

PHƯƠNG PHÁP HƯỚNG PHÂN GIÁC GIẢI BÀI TOÁN TỐI ƯU HÓA KHÔNG RÀNG BUỘC TỔNG QUÁT

NGUYỄN VĂN MẠNH - BUI MINH CHÍ
NGUYỄN THỨ LOAN

1. MỞ ĐẦU

Xét bài toán tối ưu hóa không ràng buộc tổng quát

$$\min \{ J(u) \mid u \in E^m \} \quad (1)$$

trong đó E^m là không gian Euclide m chiều, $J(u)$ là hàm phi tuyến.

Để giải bài toán (1) có vô số phương pháp, tuy nhiên các phương pháp đó tỏ ra kém hiệu quả (tốc độ hội tụ chậm) khi hàm cực tiểu có dạng « hình khe » phức tạp [4,7,8,9]. Hàm có cấu trúc khe là hàm có các mặt đẳng mức bị kéo dài thẳng hoặc cong

Định nghĩa 1. « Hướng phân giác » của hai vectơ a^1 và a^2 là hướng của vectơ c hình thành theo công thức

$$c = \frac{a^1}{\|a^1\|} + \frac{a^2}{\|a^2\|} \quad (2)$$

Ý chính của phương pháp « hướng phân giác » (sau này ta sẽ viết là HPG) là sau mỗi bước chuyển động theo hướng đối-gradien $g_A = -J'(u^A)$ vượt qua « lòng khe » từ điểm A tới điểm B ta không đi theo hướng đối-gradien $g_B = -J'(u^B)$ tại điểm B mà đi theo hướng phân giác giữa g_A và g_B .

2. THUẬT TOÁN HƯỚNG PHÂN GIÁC.

Giả sử đã chọn được điểm ban đầu u^0 nào đó, qui tắc xây dựng dãy điểm $\{u^k\}$ theo phương pháp HPG thể hiện như sau:

$$u^{k+1} = u^k + \alpha_k S_k, \quad \alpha_k > 0; \quad k = 0, 1, 2, \dots \quad (3)$$

$$S_0 = -J'(u^0), \quad S_{k+1} = -J'(u^{k+1}) + \beta_{k+1} S_k \quad (4)$$

trong đó

$$\beta_{k+1} = \begin{cases} 0, & \text{nếu } k \in I_1 \\ \|J'(u^{k+1})\| \|S_k\|^{-1}, & \text{nếu } k \in I_2 \end{cases} \quad (5)$$

Trong (3), (4), (5), S_k là vec-tơ chỉ hướng thay đổi của hàm $J(u)$ ở bước thứ k ; I_1 là tập tất cả các chỉ số k mà tại điểm u^k hướng S_{k-1} và $J'(u^k)$ trùng nhau, hay $k=0, r, 2r, \dots$ với r là số nguyên dương cho trước; I_2 là tập bổ sung của I_1 trong tập $\{0, 1, 2, \dots\}$, nghĩa là

$$I_1 \cup I_2 = \{0, 1, 2, \dots\}, \quad I_1 \cap I_2 = \emptyset \quad (6)$$

Những điểm mà tại đó $S_{k+1} = -J'(u^{k+1})$ gọi là điểm khôi phục ứng với khi $\beta_{k+1} = 0$. Trong (4) và (5) khi $k \in I_2$ thì S_{k+1} là hướng phân giác của góc tạo bởi hai vec-tơ $-J'(u^{k+1})$ và S_k . Điều này kiểm tra dễ dàng nếu ta thay S_{k+1} , $-J'(u^{k+1})$ và S_k vào công thức tính cosin của hai góc tạo bởi S_{k+1} với $-J'(u^{k+1})$ và S_k .

Giá trị của α_k trong (3) là độ dài của bước chuyển động từ u^k tới u^{k+1} theo hướng S_k . Tùy theo cách chọn giá trị của α_k mà ta sẽ có các phương án khác nhau. Ta sẽ xét ba phương án sau:

Phương án 1. Giá trị của α_k đã đạt tới lòng khe.

