả đạt được

- Nắm được các kiến thức liên quan đến bài toán tối ưu đa mục tiêu không ràng buộc.
- Hiểu và trình bày phương pháp hướng giảm gradient trong các trường hợp không lồi, lồi và lồi mạnh.
- Lập trình thử nghiệm các thuật toán cho một số ví dụ cụ thể.

3. Ý thức làm việc của sinh viên

Hà Nội, ngày ... tháng ... năm 2024

Giảng viên hướng dẫn

TS. Phạm Thị Hoài



K

Danh sách lớp học / Đồ án II / 746363

Thông tin lớp học mã 746363

Kì học : 20241

Mã học phần : MI3390

Tên học phần : Đồ án II

Mã lớp : 746363

Đồ án

Giáo viên hướng dẫn : Phạm Thị Hoài

Tên đồ án : Phương pháp hướng giảm gradient cho bài toán tối ưu đa mục tiêu không ràng buộc và ứng dụng

Nội dung : Tìm hiểu sự hội tụ của phương pháp hướng giảm gradient cho bài toán tối ưu đa mục tiêu không ràng buộc. Tiến hành lập trình thử nghiệm thuật toán để thấy được sự hiệu quả của thuật toán đối với một số tình huống cụ thể như: các hàm mục tiêu lồi, lồi mạnh hoặc không lồi.

Các mốc kiểm soát chính :

Giáo viên phản biện :

Danh sách đánh giá đồ án

Ngày đánh giá	Lần	Nội dung kế hoạch	Nội dung đã thực hiện	Điểm tích cực	Điểm nội dung	Ghi chú
28/10/2024	1	Lên dàn ý cho đồ án và tìm hiểu các chứng minh chi tiết của các thuật toán hướng giảm giải bài toán tối ưu đa mục tiêu phi tuyến trong bài báo giảng viên hướng dẫn đã giao.	Lên dàn ý cho đồ án và tìm hiểu các chứng minh chi tiết của các thuật toán hướng giảm giải bài toán tối ưu đa mục tiêu phi tuyến trong bài báo giảng viên hướng dẫn đã giao.	9.5	10	
05/12/2024	2	Triển khai hoàn thiện đồ án về mặt nội dung và lập trình tính toán thử nghiệm trên một số bài toán tối ưu đa mục tiêu lồi và không lồi để đánh giá tính hiệu quả của thuật toán hướng giảm gradient áp dụng cho tối ưu phi tuyến không ràng buộc.	Sinh viên thực hiện đầy đủ các nhiệm vụ được giao.	10	9.5	



Lời cảm ơn

Báo cáo này là kết quả của quá trình tìm hiểu và học tập của em dưới sự hướng dẫn của TS. Phạm Thị Hoài thuộc học phần Đồ án II tại Khoa Toán - Tin, Đại học Bách khoa Hà Nội.

Em xin cảm ơn giảng viên hướng dẫn - cô Phạm Thị Hoài đã giảng dạy tận tình, chi tiết để em có đủ kiến thức và vận dụng vào đồ án này. Trong quá trình tìm hiểu, em đã tích lũy thêm cho mình nhiều kiến thức bổ ích, tinh thần học tập hiệu quả.

Tuy nhiên vì kiến thức của bản thân còn hạn chế nên đồ án không tránh khỏi những thiếu sót. Em rất mong nhận được sự nhận xét, góp ý, phê bình từ phía thầy cô để đồ án được hoàn thiện hơn.

Em cũng xin được bày tỏ sự biết ơn đến các thầy cô tại Khoa Toán - Tin vì đã luôn nhiệt tình giảng dạy và cho em một môi trường học tập tốt. Kính chúc thầy cô thật nhiều sức khỏe, thành công và hạnh phúc. Em xin chân thành cảm ơn.

Hà Nội, ngày ... tháng ... năm 2024

Sinh viên thực hiện

Hoàng Kim Khánh



Mục lục

Danh mục hình vẽ

Danh sách bảng

Danh mục ký hiệu và chữ viết tắt

Lời nói đầu	1
1 Kiến thức chuẩn bị	3
1.1 Kiến thức cơ sở	3
1.2 Bài toán tối ưu đa mục tiêu	4
1.3 Tối ưu Pareto	5
1.4 Phương pháp hướng giảm gradient	6
2 Phương pháp hướng giảm gradient	9
2.1 Trường hợp các hàm mục tiêu không lồi	9
2.1.1 Thuật toán 1	9
2.1.2 Sự hội tụ của Thuật toán 1	10
2.2 Trường hợp các hàm mục tiêu lồi hoặc lồi mạnh	11
2.2.1 Thuật toán 2	11
2.2.2 Sự hội tụ của Thuật toán 2	12
3 Lập trình thử nghiệm	18
Kết luận	32



Danh sách hình vẽ

1.1	Ví dụ tập nghiệm tối ưu Pareto toàn cục, Pareto địa phương, Pareto yếu toàn cục, Pareto yếu địa phương	6
3.1	Kết quả chạy Thuật toán 1 và Thuật toán 2 trong không gian ảnh của Ví dụ 3.0.1	20
3.2	Kết quả chạy Thuật toán 1 và Thuật toán 2 trong không gian ảnh của Ví dụ 3.0.2	21
3.3	Kết quả chạy Thuật toán 1 và Thuật toán 2 trong không gian ảnh của Ví dụ 3.0.3	22
3.4	Kết quả chạy Thuật toán 1 trong không gian ảnh với trường hợp $\alpha = 0$ của Ví dụ 3.0.4	23
3.5	Kết quả chạy Thuật toán 1 trong không gian ảnh với trường hợp $\alpha = 0.5$ của Ví dụ 3.0.4	23
3.6	Kết quả chạy Thuật toán 1 trong không gian ảnh với trường hợp $\alpha = 0.9$ của Ví dụ 3.0.4	24
3.7	Kết quả chạy Thuật toán 1 trong không gian ảnh với trường hợp $\alpha = 0.99$ của Ví dụ 3.0.4	24
3.8	Kết quả chạy Thuật toán 1 trong không gian ảnh với trường hợp $\alpha = 0.999$ của Ví dụ 3.0.4	25
3.9	Kết quả chạy Thuật toán 1 trong không gian ảnh với trường hợp $\alpha = 1$ của Ví dụ 3.0.4	25
3.10	Kết quả chạy Thuật toán 1 trong không gian ảnh của Ví dụ 3.0.5	26
3.11	Kết quả chạy Thuật toán 1 trong không gian ảnh của Ví dụ 3.0.6	27
3.12	Kết quả chạy Thuật toán 1 trong không gian ảnh của Ví dụ 3.0.7	27
3.13	Kết quả chạy Thuật toán 1 trong không gian ảnh của Ví dụ 3.0.8	28
3.14	Kết quả chạy Thuật toán 1 và Thuật toán 2 trong không gian ảnh của Ví dụ 3.0.9	30



Danh sách bảng

3.1	Mô tả các ví dụ được sử dụng để thử nghiệm	19
3.2	Tổng hợp kết quả chạy thuật toán cho các ví dụ	31



Danh mục ký hiệu và chữ viết tắt

\mathbb{R}	tập các số thực
\mathbb{R}^n	không gian n chiều
\mathbb{R}_+	tập các số thực không âm
\mathbb{R}_{++}	tập các số thực dương
$JF(x)$	ma trận Jacobi của hàm F tại x
$\nabla f_i(x)$	gradient của hàm f_i tại x
vđk	với điều kiện
\mathbb{N}_0	tập số tự nhiên khác 0
$B(x^*, \delta)$	quả cầu mở tâm x^* bán kính δ
$\ \cdot\ $	chuẩn Euclid trong \mathbb{R}^n



Lời nói đầu

Trong thời đại công nghệ phát triển như hiện nay, các bài toán tối ưu ngày càng đóng vai trò quan trọng trong nhiều lĩnh vực từ kỹ thuật, khoa học máy tính cho đến kinh tế. Đặc biệt, bài toán tối ưu đa mục tiêu là một trong những bài toán khó, xuất hiện nhiều trong thực tế khi cần tìm ra giải pháp tối ưu cùng lúc cho nhiều mục tiêu. Tuy nhiên, việc giải quyết các bài toán này thường gặp nhiều thách thức, nhất là khi các hàm mục tiêu có sự xung đột với nhau.

Phương pháp hướng giảm gradient là một trong những phương pháp được sử dụng phổ biến để giải quyết các bài toán tối ưu hóa. Trong đề án này, tác giả sẽ tập trung tìm hiểu phương pháp này cho bài toán tối ưu đa mục tiêu không ràng buộc, bao gồm các trường hợp không lồi, lồi và lồi mạnh. Đề án không chỉ cung cấp cái nhìn tổng quan về lý thuyết mà còn thực hiện các thử nghiệm để đánh giá hiệu quả của phương pháp.

Đề án được tham khảo chính trong bài báo "Complexity of gradient descent for multiobjective optimization" [5] của các tác giả J. Fliege, A Ismael F Vaz và Luís Nunes Vicente, được đăng trên tạp chí *Optimization Methods and Software* năm 2019.

Đề án tập trung tìm hiểu các nội dung:

- Bài toán tối ưu đa mục tiêu không ràng buộc.
- Phương pháp hướng giảm gradient: thuật toán và chứng minh sự hội tụ của thuật toán. Nội dung được lấy từ bài báo "Complexity of gradient descent for multiobjective optimization" [5] của các tác giả J. Fliege, A Ismael F Vaz và Luís Nunes Vicente, được đăng trên tạp chí *Optimization Methods and Software* năm 2019.
- Thử nghiệm phương pháp hướng giảm gradient trong [5] trên một số ví dụ lấy từ bài báo "A Barzilai-Borwein descent method for multiobjective optimization problems" [3] của các tác giả Jian Chen, Liping Tang và Xinmin Yang, được đăng trên tạp chí *European Journal of Operational Research* năm 2023.

Với mục tiêu đặt ra như vậy, những nội dung và kết quả chính của đề án được trình bày trong ba chương như sau:

- **Chương 1** - Kiến thức chuẩn bị: Trình bày các kiến thức cơ sở để có thể tiếp cận được bài toán tối ưu đa mục tiêu không ràng buộc.

- **Chương 2** - Phương pháp hướng giảm gradient: Trình bày phương pháp hướng giảm gradient giải bài toán tối ưu đa mục tiêu không ràng buộc trong các trường hợp không lồi, lồi và lồi mạnh.
- **Chương 3** - Lập trình thử nghiệm: Tính toán thử nghiệm thuật toán trên một số bài toán ứng với các trường hợp không lồi, lồi và lồi mạnh.

Cuối cùng là kết luận và đưa ra đánh giá về những kết quả đạt được và hướng phát triển trong tương lai.

Kiến thức chuẩn bị

Chương đầu tiên tác giả sẽ nhắc lại những kiến thức cơ sở của bài toán tối ưu đa mục tiêu không ràng buộc, tối ưu Pareto và các khái niệm liên quan. Các kiến thức được tham khảo chính ở trong cuốn sách "*Nonlinear multiobjective optimization*" của tác giả Kaisa Miettinen [7] xuất bản năm 1999 và cuốn sách "*Introduction to nonlinear optimization: Theory, algorithms, and applications with MATLAB*" của tác giả Amir Beck [2] xuất bản năm 2014.

