ĐỒ ÁN 2

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

GVHD:	TS. Phạm Thị Hoài
Sinh viên thực hiện:	Hoàng Kim Khánh
MSSV:	20216839

Hà Nội, ngày 16 tháng 1 năm 2025

Nội dung

Nội dung

Giới thiệu

- 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.
- 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.
- 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.

Nội dung

Bài toán

Bài toán tối ưu đa mục tiêu không ràng buộc[1] có dạng như sau:

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

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

[1] J. Fliege, A. I. F. Vaz, and L. N. Vicente. "Complexity of gradient descent for multiobjective optimization". In: Optimization Methods and Software 34.5 (2019), pp. 949–959.

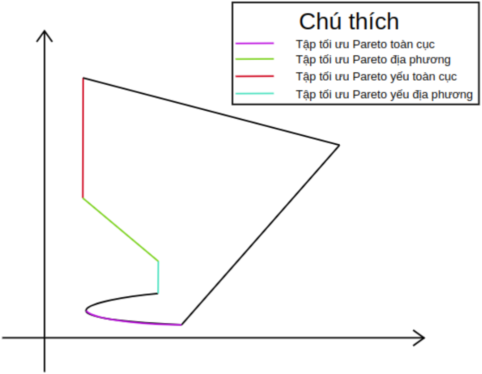
Tối ưu Pareto

Xét bài toán (1) với x^* là một vectơ trong $\mathbb{R}^n$.

1. Vectơ x^* là nghiệm tối ưu Pareto toàn cục nếu không tồn tại vectơ $x \in \mathbb{R}^n$ thỏa mãn $F(x) \leq F(x^*)$ và $F(x) \neq F(x^*)$.
2. Vectơ x^* là nghiệm tối ưu Pareto địa phương nếu tồn tại $\delta > 0$ thỏa mãn x^* là nghiệm tối ưu Pareto toàn cục trong $B(x^*, \delta)$.
3. Vectơ x^* là nghiệm tối ưu Pareto yếu toàn cục nếu không tồn tại vectơ $x \in \mathbb{R}^n$ thỏa mãn $F(x) < F(x^*)$.
4. Vectơ x^* là nghiệm tối ưu Pareto yếu địa phương nếu tồn tại $\delta > 0$ thỏa mãn x^* là nghiệm tối ưu Pareto yếu toàn cục trong $B(x^*, \delta)$.
5. Vectơ x^* được gọi là điểm tới hạn Pareto nếu:

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

Tối ưu Pareto



Hình 1: Ví dụ tập nghiệm tối ưu Pareto toàn cục, Pareto địa phương, Pareto yếu toàn cục, Pareto yếu địa phương

Nội dung

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:

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

$$\text{vđk } \nabla f_i(x^k)^\top d \leq \alpha, \quad i = 1, \ldots, m.$$

Trường hợp các hàm mục tiêu là không lồi[2].

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:

$$T^k := \left\{ t = \frac{1}{2^j} \middle| j \in \mathbb{N}_0,\ F(x^k + td^k) \le F(x^k) + \beta t \nabla F(x^k)^\top d^k \right\}. \tag{4}$$

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

[2] J. Fliege, A. I. F. Vaz, and L. N. Vicente. "Complexity of gradient descent for multiobjective optimization". In: Optimization Methods and Software 34.5 (2019), pp. 949–959.

Sự hội tụ của Thuật toán 1

Định lý 3.1

Giả sử rằng ít nhất một trong các hàm $f_1, \ldots, f_m$ *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*[3]*:*

$$\min_{0 \leq \ell \leq k-1} \|d^\ell\| \leq \sqrt{\frac{F_0^{max} - F^{min}}{M} \frac{1}{\sqrt{k}}}, \tag{5}$$

trong đó M *là một số thực dương.*

[3] J. Fliege, A. I. F. Vaz, and L. N. Vicente. "Complexity of gradient descent for multiobjective optimization". In: Optimization Methods and Software 34.5 (2019), pp. 949–959.

Thuật toán 2

Trường hợp các hàm mục tiêu lồi hoặc lồi mạnh[4].

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:

$$T^k := \left\{ t = \frac{1}{2^j} \,\middle|\, j \in \mathbb{N}_0, F(x^k + td^k) \leq F(x^k) + t\nabla F(x^k)^\top d^k + \frac{\gamma t}{2}\|d^k\|^2 e \right\}, \quad (6)$$

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.

[4] J. Fliege, A. I. F. Vaz, and L. N. Vicente. "Complexity of gradient descent for multiobjective optimization". In: Optimization Methods and Software 34.5 (2019), pp. 949–959.

Sự hội tụ của Thuật toán 2

Bằng cách chứng minh tương tự như **Định lý 3.1**, ta sẽ chứng minh được tốc độ hội tụ của Thuật toán 2 với $\min\limits_{0 \le \ell \le k-1} \|d^\ell\|$ là $O\left(\frac{1}{\sqrt{k}}\right)$.