Đó là trường hợp khi α_k thỏa mãn điều kiện cực tiểu hàm $J(u)$ theo hướng S_k :

$$\alpha_k = \alpha_k^* : J(u^k + \alpha_k^* S_k) = \inf_{\alpha \geq 0} J(u^k + \alpha S_k) \quad (7)$$

Phương án 2. Giá trị của α_k chưa đạt tới lòng khe.

Đó là trường hợp khi α_k thỏa hệ thức sau:

$$0 < \varepsilon_1 \leq \alpha_k \leq \frac{1}{\varepsilon_2 + L}, \quad \varepsilon_2 > 0 \quad (8)$$

với L là hằng số Lipschitz tức là:

$$\|J'(u) - J'(v)\| \leq L \|u - v\|, \quad u, v \in E^m \quad (9)$$

Phương án 3. Giá trị của α_k đã vượt qua lòng khe.

Định nghĩa 2. Bước "vượt qua khe" là bước có giá trị α_k thỏa mãn điều kiện sau:

$$\left. \frac{d}{d\alpha} g_k(\alpha) \right|_{\alpha = \alpha_k} > 0 \quad (10)$$

trong đó ký hiệu $g_k(\alpha) = J(u^k + \alpha S_k)$.

Điều kiện (10) cho phép chọn bước α_k lớn hơn hoặc bằng giá trị bé nhất trong tập hợp điểm dừng $N(\alpha)$ của $g_k(\alpha)$, $N(\alpha) = \{\alpha \mid g'_k(\alpha) = 0\}$

Nó còn bảo đảm cho điểm $u^{k+1} = u^k + \alpha_k S_k$ nằm trên sườn dốc lên của "lòng khe" (theo hướng S_k). Như vậy từ u^k sang u^{k+1} quỹ đạo đã vượt qua khe. Điều kiện "vượt qua khe" (10) nhằm tăng cường tốc độ hội tụ của thuật toán HPG. Để đảm bảo hội tụ phải thực hiện đồng thời điều kiện đơn điệu sau

$$J(u^k + \alpha_k S_k) \leq J(u^k + \varepsilon_1 S_k), \quad 0 < \varepsilon_1 \leq \frac{1}{\varepsilon_2 + L}, \quad \varepsilon_2 > 0 \quad (11)$$

Trước khi đi vào từng phương án cụ thể ta nêu ra một số bổ đề làm cơ sở cho việc chứng minh các định lý hội tụ sau này.

Bổ đề 1. Giả sử vec-tơ S_k xây dựng theo qui tắc (4), (5). Ngoài ra chuẩn $\|J'(u_k)\| > 0$ với mọi k thì

$$\|S_k\| = \langle S_k, S_k \rangle^{1/2} > 0, \quad k = 0, 1, 2, \dots \quad (12)$$

$$\langle J'(u^k), S_k \rangle < 0 \quad (13)$$

Ở đây dấu $\langle \dots \rangle$ là ký hiệu tích vô hướng của hai vec-tơ.

Bất đẳng thức (13) trong bổ đề 1 chứng tỏ S_k luôn luôn là hướng giảm của hàm $J(u)$. Như vậy hướng phân giác dương nhiên là hướng giảm.

Từ bổ đề 1 ta có thể rút ra hệ quả sau;

Bổ đề 2. Với điều kiện của bổ đề 1 thì

$$\lim_{k \rightarrow \infty} J'(u^k) = 0 \quad \text{và} \quad \lim_{k \rightarrow \infty} S_k = 0 \quad (14)$$

khi và chỉ khi:

$$\lim_{k \rightarrow \infty} \langle J'(u^k), S_k \rangle = 0 \quad (15)$$

Bổ đề 3. Từ qui tắc (4), (5), xây dựng hướng chuyển động S_k ta có:

$$\frac{\langle J'(u^k), S_k \rangle}{\|S_k\|^4} = \begin{cases} -1 & \text{nếu } \beta_k = 0 \\ -1/2 & \text{nếu } \beta_k = \|J'(u^k)\| \cdot \|S_{k-1}\|^{-1} \end{cases} \quad (16)$$

Bổ đề 4. Nếu hàm $J(u)$ có đạo hàm liên tục và bị chặn dưới thì tồn tại ít nhất một giá trị $\alpha > 0$ sao cho các bất đẳng thức sau

$$\frac{d}{d\alpha} J(u^k + \alpha S_k) \geq 0 \quad (17)$$

$$J(u^k + \alpha S_k) \leq J(u^k + \varepsilon_1 S_k), \quad 0 < \varepsilon_1 \leq \frac{1}{L + \varepsilon_2} \quad (18)$$

thực hiện được đồng thời.