1.1 Kiến thức cơ sở

Định nghĩa 1.1.1. Hàm số $f : \mathbb{R}^n \rightarrow \mathbb{R}$ gọi là lồi nếu với mọi $x, y \in \mathbb{R}^n$ ta có:

$$f(\beta x + (1 - \beta)y) \leq \beta f(x) + (1 - \beta)f(y) \quad \forall 0 \leq \beta \leq 1.$$

Định nghĩa 1.1.2. Hàm số $f : \mathbb{R}^n \rightarrow \mathbb{R}$ gọi là lồi mạnh với hệ số lồi mạnh μ nếu với mọi $x, y \in \mathbb{R}^n$ ta có:

$$f(y) - f(x) \geq \nabla f(x)^\top (y - x) + \frac{\mu}{2} \|y - x\|^2.$$

Định nghĩa 1.1.3. Hàm số $f : \mathbb{R}^n \rightarrow \mathbb{R}$ là khả vi tại $x^* \in \mathbb{R}^n$ nếu:

$$f(x^* + d) - f(x^*) = \nabla f(x^*)^\top d + \|d\| \varepsilon(x^*, d),$$

trong đó $\nabla f(x^*)$ là gradient của f tại x^* và $\varepsilon(x^*, d) \rightarrow 0$ khi $\|d\| \rightarrow 0$.

Định nghĩa 1.1.4. Hàm số $f : \mathbb{R}^n \rightarrow \mathbb{R}$ là khả vi liên tục tại $x^* \in \mathbb{R}^n$ nếu mọi đạo hàm riêng $\frac{\partial f(x^*)}{\partial x_j}$ ($j = 1, \dots, n$) đều liên tục tại x^* , nghĩa là mọi thành phần của gradient đều liên tục tại x^* .

Định nghĩa 1.1.5. Hàm số $f : \mathbb{R}^n \rightarrow \mathbb{R}$ là liên tục Lipschitz nếu với mọi $x, y \in \mathbb{R}^n$ tồn tại hằng số $L > 0$ thỏa mãn:

$$\|f(x) - f(y)\| \leq L \|x - y\|.$$

Chú ý: Nếu hàm số f liên tục Lipschitz với hằng số L thì cũng liên tục Lipschitz với hằng số $\bar{L} \geq L$. Do đó ta thường chọn L là hằng số nhỏ nhất có thể.

Bổ đề 1.1. [2] Cho hàm số $f : \mathbb{R}^n \rightarrow \mathbb{R}$ khả vi liên tục và ∇f liên tục Lipschitz với hằng số L . Khi đó với mọi $x, y \in \mathbb{R}^n$:

$$f(y) \leq f(x) + \nabla f(x)^\top (y - x) + \frac{L}{2} \|x - y\|^2.$$

Chứng minh. Theo định lý cơ bản của giải tích:

$$f(y) - f(x) = \int_0^1 \langle \nabla f(x + t(y - x)), y - x \rangle dt.$$

Hay tương đương:

$$f(y) - f(x) = \langle \nabla f(x), y - x \rangle + \int_0^1 \langle \nabla f(x + t(y - x)) - \nabla f(x), y - x \rangle dt.$$

Do đó:

$$\begin{aligned} |f(y) - f(x) - \langle \nabla f(x), y - x \rangle| &= \left| \int_0^1 \langle \nabla f(x + t(y - x)) - \nabla f(x), y - x \rangle dt \right| \\ &\leq \int_0^1 |\langle \nabla f(x + t(y - x)) - \nabla f(x), y - x \rangle| dt \\ &\leq \int_0^1 \|\nabla f(x + t(y - x)) - \nabla f(x)\| \cdot \|y - x\| dt \\ &\leq \int_0^1 tL \|y - x\|^2 dt \\ &= \frac{L}{2} \|y - x\|^2. \end{aligned}$$

Hoàn tất chứng minh. □

1.2 Bài toán tối ưu đa mục tiêu

Bài toán tối ưu đa mục tiêu không ràng buộc có dạng như sau:

$$\min_{x \in \mathbb{R}^n} F(x) \equiv (f_1(x), \dots, f_m(x)), \quad (1)$$

trong đó hàm mục tiêu $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ khả vi liên tục và ∇f_i liên tục Lipschitz với hằng số $L_i > 0, i = 1, \dots, m$.

Một số khái niệm:

- \mathbb{R}^n : không gian quyết định
- \mathbb{R}^m : không gian mục tiêu
- $F(x)$: vectơ hàm mục tiêu

Ví dụ 1.2.1.

$$\min_{x \in \mathbb{R}^4} (f_1(x), f_2(x), f_3(x), f_4(x)),$$

trong đó

$$f_1(x) = (2x_1 - 1)^2 + x_2^2,$$

$$f_2(x) = 2(2x_1 - x_2)^2 - x_1^2 + 2x_2^2,$$

$$f_3(x) = 3(2x_2 - x_3)^2 - 2x_2^2 + 3x_3^2,$$

$$f_4(x) = 4(2x_3 - x_4)^2 - 3x_3^2.$$

1.3 Tối ưu Pareto

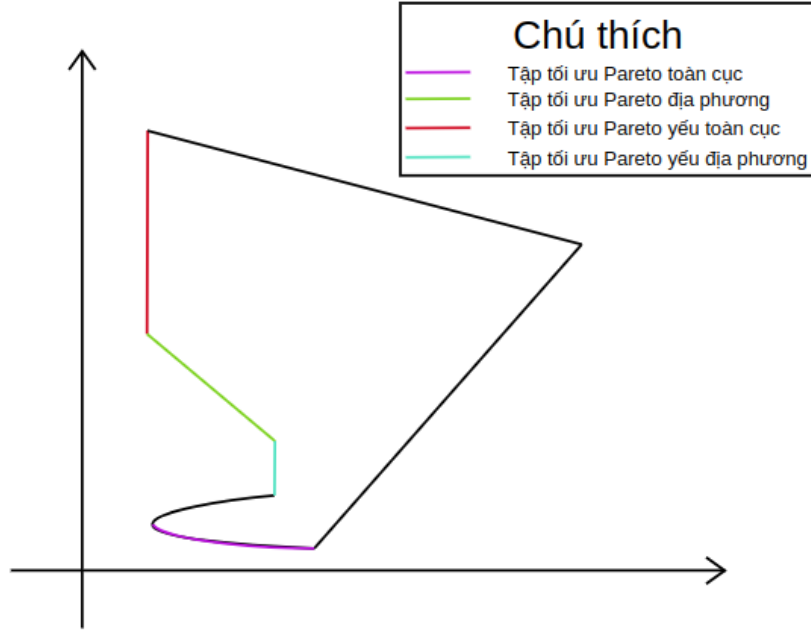
Định nghĩa 1.3.1. Một vectơ $x^* \in \mathbb{R}^n$ là nghiệm tối ưu Pareto toàn cục của bài toán (1) nếu không tồn tại vectơ $x \in \mathbb{R}^n$ thỏa mãn $F(x) \leq F(x^*)$ và $F(x) \neq F(x^*)$.

Định nghĩa 1.3.2. Một vectơ $x^* \in \mathbb{R}^n$ là nghiệm tối ưu Pareto địa phương của bài toán (1) nếu tồn tại $\delta > 0$ thỏa mãn x^* là nghiệm tối ưu Pareto toàn cục trong $B(x^*, \delta)$.

Chú ý: Mọi nghiệm tối ưu Pareto toàn cục đều là nghiệm tối ưu Pareto địa phương. Điều ngược lại chỉ đúng khi là bài toán tối ưu đa mục tiêu không ràng buộc lồi, tức là toàn bộ các hàm mục tiêu phải lồi.

Định nghĩa 1.3.3. Một vectơ $x^* \in \mathbb{R}^n$ là nghiệm tối ưu Pareto yếu toàn cục của bài toán (1) nếu không tồn tại vectơ $x \in \mathbb{R}^n$ thỏa mãn $F(x) < F(x^*)$.

Định nghĩa 1.3.4. Một vectơ $x^* \in \mathbb{R}^n$ là nghiệm tối ưu Pareto yếu địa phương của bài toán (1) nếu tồn tại $\delta > 0$ thỏa mãn x^* là nghiệm tối ưu Pareto yếu toàn cục trong $B(x^*, \delta)$.



Hình 1.1: Ví dụ tập nghiệm tối ưu Pareto toàn cục, Pareto địa phương, Pareto yếu toàn cục, Pareto yếu địa phương

Định nghĩa 1.3.5. Một vectơ $x^* \in \mathbb{R}^n$ được gọi là điểm tối hạn Pareto của bài toán (1) nếu:

$$JF(x^*)(\mathbb{R}^n) \cap (-\mathbb{R}_{++})^m = \emptyset, \quad (2)$$

trong đó:

- $JF(x)$ là ma trận Jacobi của vectơ hàm mục tiêu F .
- $JF(x)(\mathbb{R}^n) := \{JF(x)u : u \in \mathbb{R}^n\}$.
- $(-\mathbb{R}_{++})^m = \{-x : x \in \mathbb{R}_{++}^m\}$.

Nếu một điểm $x \in \mathbb{R}^n$ không phải điểm tối hạn Pareto thì tồn tại một hướng $d \in \mathbb{R}^n$ thỏa mãn:

$$JF(x)d \in (-\mathbb{R}_{++})^m$$

tức d là hướng giảm của F tại x .

1.4 Phương pháp hướng giảm gradient

Phương pháp gradient giải bài toán tối ưu đa mục tiêu xây dựng dãy $\{x^k\}$ theo công thức:

$$x^{k+1} = x^k + t^k d^k,$$

trong đó $t^k > 0$ gọi là cỡ bước và hướng d^k thu được bằng cách giải bài toán con:

$$(d^k, \alpha^k) = \arg \min_{d \in \mathbb{R}^n, \alpha \in \mathbb{R}} \left(\alpha + \frac{1}{2} \|d\|^2 \right) = q(d, \alpha), \quad (3)$$

$$\forall k \nabla f_i(x^k)^\top d \leq \alpha, \quad i = 1, \dots, m.$$

Từ điều kiện KKT của bài toán con (3) ta có:

$$d^k = - \sum_{i=1}^m \lambda_i^k \nabla f_i(x^k), \quad (4)$$

trong đó $\lambda_i^k \geq 0$ là các hệ số Lagrange với các ràng buộc bất đẳng thức tuyến tính trong (3) và:

$$\sum_{i=1}^m \lambda_i^k = 1. \quad (5)$$

Nghiệm của bài toán con (3) có liên quan mật thiết đến điểm tới hạn Pareto của bài toán (1) như trong kết quả dưới đây.

Bổ đề 1.2. [4] Gọi (d^k, α^k) là nghiệm của bài toán con (3).

1. Nếu $x^k \in \mathbb{R}^n$ là điểm tới hạn Pareto của bài toán (1) thì $d^k = 0 \in \mathbb{R}^n$ và $\alpha^k = 0$.
2. Nếu $x^k \in \mathbb{R}^n$ không phải điểm tới hạn Pareto của bài toán (1) thì:

$$\alpha^k \leq -(1/2)\|d^k\|^2 < 0, \quad (6)$$

$$\nabla f_i(x^k)^\top d^k \leq \alpha^k \quad \forall i = 1, \dots, m. \quad (7)$$

Chứng minh.