Định lý 3.2

Cho các hàm $f_1, \ldots, f_m$ 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[5]:

$$\sum_{i=1}^{m} \bar{\lambda}_i^{k-1} \left(f_i(x^k) - f_i(x^*) \right) \le \frac{\|x^0 - x^*\|^2}{Mk}, \tag{7}$$

trong đó trọng số $\bar{\lambda}_i^{k-1} = \frac{1}{k} \sum_{\ell=0}^{k-1} \lambda_i^\ell$ và M là một số thực dương. Nếu $\{\lambda^k\}$ hội tụ đến λ^* thì $\{\bar{\lambda}^k\}$ cũng hội tụ đến λ^*.

[5] J. Fliege, A. I. F. Vaz, and L. N. Vicente. "Complexity of gradient descent for multiobjective optimization". In: Optimization Methods and Software 34.5 (2019), pp. 949–959.

Sự hội tụ của Thuật toán 2

Định lý 3.3

Cho các hàm f_i là lồi mạnh với hệ số lồi mạnh $\mu_i > 0$, $i = 1, \ldots, m$. Giả sử $\{x^k\}$ hội tụ đến x^*. Phương pháp gradient được mô tả trong Thuật toán 2 sinh ra dãy $\{x^k\}$ thoả mãn[6]*:*

$$\|x^k - x^*\| \leq \left(\sqrt{1 - M}\right)^k \|x^0 - x^*\|, \tag{8}$$

với M là một số thực dương.

[6] J. Fliege, A. I. F. Vaz, and L. N. Vicente. "Complexity of gradient descent for multiobjective optimization". In: Optimization Methods and Software 34.5 (2019), pp. 949–959.

Nội dung

Các tham số cài đặt

Các ví dụ được lấy từ bài báo "A Barzilai-Borwein descent method for multiobjective optimization problems"[7]. Chương trình của tác giả được viết bằng ngôn ngữ Python. Các tham số như sau:

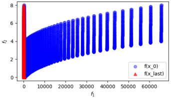
- Số lần lặp tối đa: 500.
- Sai số dừng: $\epsilon = 10^{-8}$.
- Hệ số $\beta = 10^{-4}$.
- Hệ số $\gamma = 0.99$.

[7] 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.

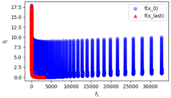
Mô tả các ví dụ được sử dụng để thực nghiệm

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

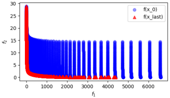
Kết quả chạy thuật toán



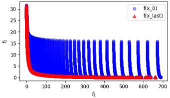
WIT 1
(Thuật toán 1)



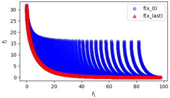
WIT 2
(Thuật toán 1)



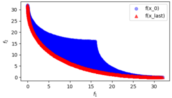
WIT 3
(Thuật toán 1)



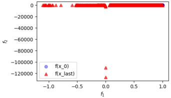
WIT 4
(Thuật toán 1)



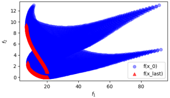
WIT 5
(Thuật toán 1)



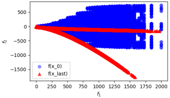
WIT 6
(Thuật toán 1)



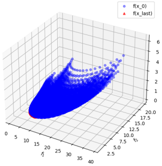
Deb
(Thuật toán 1)



PNR
(Thuật toán 1)

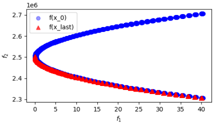


DD1
(Thuật toán 1)

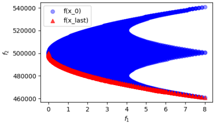


FDS
(Thuật toán 1)

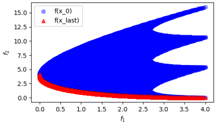
Kết quả chạy thuật toán



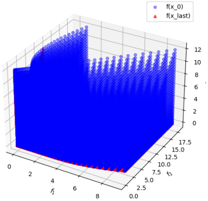
Imbalance 1
(Thuật toán 1)



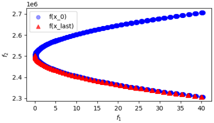
Imbalance 2
(Thuật toán 1)



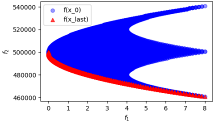
JOS1
(Thuật toán 1)



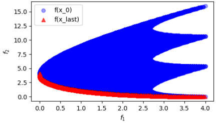
TRIDIA1
(Thuật toán 1)



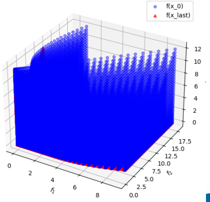
Imbalance 1
(Thuật toán 2)



Imbalance 2
(Thuật toán 2)



JOS1
(Thuật toán 2)



TRIDIA1
(Thuật toán 2)

Tổng hợp kết quả chạy thuật toán

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	- - - - - - - -		

Tổng hợp kết quả chạy thuật toán

Problem	Thuật toán 1			Thuật toán 2		
	iter	**time(ms)**	**stepsize**	**iter**	**time(ms)**	**stepsize**
WIT5	6.31	1.27	0.40	Không lồi		
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

Nội dung

Kết luận

Kết quả đồ án:

- Trình bày nội dung lý thuyết về bài toán tối ưu đa mục tiêu không ràng buộc tổng quát, nghiệm tối ưu Pareto cùng các khái niệm liên quan.
- 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:

- 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.

Cảm ơn thầy cô và các bạn đã chú ý lắng nghe!