Để dễ dàng chứng minh định lý về tốc độ hội tụ của phương pháp HPG chúng ta sẽ chọn tập những thời điểm khởi phục I_1 hơi khác ở trên một chút. Cụ thể là I_1 bao gồm những chỉ số $k = 0, r, 2r, \dots$, và những chỉ số k khi

$$1 - \cos \varphi_k \leq \varepsilon_3 \leq 1/2, \quad \cos \varphi_k = \frac{\langle J'(u^k), S_{k-1} \rangle}{\|J'(u^k)\| \|S_{k-1}\|}, \quad \varepsilon_3 = \text{const} \quad (19)$$

Chọn ε_3 càng bé thì khả năng đi dọc lòng khe của thuật toán càng cao. Nếu chọn $\varepsilon_3 = 0$ thì qui tắc chọn tập I_1 ở đây sẽ trùng với trên.

Bổ đề 5. Giả sử β_k và S_k được xây dựng theo quy tắc (4), (5) và (10), khi đó ta có bất đẳng thức

$$\|S_k\|^2 \geq 2\varepsilon_3 \|J'(u^k)\|^2 \quad (20)$$

3. CÁC ĐỊNH LÝ HỘI TỤ

Các định lý hội tụ của thuật toán sẽ được phát biểu và chứng minh cho các phương án 2 và 3 nêu trong phần thuật toán.

Ta không quan tâm nhiều lắm đến phương án 1 trong bài báo này bởi lẽ để thực hiện điều kiện (7) nói chung phải giải bài toán tối ưu hàm một biến, ví dụ theo phương pháp « lát cắt vàng », là một thủ tục khá mất thì giờ và phức tạp. Phương án này không thể hiện rõ tính ưu việt của phương pháp HPG.

Đối với phương án 2 chọn α_k theo (8) cho ta một quỹ đạo đơn điệu nằm về một phía của lòng khe tức là chưa vượt qua que. Phương án này có tốc độ hội tụ chưa phải là thật cao nhưng chúng ta quan tâm tới nó để thuận lợi cho việc chứng minh sự hội tụ của phương án 3.

Định lý 1. Giả sử $J(u)$ là hàm trơn bị chặn dưới, $J'(u)$ thỏa điều kiện Lipschitz tức là

$$\inf_{u \in E^m} J(u) = J^h > -\infty \quad (21)$$

$$\|J'(u) - J'(v)\| \leq L \|u - v\|, \quad L = \text{const} \quad (22)$$

$u, v \in E^m$ và u^0 là điểm tùy ý đã chọn thì dãy điểm $\{u^k\}$ xây dựng theo quy tắc (3), (4), (5) và (8) sẽ hội tụ về điểm dừng của $J(u)$.

Đối với phương án 3 chọn giá trị α_k để vượt qua lòng khe. Đây là phương án mà chúng tôi tập trung chú ý trong bài báo này. Qua tính toán thử nghiệm cho thấy rằng đối với phương án này thuật toán HPG hội tụ nhanh hơn nhiều so với phương pháp Gradient.

Định lý 2. Giả sử $J(u)$ có đạo hàm liên tục, bị chặn dưới; $J'(u)$ thỏa điều kiện Lipschitz; u^0 là điểm ban đầu tùy ý thì dãy điểm $\{u^k\}$ xây dựng theo quy tắc (3), (4), (5), (10) và (11) sẽ hội tụ về điểm dừng của $J(u)$.

Ngoài ra nếu $J(u)$ làm lồi và tập hợp $M(u^0) = \{u | J(u) \leq J(u^0)\}$ giới nội thì $\{u^k\}$ là dãy cực tiểu hóa của $J(u)$. Trong trường hợp điểm cực tiểu là duy nhất thì mọi dãy $\{u^k\}$ đều hội tụ về điểm đó.