1. Nếu $x^k \in \mathbb{R}^n$ là điểm tới hạn Pareto của bài toán (1) thì không tồn tại $d \in \mathbb{R}^n$ thỏa mãn:

$$\nabla f_i(x^k)^\top d < 0 \quad \forall i = 1, \dots, m.$$

Do đó tồn tại chỉ số $\bar{i} \in \{1, \dots, m\}$ thỏa mãn:

$$\nabla f_{\bar{i}}(x^k)^\top d^k \geq 0.$$

Kết hợp với ràng buộc của (3) ta phải có:

$$\alpha^k \geq \nabla f_{\bar{i}}(x^k)^\top d^k \geq 0. \quad (*)$$

Mặt khác do $(d, \alpha) = (0, 0)$ thuộc tập chấp nhận được của (3) nên:

$$\begin{aligned} q(d^k, \alpha^k) &\leq q(0, 0) = 0 \\ \Leftrightarrow \alpha^k + \frac{1}{2}\|d^k\|^2 &\leq 0. \end{aligned} \quad (**)$$

Vậy từ (*) và (**) ta phải có $d^k = 0 \in \mathbb{R}^n$ và $\alpha^k = 0$.

2. Nếu $x^k \in \mathbb{R}^n$ không là điểm tối hạn Pareto của bài toán (1) thì tồn tại $d \in \mathbb{R}^n$ thỏa mãn:

$$\nabla f_i(x^k)^\top d < 0 \quad \forall i = 1, \dots, m.$$

Do đó bất đẳng thức bên phải của (6) là đúng. Hiển nhiên (7) cũng đúng bởi vì nó chính là ràng buộc của bài toán (3).

Do $(d, \alpha) = (0, 0)$ thuộc tập chấp nhận được của (3) nên:

$$q(d^k, \alpha^k) \leq q(0, 0) = 0$$

$$\Leftrightarrow \alpha^k + \frac{1}{2} \|d^k\|^2 \leq 0$$

$$\Leftrightarrow \alpha^k \leq -\frac{1}{2} \|d^k\|^2.$$

Hoàn tất chứng minh. □

Kết luận: Chương 1 của đề án đã trình bày một số kiến thức cơ sở của bài toán tối ưu đa mục tiêu không ràng buộc cùng các kết quả liên quan. Chương 2 của đề án sẽ trình bày chi tiết về thuật toán hướng giảm gradient giải bài toán tối ưu đa mục tiêu không ràng buộc.

Phương pháp hướng giảm gradient

Chương này ta sẽ trình bày hai thuật toán giải bài toán tối ưu đa mục tiêu không ràng buộc tương ứng với hai trường hợp, đó là trường hợp các hàm mục tiêu không lồi và trường hợp các hàm mục tiêu lồi hoặc lồi mạnh. Đi kèm với thuật toán sẽ là bổ đề về giới hạn của cỡ bước, tiếp đó là chứng minh sự hội tụ của thuật toán. Các nội dung được tham khảo chính từ bài báo "Complexity of gradient descent for multiobjective optimization" [5] của các tác giả J. Fliege, A. Ismael F Vaz và Luís Nunes Vicente, được đăng trên tạp chí *Optimization Methods and Software* năm 2019.

2.1 Trường hợp các hàm mục tiêu không lồi

2.1.1 Thuật toán 1

Thuật toán 1

- 1: Chọn $\beta \in (0, 1)$ và $x^0 \in \mathbb{R}^n$. Đặt $k := 0$.
- 2: Tính d^k bằng cách giải bài toán con (3).
- 3: Kết thúc nếu x^k là điểm tới hạn Pareto.
- 4: Tính cỡ bước $t^k \in (0, 1]$ là giá trị lớn nhất của:

$$T^k := \left\{ t = \frac{1}{2^j} \mid j \in \mathbb{N}_0, F(x^k + td^k) \leq F(x^k) + \beta t \nabla F(x^k)^\top d^k \right\}. \quad (8)$$

- 5: Đặt $x^{k+1} := x^k + t^k d^k$, $k := k + 1$ và quay lại bước 2.
-

Bổ đề 2.1. Cỡ bước trong Thuật toán 1 luôn luôn thỏa mãn $t^k \geq t_{\min} = \min \left\{ \frac{1 - \beta}{2L_{\max}}, 1 \right\}$ với $L_{\max} = \max\{L_1, \dots, L_m\}$ (L_i là hằng số Lipschitz của ∇f_i , $i = 1, \dots, m$) và $\beta \in (0, 1)$ là tham số của điều kiện giảm đủ (8).

Chứng minh. Khi $2t$ không thỏa mãn điều kiện giảm đủ (8) trong Thuật toán 1, tồn tại chỉ số $i \in \{1, \dots, m\}$ thỏa mãn:

$$f_i(x^k + (2t)d^k) - f_i(x^k) > \beta(2t)\nabla f_i(x^k)^\top d^k. \quad (9)$$

Áp dụng **Bổ đề 1.1** ta có:

$$f_i(x^k + (2t)d^k) - f_i(x^k) \leq (2t)\nabla f_i(x^k)^\top d^k + \frac{L_i}{2}\|(2t)d^k\|^2. \quad (10)$$

Kết hợp (9) với (10) ta suy ra:

$$\begin{aligned} \beta(2t)\nabla f_i(x^k)^\top d^k &< (2t)\nabla f_i(x^k)^\top d^k + \frac{L_i}{2}\|(2t)d^k\|^2 \\ \Leftrightarrow 0 &< (2t)(1-\beta)\nabla f_i(x^k)^\top d^k + 2L_it^2\|d^k\|^2 \\ \Leftrightarrow -L_it\|d^k\|^2 &< (1-\beta)\nabla f_i(x^k)^\top d^k. \end{aligned}$$

Áp dụng (6) và (7) ta được:

$$(1-\beta)\nabla f_i(x^k)^\top d^k < -\frac{1-\beta}{2}\|d^k\|^2.$$

Do đó ta phải có:

$$\begin{aligned} -L_it\|d^k\|^2 &< -\frac{1-\beta}{2}\|d^k\|^2 \\ \Leftrightarrow t &> \frac{1-\beta}{2L_i}. \end{aligned}$$

Chú ý rằng t không bao giờ lớn hơn 1 và $L_{\max} = \max\{L_1, \dots, L_m\}$. Do đó ta có điều phải chứng minh. \square

2.1.2 Sự hội tụ của Thuật toán 1

Định lý 2.1. Giả sử rằng ít nhất một trong các hàm f_1, \dots, f_m bị chặn dưới. Gọi f_i^{\min} là cận dưới của hàm f_i khi nó bị chặn dưới. Đối với chỉ số i , gọi F^{\min} là giá trị nhỏ nhất của cận dưới f_i^{\min} và F_0^{\max} là giá trị lớn nhất của $f_i(x_0)$.

Phương pháp hướng giảm gradient được mô tả trong Thuật toán 1 sinh ra dãy $\{x^k\}$ thoả mãn:

$$\min_{0 \leq \ell \leq k-1} \|d^\ell\| \leq \sqrt{\frac{F_0^{\max} - F^{\min}}{M}} \frac{1}{\sqrt{k}},$$

trong đó $M = \frac{\beta t_{\min}}{2}$ và t_{\min} được cho trong **Bổ đề 2.1**.

Chứng minh. Gọi i là chỉ số mà hàm f_i bị chặn dưới. Từ điều kiện giảm đủ (8) ta có:

$$f_i(x^k + t^k d^k) - f_i(x^k) \leq \beta t^k \nabla f_i(x^k)^\top d^k.$$

Áp dụng (6) và (7) ta có:

$$\beta t^k \nabla f_i(x^k)^\top d^k \leq -\beta \frac{t^k}{2} \|d^k\|^2.$$

Do đó:

$$\begin{aligned} f_i(x^k + t^k d^k) - f_i(x^k) &\leq -\beta \frac{t^k}{2} \|d^k\|^2 \\ \Leftrightarrow f_i(x^k) - f_i(x^k + t^k d^k) &\geq \beta \frac{t^k}{2} \|d^k\|^2. \end{aligned}$$

Lại từ **Bổ đề 2.1** ta có:

$$\frac{\beta t^k}{2} \|d^k\|^2 \geq \frac{\beta t_{\min}}{2} \|d^k\|^2 \equiv M \|d^k\|^2.$$

Vậy tóm lại ta có:

$$f_i(x^k) - f_i(x^k + t^k d^k) \geq M \|d^k\|^2.$$

Cộng cho đến lần lặp thứ $k-1$ ta được:

$$\begin{aligned} f_i(x^0) - f_i(x^{k-1} + t^{k-1} d^{k-1}) &= \sum_{\ell=0}^{k-1} \left(f_i(x^\ell) - f_i(x^\ell + t^\ell d^\ell) \right) \\ &\geq M \sum_{\ell=0}^{k-1} \|d^\ell\|^2 \geq M.k. \left(\min_{0 \leq \ell \leq k-1} \|d^\ell\| \right)^2. \end{aligned}$$

Theo định nghĩa của F_0^{\max} và F^{\min} ta thu được:

$$\min_{0 \leq \ell \leq k-1} \|d^\ell\| \leq \sqrt{\frac{f_i(x^0) - f_i(x^{k-1} + t^{k-1} d^{k-1})}{M}} \frac{1}{\sqrt{k}} \leq \sqrt{\frac{F_0^{\max} - F^{\min}}{M}} \frac{1}{\sqrt{k}}.$$

Hoàn tất chứng minh. □

Nhận xét: Tốc độ hội tụ của Thuật toán 1 với $\min_{0 \leq \ell \leq k-1} \|d^\ell\|$ là $O\left(\frac{1}{\sqrt{k}}\right)$.

2.2 Trường hợp các hàm mục tiêu lồi hoặc lồi mạnh

2.2.1 Thuật toán 2

Thuật toán 2

- 1: Chọn $\gamma \in (0, 1)$ và $x^0 \in \mathbb{R}^n$. Đặt $k := 0$.
- 2: Tính d^k bằng cách giải bài toán con (3).
- 3: Dừng lại nếu x^k là điểm tới hạn Pareto.
- 4: Tính cỡ bước $t^k \in (0, 1]$ là giá trị lớn nhất của:

$$T^k := \left\{ t = \frac{1}{2^j} \mid j \in \mathbb{N}_0, F(x^k + td^k) \leq F(x^k) + t \nabla F(x^k)^\top d^k + \frac{\gamma t}{2} \|d^k\|^2 e \right\}, \quad (11)$$

trong đó e là vectơ một trong \mathbb{R}^m .

- 5: Đặt $x^{k+1} := x^k + t^k d^k$, $k := k + 1$ và quay lại bước 2.
-

Bổ đề 2.2. Trong Thuật toán 2, cỡ bước luôn luôn thỏa mãn $t^k \geq t_{\min} \equiv \min \left\{ \frac{\gamma}{2L_{\max}}, 1 \right\}$, với $L_{\max} = \max\{L_1, \dots, L_m\}$.