Định lý 3. Giả sử hàm $J(u)$ lồi, có đạo hàm liên tục trong E^m , $J'(u)$ thỏa điều kiện Lipschitz; dãy điểm $\{u^k\}$ xây dựng theo quy tắc (3), (4), (5), (17), (18) và (19) thì với mọi k ,

$$0 \leq J(u^k) - J^* \leq \frac{D(L + \varepsilon_2)}{\varepsilon_1 \varepsilon_2 \varepsilon_3} \cdot \frac{1}{k} \quad (23)$$

Ở đây $D = \sup_{u, v \in M(u^0)} \|u - v\|$ là đường kính lớn nhất của tập $M(u^0)$;

$J^* = \inf_{u \in E^m} J(u) = J(u^*)$; L là hằng số Lipschitz.

Nếu $J(u)$ là lồi mạnh trong E^m thì

$$0 \leq J(u^k) - J^* \leq (J(u^0) - J^*) q^k \quad (24)$$

$$\|u^k - u^*\|^2 \leq \frac{2}{\%} (J(u^0) - J^*) q^k \quad (25)$$

$$q = 1 - \frac{\varepsilon_1 \varepsilon_2 \varepsilon_3 \mu}{L + \varepsilon_2}, \quad 0 < q < 1, \quad \%, \mu = \text{const} > 0.$$

4. THỰC NGHIỆM TÍNH TOÁN VÀ KẾT LUẬN

Kết quả tính toán tìm tối ưu thực hiện trên máy vi tính; chương trình được viết bằng ngôn ngữ BASIC. Để dễ so sánh hiệu quả của thuật toán HPG so với hai phương pháp phổ biến là « Gradien liên hợp » và « Hạ nhanh nhất », cả ba thuật toán trên đều được lập trình bằng ngôn ngữ BASIC và được thử cùng với 4 ví dụ khác nhau.

Ví dụ 1. $J_1(u) = \sum_{i=1}^{11} \left(Y_i - \frac{u_2}{\sqrt{u_1^2 + W_i^2}} \right)^2$

Đây là bài toán nhận dạng khâu quán tính bậc nhất lấy trong [9] theo các số liệu sau:

Số T. T.	1	2	3	4	5	6	7	8	9	10	11
W_i	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
Y_i	71	67	64	61	58	55	53	51	49	47	45

Điểm ban đầu u^0 bằng $(u_1^0, u_2^0) = (0, 7, 70)$;

$$J(u_0) = 1865.05.$$

Kết quả tính toán được biểu diễn trên hình 1.

Ví dụ 2. $J_2(u) = 100(u_2 - u_1^2)^2 + (1 - u_1)^2$.

Đây là hàm thử do Rosenbrock H. H. đưa ra, có lòng khe sâu. Điểm xuất phát chọn theo [15] $u^0 = (-1.2, 1)$, $J(u^0) = 24.2$; điểm cực tiểu là $u^* = (1, 1)$, $J(u^*) = 0$. Kết quả tính toán được biểu diễn trên hình 2.

Ví dụ 3. $J_3(u) = u_1^2 + 10u_2^2 + 100u_3^2 + 200u_4^2$

Đây là hàm toàn phương 4 biến số có khe tương đối sâu. Điểm xuất phát $u^0 = (1, 1, 1, 2)$, $J(u^0) = 911$. Điểm cực tiểu là $u^* = (0, 0, 0, 0)$, $J(u^*) = 0$. Kết quả tính toán được biểu diễn trên hình 3.

Ví dụ 4. $J_4(u) = (u_1 + 10u_2)^2 + 5(u_3 - u_4)^2 + (u_2 - 2u_3)^4 + 10(u_1 - u_4)^4$.

Đây là hàm thử do Powell M. J. D. đưa ra năm 1964, lấy trong [15]. Điểm xuất phát $u^0 = (3, -1, 0, 1)$, $J(u^0) = 215$. Điểm cực tiểu $u^* = (0, 0, 0, 0)$, $J(u^*) = 0$. Hàm này có đặc điểm là tại điểm cực tiểu ma trận đạo hàm bậc 2 suy biến. Kết quả tính toán được biểu diễn trên hình 4.