Chứng minh. Áp dụng **Bổ đề 1.1** ta có:

$$f_i(x^k + td^k) \leq f_i(x^k) + t \nabla f_i(x^k)^\top d^k + \frac{L_i}{2} \|td^k\|^2.$$

Với mọi $t \in \left(0, \frac{\gamma}{L_i}\right]$ thì:

$$\frac{L_i}{2} \|td^k\|^2 \leq \frac{\gamma}{2t} \|td^k\|^2 = \frac{\gamma t}{2} \|d^k\|^2.$$

Suy ra:

$$f_i(x^k + td^k) \leq f_i(x^k) + t \nabla f_i(x^k)^\top d^k + \frac{\gamma t}{2} \|d^k\|^2.$$

Do đó điều kiện giảm đủ (11) được thỏa mãn với $t \in \left(0, \frac{\gamma}{L_{\max}}\right]$ và thu được điều phải chứng minh vì thuật toán quay lui bắt đầu với cỡ bước là một và giảm một nửa mỗi vòng lặp. \square

2.2.2 Sự hội tụ của Thuật toán 2

Nhận xét: Bằng cách chứng minh tương tự như **Định lý 2.1**, ta sẽ chứng minh được tốc độ hội tụ của Thuật toán 2 với $\min_{0 \leq \ell \leq k-1} \|d^\ell\|$ là $\frac{1}{\sqrt{k}}$. Cụ thể như sau:

Áp dụng (11) kết hợp với (6) và (7), $\forall i = 1, \dots, m$:

$$\begin{aligned} f_i(x^k + t^k d^k) &\leq f_i(x^k) + t^k \nabla f_i(x^k)^\top d^k + \frac{\gamma t^k}{2} \|d^k\|^2 \\ &\leq f_i(x^k) - \frac{t^k}{2} \|d^k\|^2 + \frac{\gamma t^k}{2} \|d^k\|^2 \\ &= f_i(x^k) - \frac{1-\gamma}{2} t^k \|d^k\|^2. \end{aligned}$$

Do đó:

$$f_i(x^k) - f_i(x^k + t^k d^k) \geq \frac{1-\gamma}{2} t^k \|d^k\|^2.$$

Áp dụng **Bổ đề 2.2** ta có:

$$\frac{1-\gamma}{2} t^k \|d^k\|^2 \geq \frac{1-\gamma}{2} t_{\min} \|d^k\|^2 \equiv M \|d^k\|^2.$$

Vậy tóm lại ta có:

$$f_i(x^k) - f_i(x^k + t^k d^k) \geq M \|d^k\|^2.$$

Cộng cho đến lần lặp thứ $k-1$ ta được:

$$\begin{aligned} f_i(x^0) - f_i(x^{k-1} + t^{k-1} d^{k-1}) &= \sum_{\ell=0}^{k-1} (f_i(x^\ell) - f_i(x^\ell + t^\ell d^\ell)) \\ &\geq M \sum_{\ell=0}^{k-1} \|d^\ell\|^2 \geq M.k. \left(\min_{0 \leq \ell \leq k-1} \|d^\ell\| \right)^2. \end{aligned}$$

Theo định nghĩa của F_0^{\max} và F^{\min} ta thu được:

$$\min_{0 \leq \ell \leq k-1} \|d^\ell\| \leq \sqrt{\frac{f_i(x^0) - f_i(x^{k-1} + t^{k-1} d^{k-1})}{M}} \frac{1}{\sqrt{k}} \leq \sqrt{\frac{F_0^{\max} - F^{\min}}{M}} \frac{1}{\sqrt{k}}.$$

Hoàn tất chứng minh.

Nếu $L(x^0) = \{x \in \mathbb{R}^n : F(x) \leq F(x^0)\}$ bị chặn, dãy $\{x^k\}$ có giới hạn là x^* thì x^* là điểm tới hạn Pareto. Vì các hệ số λ^k nằm trong một tập bị chặn, tồn tại dãy con K sao cho x^k hội tụ tới x^* và λ^k hội tụ tới λ^* thỏa mãn:

$$\sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0, \quad \sum_{i=1}^m \lambda_i^* = 1, \quad \lambda_i^* \geq 0, \quad i = 1, \dots, m \quad (12)$$

Trong việc suy ra tốc độ hội tụ toàn cục cho trường hợp lồi, ta giả sử mạnh hơn rằng (x^k, λ^k) hội tụ tới (x^*, λ^*) . Với giả thiết các hàm mục tiêu f_i lồi, điểm x^* là điểm Pareto yếu, và nếu thêm thêm điều kiện x^* là cực tiểu duy nhất của hàm vô hướng $\sum_i \lambda_i^* f_i$ thì x^* là điểm Pareto [6].

Bổ đề 2.3. Giả sử rằng $\{x^k\}$ hội tụ đến x^* . Nếu f_i là hàm lồi hoặc lồi mạnh với hệ số lồi mạnh $\mu_i > 0 \forall i = 1, \dots, m$, khi đó:

$$\sum_{i=1}^m \lambda_i^k (f_i(x^{k+1}) - f_i(x^*)) \leq \frac{1}{2t_{\min}} (\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2) - \frac{\mu}{2} \|x^k - x^*\|^2, \quad (13)$$

trong đó $\mu = \min_{1 \leq i \leq m} \mu_i$ và t_{\min} được cho trong **Bổ đề 2.2**.

Chứng minh. Áp dụng **Bổ đề 1.1** kết hợp điều kiện $t^k \in \left(0, \frac{\gamma}{L_i}\right]$, với mọi $i = 1, \dots, m$ ta có:

$$\begin{aligned} f_i(x^{k+1}) &\equiv f_i(x^k + t^k d^k) \leq f_i(x^k) + \nabla f_i(x^k)^\top (t^k d^k) + \frac{L_i}{2} \|t^k d^k\|^2 \\ &\leq f_i(x^k) + \nabla f_i(x^k)^\top (t^k d^k) + \frac{\gamma}{2t^k} \|t^k d^k\|^2. \end{aligned}$$

Vì $\gamma < 1$ nên suy ra:

$$f_i(x^{k+1}) \leq f_i(x^k) + \nabla f_i(x^k)^\top (t^k d^k) + \frac{1}{2t^k} \|t^k d^k\|^2. \quad (14)$$

Sử dụng tính lồi ($\mu_i = 0$) / lồi mạnh ($\mu_i > 0$) của f_i ta được:

$$\begin{aligned} f_i(x^k) &\leq f_i(x^*) + \nabla f_i(x^k)^\top (x^k - x^*) - \frac{\mu_i}{2} \|x^k - x^*\|^2 \\ &\leq f_i(x^*) + \nabla f_i(x^k)^\top (x^k - x^*) - \frac{\mu}{2} \|x^k - x^*\|^2. \end{aligned} \quad (15)$$

Kết hợp (14) và (15) ta được:

$$\begin{aligned} f_i(x^{k+1}) &\leq f_i(x^*) + \nabla f_i(x^k)^\top (x^k - x^*) - \frac{\mu}{2} \|x^k - x^*\|^2 \\ &\quad + \nabla f_i(x^k)^\top (t^k d^k) + \frac{1}{2t^k} \|t^k d^k\|^2 \quad \forall i = 1, \dots, m. \end{aligned}$$

Tương đương với:

$$f_i(x^{k+1}) - f_i(x^*) \leq \nabla f_i(x^k)^\top (x^k - x^* + t^k d^k) + \frac{t^k}{2} \|d^k\|^2 - \frac{\mu}{2} \|x^k - x^*\|^2 \quad \forall i = 1, \dots, m.$$

Nhân với λ_i^k và cộng theo vế với $i = 1, \dots, m$ ta được:

$$\sum_{i=1}^m \lambda_i^k \left(f_i(x^{k+1}) - f_i(x^*) \right) \leq \left(\sum_{i=1}^m \lambda_i^k \nabla f_i(x^k) \right)^\top (x^k - x^* + t^k d^k) + \left(\frac{t^k}{2} \|d^k\|^2 - \frac{\mu}{2} \|x^k - x^*\|^2 \right) \sum_{i=1}^m \lambda_i^k.$$

Áp dụng (4) và (5) ta thu được:

$$\begin{aligned}
\sum_{i=1}^m \lambda_i^k \left(f_i(x^{k+1}) - f_i(x^*) \right) &\leq -(d^k)^\top (x^k - x^* + t^k d^k) + \frac{t^k}{2} \|d^k\|^2 - \frac{\mu}{2} \|x^k - x^*\|^2 \\
&= -\frac{1}{2t^k} \left(2(t^k d^k)^\top (x^k - x^*) + \|t^k d^k\|^2 \right) - \frac{\mu}{2} \|x^k - x^*\|^2 \\
&= -\frac{1}{2t^k} \left(\|x^k - x^* + t^k d^k\|^2 - \|x^k - x^*\|^2 \right) - \frac{\mu}{2} \|x^k - x^*\|^2 \\
&= \frac{1}{2t^k} \left(\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2 \right) - \frac{\mu}{2} \|x^k - x^*\|^2 \\
&\leq \frac{1}{2t_{\min}} \left(\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2 \right) - \frac{\mu}{2} \|x^k - x^*\|^2.
\end{aligned}$$

Hoàn tất chứng minh. □

Nhận xét: Do $\lambda_i^k [f_i(x^{k+1}) - f_i(x^*)] \geq 0$ dẫn đến $\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2 \geq 0$.

Nhận xét: Nếu sử dụng một cỡ bước cố định $t^k = \bar{t}$ không đổi (thay vì dùng điều kiện giảm đủ) thỏa mãn $0 < \bar{t} \leq \frac{1}{L_{\max}}$ thì kết quả của **Bổ đề 2.3** vẫn đúng mà không cần giả sử $\{x^k\}$ hội tụ tới x^* . Trong trường hợp này, chúng ta sẽ không biết liệu vế trái của (13) có không âm hay không, và do đó không chắc chắn liệu kết quả đó sau này có thể được sử dụng để chứng minh tốc độ hiệu quả hay không.

Chúng ta tách trường hợp lồi ra khỏi trường hợp lồi mạnh. Tiếp theo ta giải quyết trường hợp lồi bằng cách thiết lập tốc độ mong muốn $O\left(\frac{1}{k}\right)$ cho chuỗi các trọng số $\{\bar{\lambda}^k\}$ hội tụ đến trọng số tối ưu λ^* trong (12), khi các hệ số $\{\lambda^k\}$ cũng hội tụ đến λ^* .

Định lý 2.2. Cho các hàm f_1, \dots, f_m là lồi. Giả sử dãy $\{x^k\}$ hội tụ đến x^* .

Phương pháp gradient được mô tả trong Thuật toán 2 sinh ra dãy $\{x^k\}$ thỏa mãn:

$$\sum_{i=1}^m \bar{\lambda}_i^{k-1} f_i(x^k) - \sum_{i=1}^m \bar{\lambda}_i^{k-1} f_i(x^*) \leq \frac{\|x^0 - x^*\|^2}{2t_{\min} k},$$

trong đó trọng số $\bar{\lambda}_i^{k-1} = \frac{1}{k} \sum_{\ell=0}^{k-1} \lambda_i^\ell$ thỏa mãn:

$$\sum_{i=1}^m \bar{\lambda}_i^{k-1} = 1 \quad \text{và} \quad \bar{\lambda}_i^{k-1} \geq 0 \quad \forall i = 1, \dots, m.$$

Nếu $\{\lambda^k\}$ hội tụ đến λ^* thì $\{\bar{\lambda}^k\}$ cũng hội tụ đến λ^* .

Chứng minh. Từ (13) ta suy ra:

$$\sum_{i=1}^m \lambda_i^\ell \left(f_i(x^{\ell+1}) - f_i(x^*) \right) \leq \frac{1}{2t_{\min}} \left(\|x^\ell - x^*\|^2 - \|x^{\ell+1} - x^*\|^2 \right). \quad (16)$$

Lấy tổng của (16) từ $\ell = 0, \dots, k-1$ ta được:

$$\sum_{\ell=0}^{k-1} \sum_{i=1}^m \lambda_i^\ell \left(f_i(x^{\ell+1}) - f_i(x^*) \right) \leq \frac{1}{2t_{\min}} \left(\|x^0 - x^*\|^2 - \|x^k - x^*\|^2 \right) \leq \frac{1}{2t_{\min}} \|x^0 - x^*\|^2.$$

Do $f_i(x^\ell) \leq f_i(x^{\ell-1})$ với mọi i, ℓ , suy ra:

$$\sum_{\ell=0}^{k-1} \sum_{i=1}^m \lambda_i^\ell \left(f_i(x^{\ell+1}) - f_i(x^*) \right) \geq \sum_{\ell=0}^{k-1} \sum_{i=1}^m \lambda_i^\ell \left(f_i(x^k) - f_i(x^*) \right).$$

Vì vậy:

$$\sum_{i=1}^m \left(\sum_{\ell=0}^{k-1} \lambda_i^\ell \right) \left(f_i(x^k) - f_i(x^*) \right) \leq \frac{1}{2t_{\min}} \|x^0 - x^*\|^2.$$

Chia hai vế cho k ta thu được:

$$\sum_{i=1}^m \bar{\lambda}_i^{k-1} \left(f_i(x^k) - f_i(x^*) \right) \leq \frac{1}{2t_{\min}k} \|x^0 - x^*\|^2.$$

Hoàn tất chứng minh. □

Nhận xét: Tốc độ hội tụ của Thuật toán 2 trong trường hợp các hàm mục tiêu lồi với

$$\sum_{i=1}^m \bar{\lambda}_i^{k-1} \left(f_i(x^k) - f_i(x^*) \right) \text{ là } O\left(\frac{1}{k}\right).$$

Bây giờ chúng ta chứng minh rằng phương pháp hướng giảm gradient cũng đạt được tốc độ hội tụ tuyến tính cho bài toán tối ưu đa mục tiêu trong trường hợp lồi mạnh.

Định lý 2.3. Cho các hàm f_i là lồi mạnh với hệ số lồi mạnh $\mu_i > 0$, $i = 1, \dots, m$. Giả sử $\{x^k\}$ hội tụ đến x^* .

Phương pháp gradient được mô tả trong Thuật toán 2 sinh ra dãy $\{x^k\}$ thỏa mãn:

$$\|x^k - x^*\| \leq \left(\sqrt{1 - t_{\min} \mu} \right)^k \|x^0 - x^*\|. \quad (17)$$

Chứng minh. Viết lại (13) thành:

$$\sum_{i=1}^m \lambda_i^k \left(f_i(x^{k+1}) - f_i(x^*) \right) \leq \frac{1}{2t_{\min}} \left((1 - t_{\min} \mu) \|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2 \right).$$

Do về trái bất đẳng thức là không âm nên phải có:

$$\|x^{k+1} - x^*\|^2 \leq (1 - t_{\min}\mu)\|x^k - x^*\|^2.$$

Áp dụng liên tiếp bất đẳng thức trên ta thu được điều phải chứng minh. \square

Nếu cặp (x^*, λ^*) là tối hạn Pareto và các hàm f_i là lồi thì $f^* \equiv \sum_{i=1}^m \lambda_i^* f_i$ là hàm lồi với điểm cực tiểu x^* . Khi tất cả các hàm f_i đều lồi mạnh thì f^* lồi mạnh với hệ số lồi mạnh $\mu^* = \min_{\lambda_i^* > 0} \mu_i$. Hơn nữa hàm f^* có gradient liên tục Lipschitz với hằng số $L^* = \max_{\lambda_i^* > 0} L_i$. Theo [1] thì:

$$f^*(x^k) - f^*(x^*) \leq \frac{L^*}{2} \|x^k - x^*\|^2.$$

và từ (17) ta cũng suy ra tốc độ hội tụ tuyến tính cho độ chênh lệch tối ưu trong f^* :

$$\sum_{i=1}^m \lambda_i^* f_i(x^k) - \sum_{i=1}^m \lambda_i^* f_i(x^*) \leq \frac{L^*}{2} \|x^k - x^*\|^2 \leq \frac{L^*}{2} (1 - t_{\min}\mu)^k \|x^0 - x^*\|^2.$$

Nhận xét: Tốc độ hội tụ của Thuật toán 2 trong trường hợp các hàm mục tiêu lồi mạnh với $\sum_{i=1}^m \lambda_i^* (f_i(x^k) - f_i(x^*))$ là $O(r^k)$ với $r \in (0, 1)$.

Kết luận: Chương 2 đã trình bày hai thuật toán và các kết quả liên quan. Tiếp theo, Chương 3 của đề án sẽ tiến hành chạy thử nghiệm các thuật toán trên các ví dụ điển hình.

Lập trình thử nghiệm

Chương này ta sẽ chạy thử nghiệm hai thuật toán trên một số ví dụ điển hình. Các ví dụ được lấy từ bài báo "A Barzilai-Borwein descent method for multiobjective optimization problems" [3] của các tác giả Jian Chen, Liping Tang và Xinmin Yang, được đăng trên tạp chí *European Journal of Operational Research* năm 2023. Tác giả sẽ đánh giá kết quả dựa trên các tiêu chí sau: Số vòng lặp trung bình (iter), thời gian trung bình (time (ms)) và cỡ bước trung bình (stepsize).

Ở chương trước ta đã tìm hiểu hai thuật toán và cả hai thuật toán này đều có chung một phần là giải bài toán con (3), tuy nhiên việc giải trực tiếp bài toán con này rất phức tạp do nó có m ràng buộc bất đẳng thức tuyến tính. Do vậy mà để đơn giản hơn ta sẽ giải bài toán đối ngẫu của nó được phát biểu như sau:

$$\max_{\lambda \geq 0} -\frac{1}{2} \left\| \sum_{i=1}^m \lambda_i \nabla f_i(x^k) \right\|_2^2,$$

$$\text{với } \lambda_i \geq 0,$$

$$\sum_{i=1}^m \lambda_i = 1.$$

Chương trình của tác giả được viết bằng ngôn ngữ Python. Các tham số như sau:

- Số lần lặp tối đa: 500.
- Sai số dừng: $\varepsilon = 10^{-8}$.
- Hệ số $\beta = 10^{-4}$.
- Hệ số $\gamma = 0.99$.

Bảng sau đây mô tả các ví dụ dựa trên số lượng biến (n), số lượng hàm mục tiêu (m), cận dưới (x_L), cận trên (x_U), và số điểm trên mỗi trục (num_points).

Ví dụ	n	m	x_L	x_U	num_points
Imbalance1	2	2	[-2,-2]	[2,2]	100
Imbalance2	2	2	[-2,-2]	[2,2]	100
JOS1	3	2	[-2,-2,-2]	[2,2,2]	100
WIT1	2	2	[-2,-2]	[2,2]	100
WIT2	2	2	[-2,-2]	[2,2]	100
WIT3	2	2	[-2,-2]	[2,2]	100
WIT4	2	2	[-2,-2]	[2,2]	100
WIT5	2	2	[-2,-2]	[2,2]	100
WIT6	2	2	[-2,-2]	[2,2]	100
Deb	2	2	[0.1,0.1]	[1,1]	100
PNR	2	2	[-2,-2]	[2,2]	100
DD1	5	2	-20[1,...,1]	20[1,...,1]	10
FDS	3	3	[-2,-2,-2]	[2,2,2]	20
TRIDIA1	3	3	[-1,-1,-1]	[1,1,1]	50

Bảng 3.1: Mô tả các ví dụ được sử dụng để thử nghiệm

Ví dụ 3.0.1. (*Imbalance1*)

$$\min_{\mathbf{x} \in \mathbb{R}^2} (f_1(x), f_2(x)),$$

trong đó

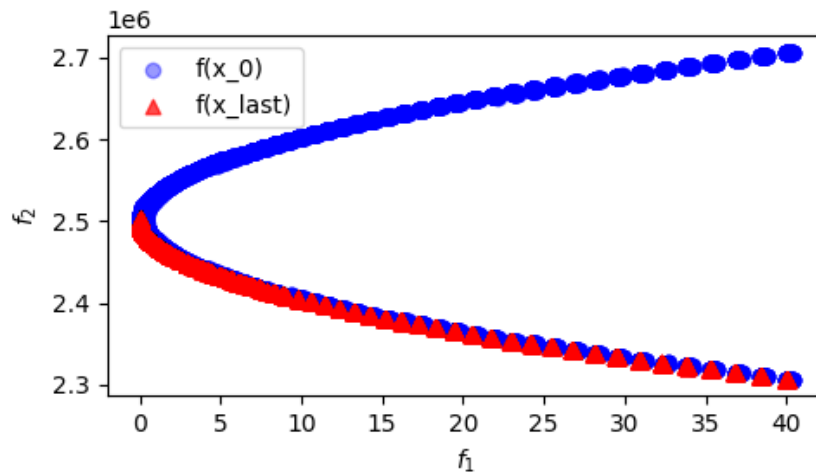
$$f_1(x) = 0.1x_1^2 + 10x_2^2,$$

$$f_2(x) = (x_1 - 50)^2 + 1000(x_2 + 50)^2,$$

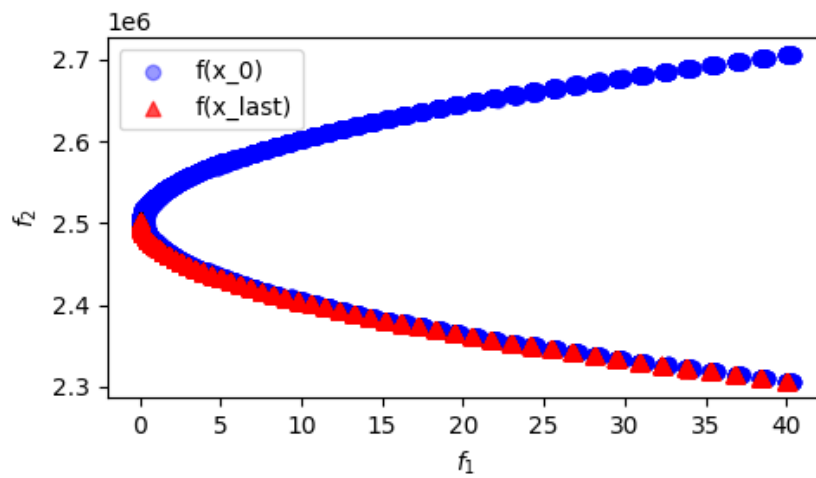
$$\text{Với } -2 \leq x_i \leq 2, \quad i = 1, 2.$$

Ở ví dụ này, cả hai hàm mục tiêu đều là hàm lồi cho nên có thể áp dụng cả hai thuật toán.