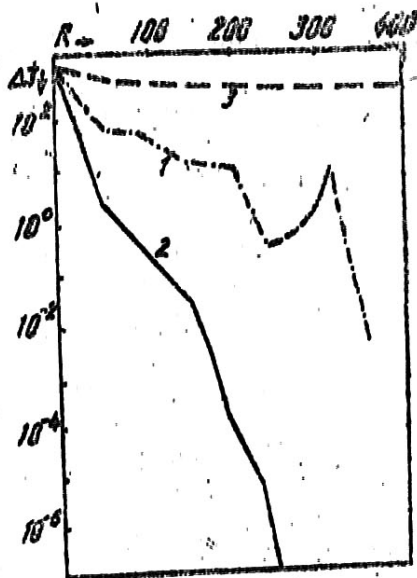
Trên các đồ thị 1, 2, 3, 4 có sử dụng các kí hiệu như sau; GLH là kí hiệu thuật toán « Gradien liên hợp », HPG - thuật toán « Hướng phân giác », HNN = thuật toán « Hạ nhanh nhất » R - số lần tính giá trị của hàm cực tiểu hóa, $\Delta J = J(u) - J^*$ là độ sai lệch của giá trị hiện tại của hàm số so với giá trị cực tiểu.

Nhận xét kết quả tính toán

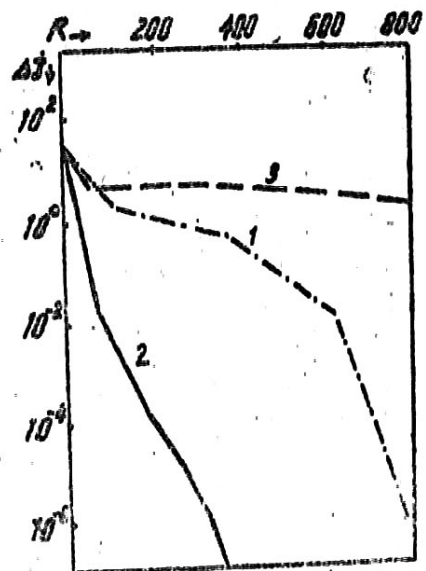
Kết quả tính toán ví dụ 1 và 2 cho ta thấy thuật toán HPG hội tụ nhanh hơn phương pháp « Gradien liên hợp » và nhanh gấp hàng chục lần so với phương pháp « Hạ nhanh nhất ». Còn đối với các ví dụ 3 và 4 thì tốc độ hội tụ của phương pháp HPG và phương pháp « Gradien liên hợp » tương đương nhau, nhưng gấp hàng chục lần tốc độ hội tụ của phương pháp « Hạ nhanh nhất ».

KẾT LUẬN

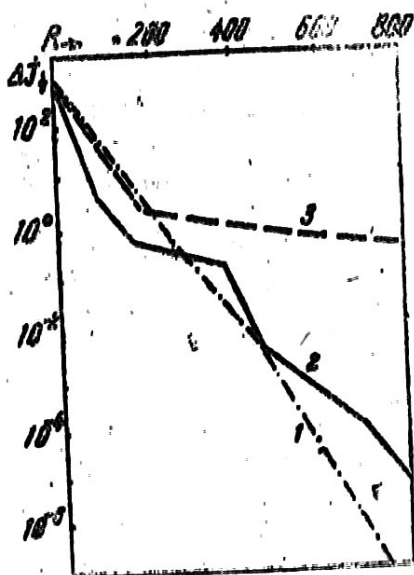
- a) Phương pháp HPG có tốc độ hội tụ cấp số nhân;
- b) Thuật toán HPG đã cho một hiệu quả tăng tốc độ hội tụ rất đáng kể so với phương pháp Gradient. Trong nhiều trường hợp, thuật toán HPG có tốc độ hội tụ thực tế xấp xỉ hoặc có khi hơn tốc độ hội tụ thực tế của thuật toán « Gradient liên hợp »;
- c) Ở giai đoạn đầu đối với các hàm khe phức tạp, thuật toán HPG nói chung hội tụ nhanh hơn thuật toán « Gradient liên hợp ».