Kết quả như sau:



(a) Thuật toán 1



(b) Thuật toán 2

Hình 3.1: Kết quả chạy Thuật toán 1 và Thuật toán 2 trong không gian ảnh của Ví dụ 3.0.1

Ví dụ 3.0.2. (*Imbalance2*)

$$\min_{\mathbf{x} \in \mathbb{R}^2} (f_1(x), f_2(x)),$$

trong đó

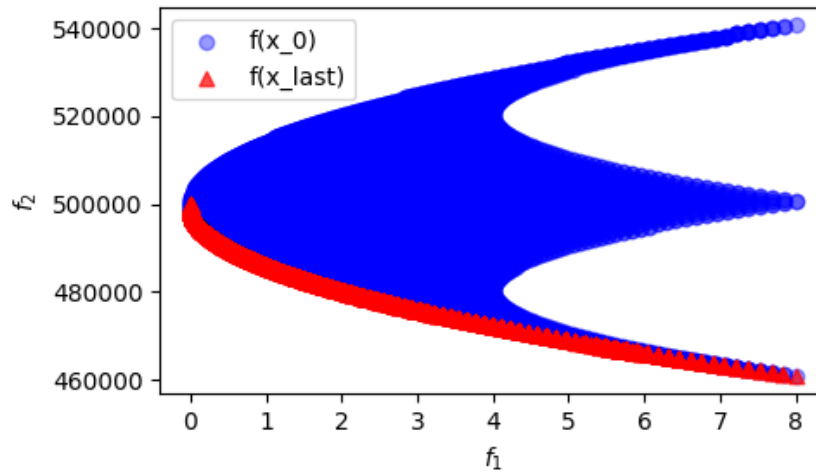
$$f_1(x) = x_1^2 + x_2^2,$$

$$f_2(x) = 100(x_1 - 50)^2 + 100(x_2 + 50)^2,$$

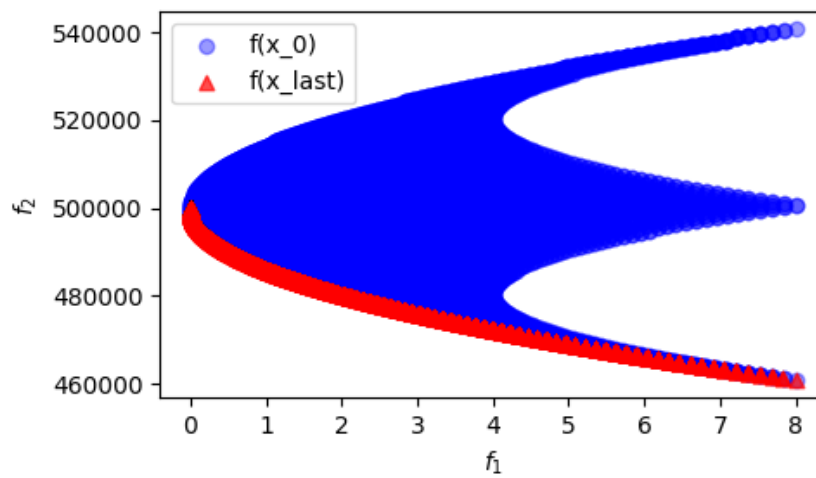
$$\text{Với } -2 \leq x_i \leq 2, \quad i = 1, 2.$$

Ở ví dụ này, cả hai hàm mục tiêu đều là hàm lồi cho nên có thể áp dụng cả hai thuật toán.

Kết quả như sau:



(a) Thuật toán 1



(b) Thuật toán 2

Hình 3.2: Kết quả chạy Thuật toán 1 và Thuật toán 2 trong không gian ảnh của Ví dụ 3.0.2

Ví dụ 3.0.3. (JOS1)

$$\min_{\mathbf{x} \in \mathbb{R}^3} (f_1(x), f_2(x)),$$

trong đó

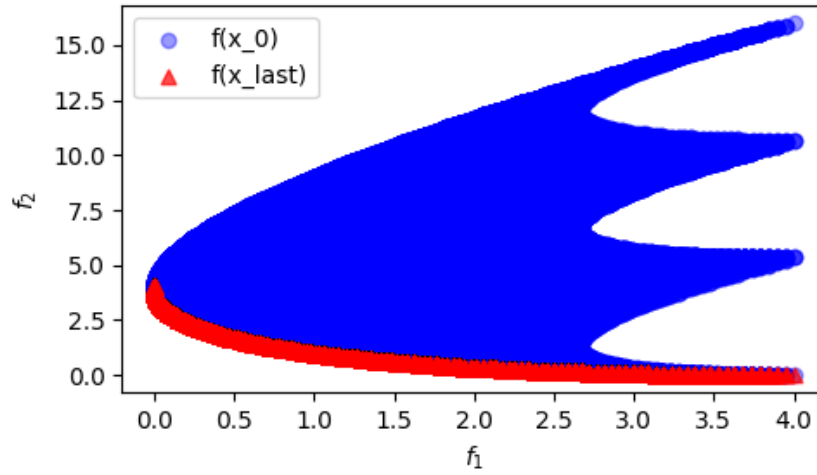
$$f_1(x) = \frac{1}{3} \sum_{i=1}^3 x_i^2,$$

$$f_2(x) = \frac{1}{3} \sum_{i=1}^3 (x_i - 2)^2,$$

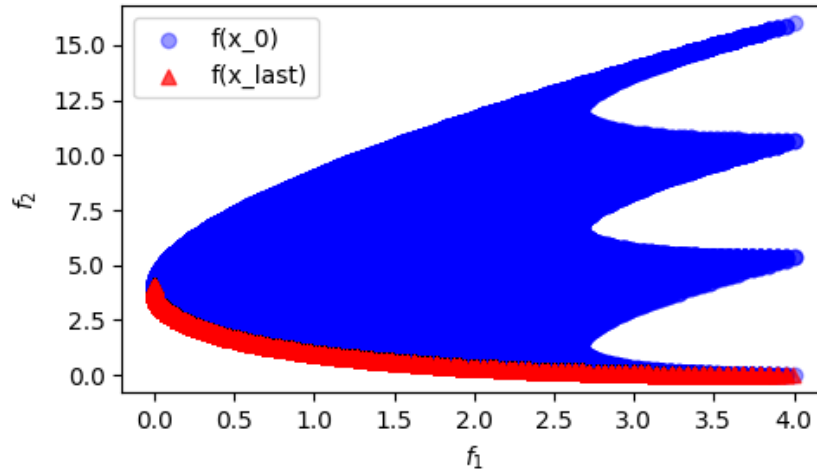
$$\text{Với } -2 \leq x_i \leq 2, \quad i = 1, 2, 3.$$

Ở ví dụ này, cả hai hàm mục tiêu đều là hàm lồi cho nên có thể áp dụng cả hai thuật toán.

Kết quả như sau:



(a) Thuật toán 1



(b) Thuật toán 2

Hình 3.3: Kết quả chạy Thuật toán 1 và Thuật toán 2 trong không gian ảnh của Ví dụ 3.0.3

Ví dụ 3.0.4. (WIT)

$$\min_{\mathbf{x} \in \mathbb{R}^2} (f_1(x, \lambda), f_2(x, \lambda)),$$

trong đó

$$f_1(x, \lambda) = \lambda ((x_1 - 2)^2 + (x_2 - 2)^2) + (1 - \lambda) ((x_1 - 2)^4 + (x_2 - 2)^8),$$

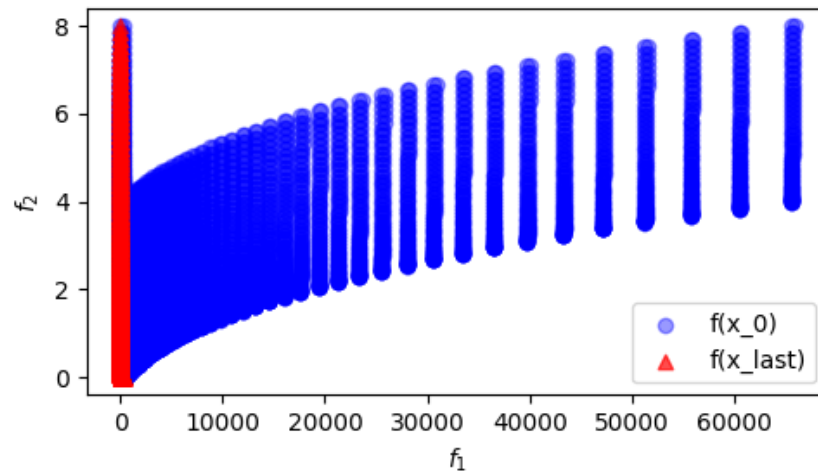
$$f_2(x, \lambda) = (x_1 + 2\lambda)^2 + (x_2 + 2\lambda)^2,$$

$$\text{Với } \lambda = 0, 0.5, 0.9, 0.99, 0.999, 1 \text{ và } -2 \leq x_i \leq 2, \quad i = 1, 2.$$

Ở ví dụ này, hàm mục tiêu $f_1(x, \lambda)$ không lồi cho nên chỉ có thể áp dụng Thuật toán 1.

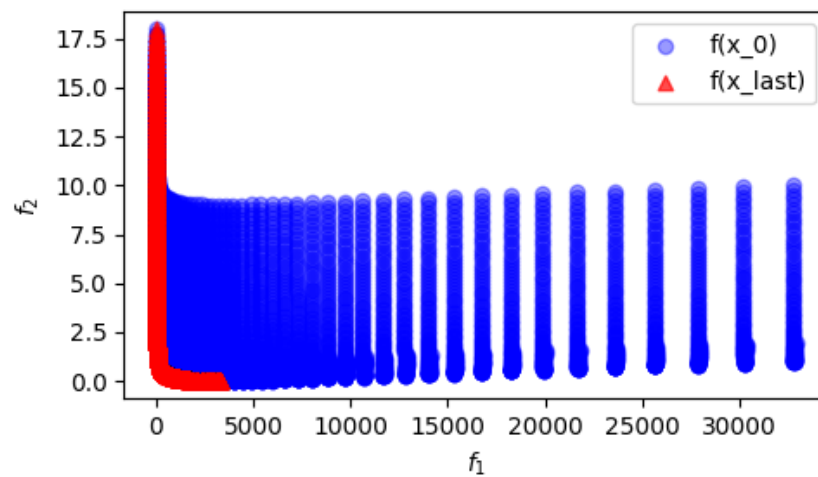
Kết quả như sau:

Với $\alpha = 0$



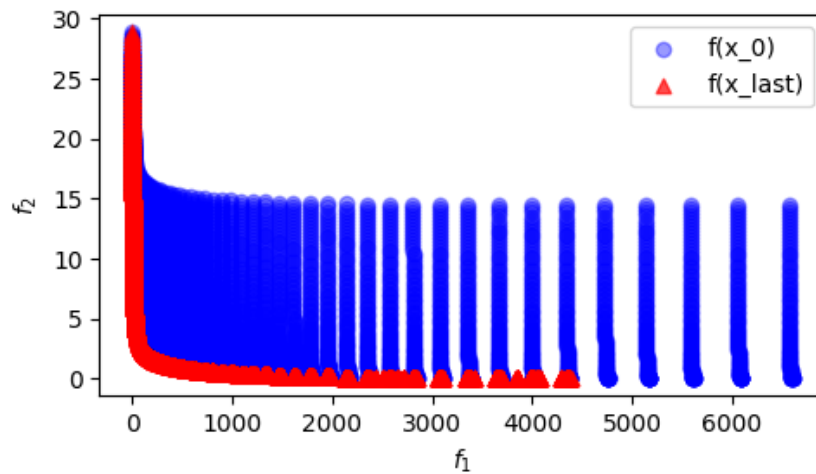
Hình 3.4: Kết quả chạy Thuật toán 1 trong không gian ảnh với trường hợp $\alpha = 0$ của Ví dụ 3.0.4

Với $\alpha = 0.5$



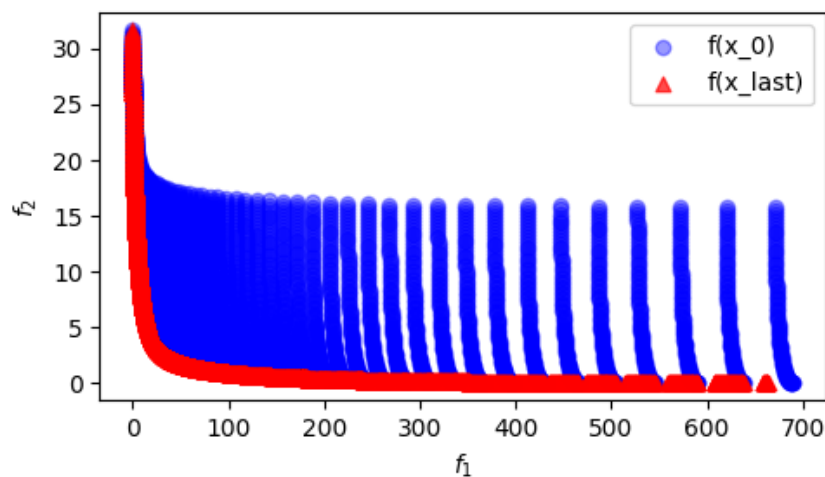
Hình 3.5: Kết quả chạy Thuật toán 1 trong không gian ảnh với trường hợp $\alpha = 0.5$ của Ví dụ 3.0.4

Với $\alpha = 0.9$



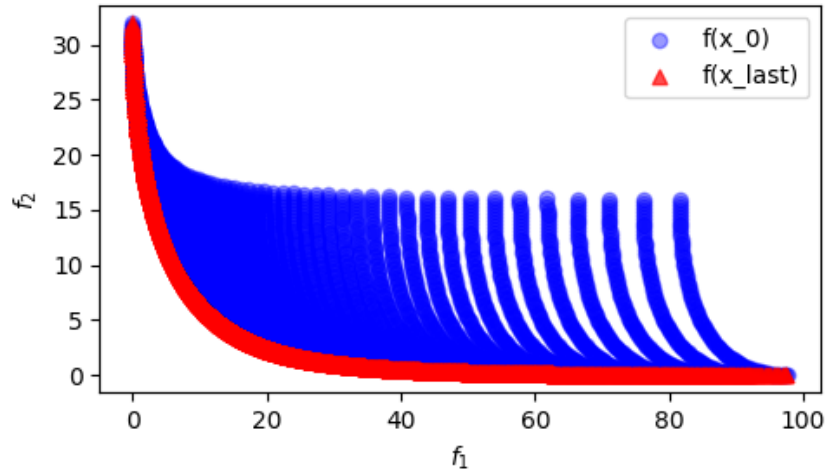
Hình 3.6: Kết quả chạy Thuật toán 1 trong không gian ảnh với trường hợp $\alpha = 0.9$ của Ví dụ 3.0.4

Với $\alpha = 0.99$



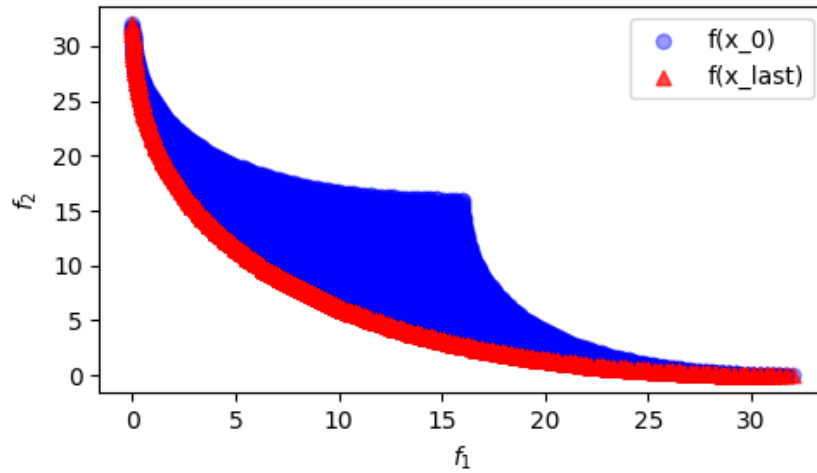
Hình 3.7: Kết quả chạy Thuật toán 1 trong không gian ảnh với trường hợp $\alpha = 0.99$ của Ví dụ 3.0.4

Với $\alpha = 0.999$



Hình 3.8: Kết quả chạy Thuật toán 1 trong không gian ảnh với trường hợp $\alpha = 0.999$ của Ví dụ 3.0.4

Với $\alpha = 1$



Hình 3.9: Kết quả chạy Thuật toán 1 trong không gian ảnh với trường hợp $\alpha = 1$ của Ví dụ 3.0.4

Ví dụ 3.0.5. (Deb)

$$\min_{x_1 > 0} \left(x_1, \frac{g(x_2)}{x_1} \right),$$

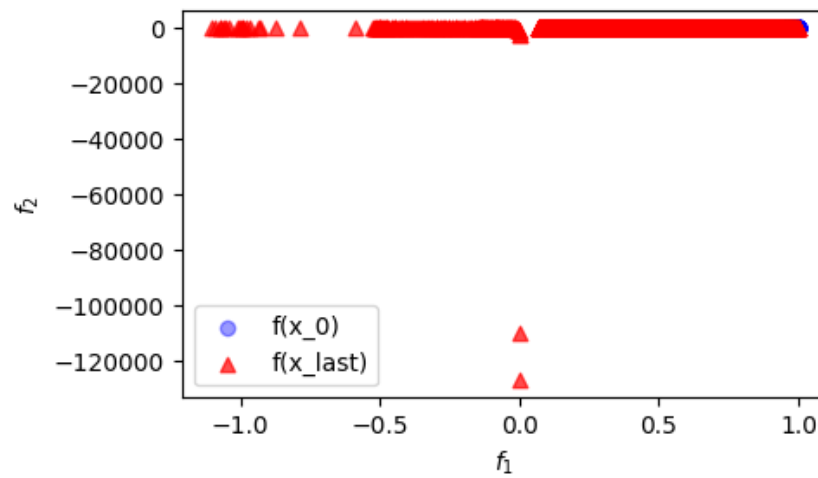
trong đó

$$g(x_2) = 2 - \exp \left\{ - \left(\frac{x_2 - 0.2}{0.004} \right)^2 \right\} - 0.8 \exp \left\{ - \left(\frac{x_2 - 0.6}{0.4} \right)^2 \right\},$$

$$\text{Với } 0.1 \leq x_i \leq 1, \quad i = 1, 2.$$

Ở ví dụ này, hàm mục tiêu $\frac{g(x_2)}{x_1}$ không lồi cho nên chỉ có thể áp dụng Thuật toán 1.

Kết quả như sau:



Hình 3.10: Kết quả chạy Thuật toán 1 trong không gian ảnh của Ví dụ 3.0.5

Ví dụ 3.0.6. (PNR)

$$\min_{\mathbf{x} \in \mathbb{R}^2} (f_1(x), f_2(x)),$$

trong đó

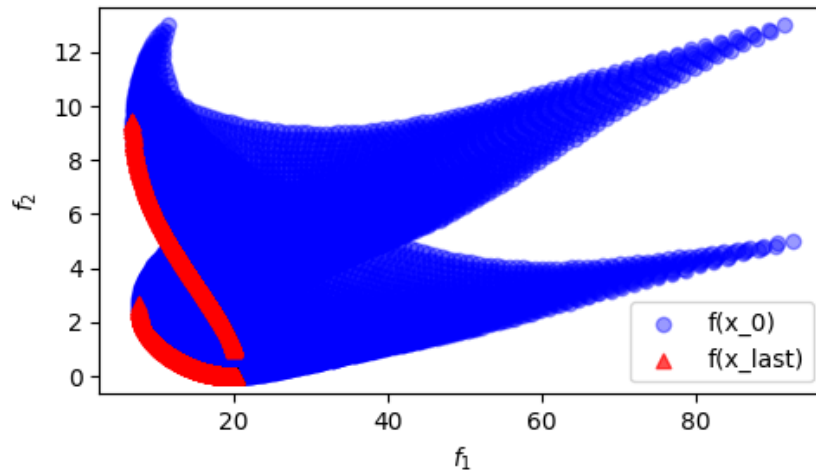
$$f_1(x) = x_1^4 + x_2^4 - x_1^2 + x_2^2 - 10x_1x_2 + 0.25x_1 + 20,$$

$$f_2(x) = (x_1 - 1)^2 + x_2^2,$$

$$\text{Với } -2 \leq x_i \leq 2, \quad i = 1, 2.$$

Ở ví dụ này, hàm mục tiêu $f_1(x)$ không lồi cho nên chỉ có thể áp dụng Thuật toán 1.

Kết quả như sau:



Hình 3.11: Kết quả chạy Thuật toán 1 trong không gian ảnh của Ví dụ 3.0.6

Ví dụ 3.0.7. (DD1)

$$\min_{x \in \mathbb{R}^5} (f_1(x), f_2(x)),$$

trong đó

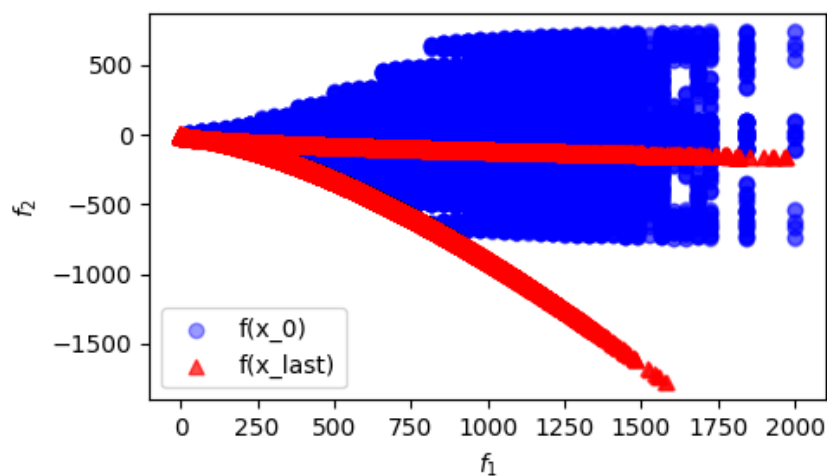
$$f_1(x) = \sum_{i=1}^5 x_i^2,$$

$$f_2(x) = 3x_1 + 2x_2 - \frac{x_3}{3} + 0.01(x_4 - x_5)^3,$$

$$\text{Với } -20 \leq x_i \leq 20, \quad i = 1, \dots, 5.$$

Ở ví dụ này, hàm mục tiêu $f_2(x)$ không lồi cho nên chỉ có thể áp dụng Thuật toán 1.