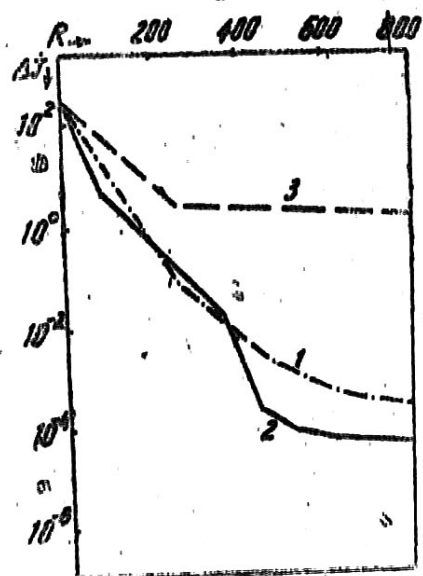
Hình 1



Hình 2



Hình 3



Hình 4

Trên các hình 1, 2, 3, 4 đường cong số 1 biểu diễn kết quả tìm tối ưu (tức là đường cong phụ thuộc giữa $\Delta J = J(u) - J^*$ và số R lần tính giá trị hàm $J(u)$) theo phương pháp « Gradient liên hợp »; đường số 2 như trên, theo phương pháp « lượng phân giác »; đường số 3 cũng vậy, theo phương pháp « li nhanh nhất ».

Nhận ngày 10-4-1987

ТÀI LIỆU THAM KHẢO

1. Конторович Л.В., Акилов Г.П., функциональный анализ. Издат. «Наука», М., 1977, 741 с.
2. Поляк Б.Т., Методы минимизации функций многих переменных. «Экономика и Матем. методы», №6, М., 1967 (обзор).
3. Поляк Б.Т., Методы сопряженных градиентов в задачах на экстремум, «Журнал вычисл. Математики и матем. Физики», №4, М., 1969.
4. Поляк Б.Т., Минимизация негладких функционалов. «Журнал вычисл. математики и матем. Физики», №3, М., 1969.
5. Шор Н.Э., Методы минимизации недифференцируемых функций и их приложения. Издат. «Наукова думка», К., 1979, 198с.
6. Нурминский Е.А., Численные методы решения детерминированных и стохастических минимаксных задач. Издат. «Наукова думка», К., 1979.
7. Моисеев Н.Н., Элементы теории оптимальных систем. Издат. «Наука», М., 1975.
8. Васильев Ф.П., Численные методы решения экстремальных задач. Издат. «Наука», М., 1974 и М., 1980.
9. Ракитский Ю.В., Устинов С.М., Чернолуцкий И.Г., Численные методы решения жестких систем. Издат. «Наука», М., 1979.
10. Huang. H., Unified approach to quadratically convergent algorithms for function minimization. J. Optimization theory and Application, 5, №6, 1970.
11. Ермольев Ю.М., Методы решения нелинейных экстремальных задач. «Кибернетика», №4, К., 1966.
12. Поляк Б.Т., Градиентные методы минимизации функционалов. «Журнал вычисл. Математики и математической физики», №3, 4, М., 1963.
13. Конторович Л.В., Приближенное решение функциональных уравнений «Успехи матем. науки», выпуск 6, 1956.
14. Fletcher R., Reeves C.M. Function Minimization by Conjugate gradients. Computer J., 1964, V.7, №2.
15. Химельблау Д., Прикладной нелинейное программирование. Издат. «Мир», М., 1975. Перевод с английского.
16. Fletcher R., Powell M.J.D. A rapidly convergent descent method for minimization. Computer J., 1963, V.6, №2.
17. Davidon W.C., Variable metric methods for minimization. AEC Research and development, Rept. ANL 5090 (Rev.), 1959.
18. Гупал А.М., Стохастические методы решения негладких экстремальных задач. Издат. «Наукова думка» К., 1979.

ABSTRACT

METHOD OF «BISECTOR DIRECTION» FOR SOLVING THE UNCONSTRAINED OPTIMISATION PROBLEM

In this paper the general unconstrained optimisation problem is considered. Here are new method, called the method of «bisector direction», is proposed for solving this problem. Our algorithm is based on the convergence theorems and it has a high rate of convergence. Some examples are given also.