Kết quả như sau:



Hình 3.12: Kết quả chạy Thuật toán 1 trong không gian ảnh của Ví dụ 3.0.7

Ví dụ 3.0.8. (FDS)

$$\min_{\mathbf{x} \in \mathbb{R}^3} (f_1(x), f_2(x), f_3(x)),$$

trong đó

$$f_1(x) = \frac{1}{n} \sum_{i=1}^n i(x_i - i)^2,$$

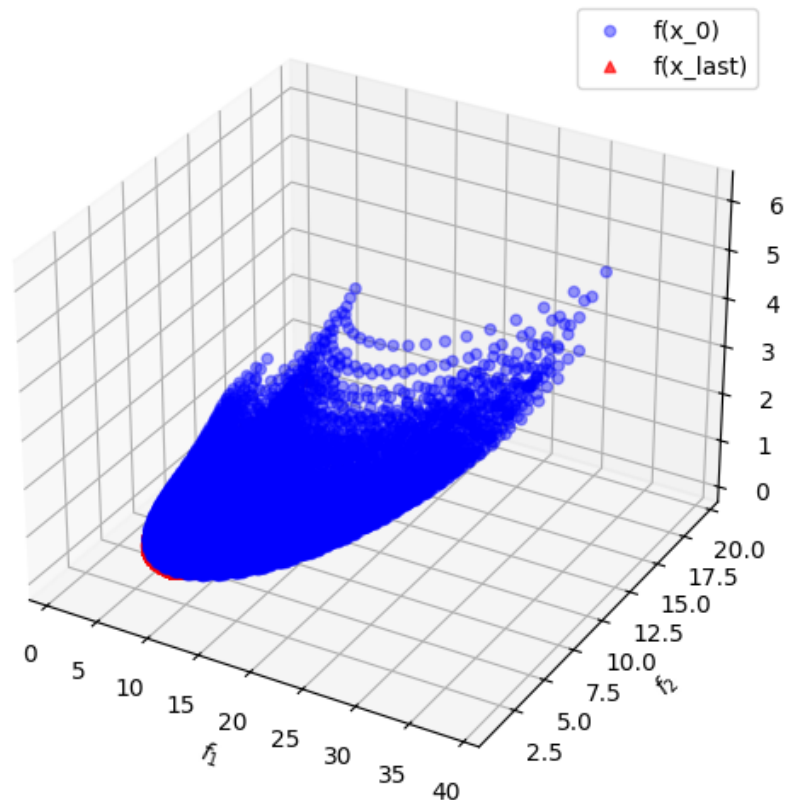
$$f_2(x) = \exp\left(\sum_{i=1}^n \frac{x_i}{n}\right) + \|x\|_2^2,$$

$$f_3(x) = \frac{1}{n(n+1)} \sum_{i=1}^n i(n-i+1) e^{-x_i},$$

Với $n = 3$ và $-2 \leq x_i \leq 2$, $i = 1, 2, 3$.

Ở ví dụ này, hàm mục tiêu $f_2(x)$ và $f_3(x)$ không lồi cho nên chỉ có thể áp dụng Thuật toán 1.

Kết quả như sau:



Hình 3.13: Kết quả chạy Thuật toán 1 trong không gian ảnh của Ví dụ 3.0.8

Ví dụ 3.0.9. (TRIDIAI)

$$\min_{x \in \mathbb{R}^3} (f_1(x), f_2(x), f_3(x)),$$

trong đó

$$f_1(x) = (2x_1 - 1)^2,$$

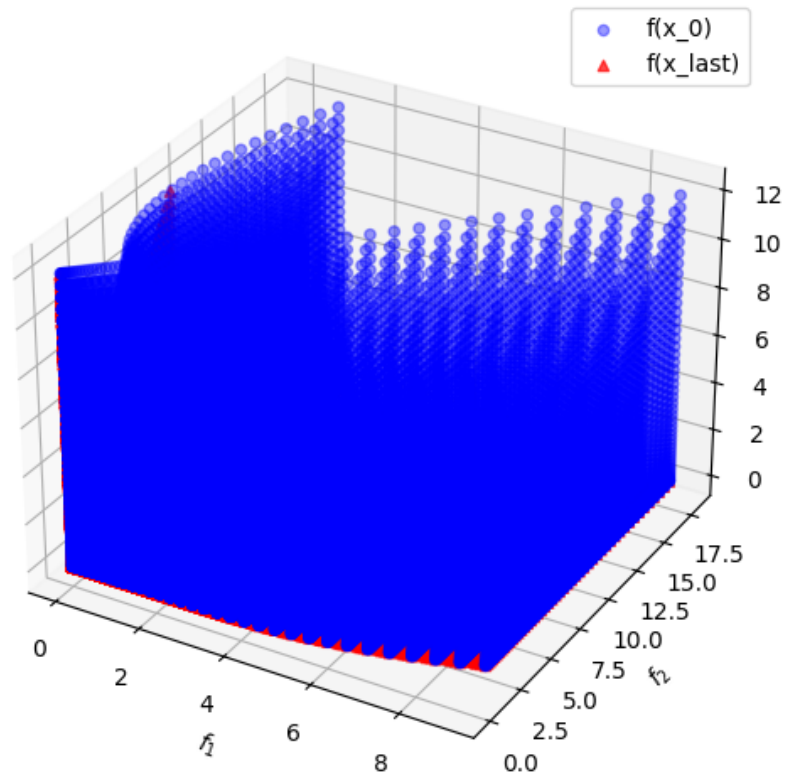
$$f_2(x) = 2(2x_1 - x_2)^2,$$

$$f_3(x) = 3(x_2 - x_3)^2,$$

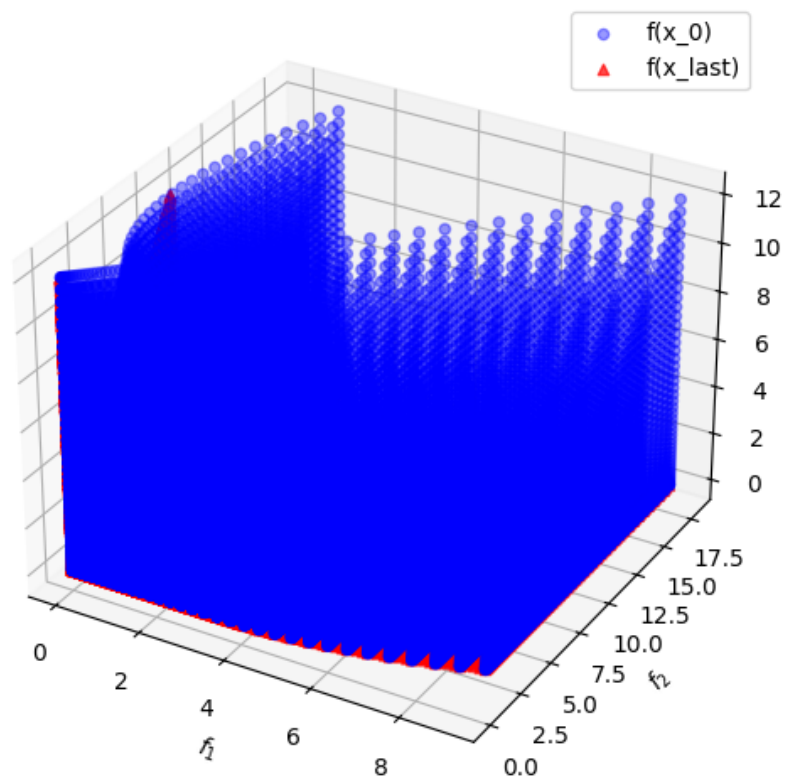
$$\text{Với } -1 \leq x_i \leq 1, \quad i = 1, 2, 3.$$

Ở ví dụ này, cả hai hàm mục tiêu đều là hàm lồi cho nên có thể áp dụng cả hai thuật toán.

Kết quả như sau:



(a) Thuật toán 1



(b) Thuật toán 2

Hình 3.14: Kết quả chạy Thuật toán 1 và Thuật toán 2 trong không gian ảnh của Ví dụ 3.0.9

Kết quả số của các ví dụ trên được tổng hợp trong bảng sau:

Problem	Thuật toán 1			Thuật toán 2		
	iter	time (ms)	stepsize	iter	time (ms)	stepsize
Imbalance1	57.25	12.96	0.50	129.13	26.05	0.40
Imbalance2	487.95	124.56	0.0082	495.12	172.37	0.0039
JOS1	8.49	1.19	0.9999	8.48	1.25	1.00
WIT1	131.02	34.08	0.06	Không lỗi		
WIT2	129.68	34.07	0.06	-----		
WIT3	41.08	9.54	0.12	-----		
WIT4	8.76	1.90	0.28	-----		
WIT5	6.31	1.27	0.40	-----		
WIT6	0.99	0.24	0.50	-----		
Deb	96.34	37.09	0.09	-----		
PNR	10.51	2.34	0.18	-----		
DD1	79.37	12.96	0.53	-----		
FDS	97.11	369.41	0.45	-----		
TRIDIA1	9.03	13.00	0.47	15.15	15.31	0.0916

Bảng 3.2: Tổng hợp kết quả chạy thuật toán cho các ví dụ

Nhận xét: Về lí thuyết thì Thuật toán 2 được chứng minh là hội tụ nhanh hơn so với Thuật toán 1 nhưng thử nghiệm 4 ví dụ ta thấy điều ngược lại. Nguyên nhân là do số lượng hàm mục tiêu chưa đủ nhiều. Ngoài ra khi áp dụng Thuật toán 2 thu được cỡ bước thường nhỏ hơn khi áp dụng Thuật toán 1.

Kết luận: Chương cuối của đồ án đã trình bày kết quả thử nghiệm các thuật toán trên một số ví dụ. Kết quả là hai thuật toán đều có độ chính xác cao.

Kết luận

Kết quả đồ án:

Đồ án "Phương pháp hướng giảm gradient cho bài toán tối ưu đa mục tiêu không ràng buộc: lý thuyết và thử nghiệm" đã trình bày phương pháp hướng giảm gradient cho bài toán tối ưu đa mục tiêu không ràng buộc trong các trường hợp không lồi, lồi và lồi mạnh. Sau đó tiến hành tính toán thử nghiệm các thuật toán với một số bài toán cụ thể. Các kết quả đạt được của đồ án được tổng hợp như sau:

- Trình bày nội dung lý thuyết về bài toán tối ưu đa mục tiêu không ràng buộc tổng quát.
- Trình bày tổng quan về nghiệm tối ưu Pareto, tối ưu Pareto yếu, điểm tối hạn Pareto.
- Nêu thuật toán hướng giảm gradient cho các trường hợp không lồi, lồi và lồi mạnh đi kèm chứng minh sự hội tụ chi tiết.
- Tính toán thử nghiệm minh họa thuật toán trên một số ví dụ cụ thể.

Hướng phát triển của đồ án:

Trong thời gian tới, tác giả sẽ tiếp tục nghiên cứu đề tài để có thể thực hiện các ý tưởng sau:

- Tìm hiểu thêm các phương pháp khác để giải bài toán tối ưu đa mục tiêu không ràng buộc.
- Triển khai phương pháp thành phương pháp hướng giảm gradient với cỡ bước thích nghi.

Tài liệu tham khảo

- [1] Amir Beck. *First-order methods in optimization*. SIAM, 2017.
- [2] Amir Beck. *Introduction to nonlinear optimization: Theory, algorithms, and applications with MATLAB*. SIAM, 2014.
- [3] Jian Chen, Liping Tang, and Xinmin Yang. “A Barzilai-Borwein descent method for multiobjective optimization problems”. In: *European Journal of Operational Research* 311.1 (2023), pp. 196–209.
- [4] Jörg Fliege and Benar Fux Svaiter. “Steepest descent methods for multicriteria optimization”. In: *Mathematical methods of operations research* 51 (2000), pp. 479–494.
- [5] Jörg Fliege, A Ismael F Vaz, and Luís Nunes Vicente. “Complexity of gradient descent for multiobjective optimization”. In: *Optimization Methods and Software* 34.5 (2019), pp. 949–959.
- [6] Johannes Jahn et al. *Vector optimization*. Springer, 2009.
- [7] Kaisa Miettinen. *Nonlinear multiobjective optimization*. Vol. 12. Springer Science & Business Media, 1999.