PHƯƠNG PHÁP CHIẾU DƯỚI GRADIENT XẤP XỈ GIẢI BÀI TOÁN QUY HOẠCH LỒI

Quách Thị Mai Liên*, Hoàng Phương Khánh
Trường Đại học Công nghệ thông tin và Truyền thông - ĐH Thái Nguyên

TÓM TẮT

Từ trước đến nay, có rất nhiều phương pháp giải bài toán Quy hoạch lồi, trong đó chủ yếu là các phương pháp sử dụng đạo hàm. Bài báo trình bày một phương pháp giải bài toán quy hoạch lồi là phương pháp chiếu gradient xấp xỉ. Phương pháp này áp dụng tốt cho bài toán $\min_{x\in D} f(x)$(P) với hàm mục tiêu $f(x)$ không khả vi. Bài báo trình bày các khái niệm cơ bản để phục vụ việc chứng minh điều kiện để thuật toán chiếu dưới gradient xấp xỉ hội tụ về nghiệm của bài toán (P) là chỉ cần tập ε-dưới vi phân của hàm f bị chặn và hàm f nửa liên tục dưới.

Từ khóa: *Quy hoạch lồi, Dưới vi phân; Phương pháp chiếu dưới gradient; Hội tụ; Bị chặn*

GIỚI THIỆU

Lý thuyết về bài toán quy hoạch lồi đã được quan tâm nghiên cứu nhiều và đã thu được nhiều kết quả quan trọng dựa trên lý thuyết của Quy hoạch lồi và Tối ưu hóa. Về phương diện tính toán, đã có khá nhiều phương pháp hữu hiệu cho lớp bài toán này. Các phương pháp đó đã được giới thiệu trong cuốn Tối ưu lồi (Convex Optimization) của các tác giả Stephen Boyd and Lieven Vandenberghe do nhà xuất bản Cambridge University Press in năm 2004. Bài báo trình bày một phương pháp giải quyết bài toán quy hoạch lồi với lớp hàm mục tiêu không khả vi.

TẬP DƯỚI VI PHÂN, ξ - CHIẾU XUỐNG TẬP LỒI

Định nghĩa 2.1.[1], [2], [4], [6] *Cho* $\varepsilon > 0$. *Một véc tơ* $\omega \in R^n$ *được gọi là một* ε*-dưới gradient của* f *tại điểm* $x_0 \in R^n$ *nếu*

$$\langle \omega, x - x^0 \rangle \le f(x) - f(x^0) + \varepsilon, \quad \forall x \in R^n$$

Tập hợp tất cả các ε *- dưới gradient của hàm* f *tại* x^0 *gọi là* ε *- dưới vi phân hàm* f *tại* x_0*, ký hiệu là*

$$\partial_\varepsilon f(x^0) = \left\{ \omega \in R^n : \langle \omega, x - x^0 \rangle \le f(x) - f(x^0) + \varepsilon, \ \forall x \in R^n \right\}$$

Định nghĩa 2.2.[1], [2], [4], [6] *Véc tơ* $\omega \in R^n$ *được gọi là một dưới gradient của* f *tại điểm* $x_0 \in R^n$ *nếu*

$$\langle \omega, x - x^0 \rangle \le f(x) - f(x^0), \quad \forall x \in R^n$$

Tập hợp tất cả các dưới gradient của hàm f *tại* x^0 *gọi là dưới vi phân của f tại* x_0*, ký hiệu là*

$$\partial f(x^0) = \left\{ \omega \in R^n : \langle \omega, x - x^0 \rangle \le f(x) - f(x^0), \forall x \in R^n \right\}$$

Định nghĩa 2.3.[3] *Giả sử* $\xi > 0$ và $x \in R^n$. *Điểm* $p_x \in C$ *được gọi là* ξ *- chiếu của* x vào C *nếu* p_x *là* ξ*-nghiệm của bài toán* $\min_{y \in C} \left\{ \frac{1}{2} \|x - y\|^2 \right\}$*, nghĩa là*

$$\frac{1}{2}\|x - p_x\|^2 \le \frac{1}{2}\|x - P_C(x)\|^2 + \xi$$

trong đó $P_C(x)$ *là hình chiếu khoảng cách của* x *lên* C.

Nhận xét 2.1. *Từ định nghĩa ta thấy,* $p_x \in C$ là ξ - chiếu của x vào C tương đương với

$$\langle x - p_x, p_x - y \rangle \ge -\xi, \quad \forall y \in C.$$

THUẬT TOÁN CHIẾU DƯỚI GRADIENT XẤP XỈ

Cho ρ, ε là các tham số dương và các dãy số thực $\{\rho_k\}, \{\beta_k\}, \{\varepsilon_k\}, \{\xi_k\}$ thỏa mãn các điều kiện sau:

$$\rho_k > \rho, \beta_k > 0, \varepsilon_k \ge 0, \xi_k \ge 0, \forall k \in N \qquad (3.1)$$

$$\sum \frac{\beta_k}{\rho_k} = +\infty, \sum \beta_k^2 < +\infty \qquad (3.2)$$

* Tel: 0973 252666, Email: qtmlien@ictu.edu.vn

$$\sum \frac{\beta_k \varepsilon_k}{\rho_k} < +\infty, \sum \xi_k < +\infty \qquad (3.3)$$

Chẳng hạn có thể lấy

$$\beta_k = \frac{1}{k}, \rho_k = 1, \varepsilon_k = \frac{1}{k}, \xi_k = \frac{1}{k(k+1)}$$

Thuật toán: Chọn $x^0 \in C$. Tại mỗi bước lặp k = 0, 1,... có x^k

Bước 1: Giả sử $x^k \in C$. Lấy $g^k \in \partial_{\varepsilon_k} f(x^k)$.

Ta định nghĩa

$\alpha_k = \dfrac{\beta_k}{\gamma_k}$ trong đó $\gamma_k = max\left\{\rho_k, \left\|g^k\right\|\right\}$.

Nếu $g^k = 0$ thì x^k là ε_k-nghiệm của bài toán (P) và dừng thuật toán nếu $\varepsilon_k \le \varepsilon$. Trái lại chuyển sang Bước 2.

Bước 2 Tính $x^{k+1} \in C$

$\langle \alpha_k g^k + x^{k+1} - x^k, x - x^{k+1}\rangle \ge -\xi_k, \ \forall x \in C$

Nếu $\left\|x^{k+1} - x^k\right\| \le \varepsilon$ thì dừng và ta được một nghiệm xấp xỉ.

Ngược lại quay về Bước 1.

Chú ý 3.1. Ta có

- Theo nhận xét 2.1 điểm x^{k+1} được gọi là ξ_k-chiếu của $\left(x_k - \alpha_k g^k\right)$ vào C. Đặc biệt, nếu $\xi_k = 0$ thì x^{k+1} là hình chiếu vuông góc của $\left(x_k - \alpha_k g^k\right)$ hay $x^{k+1} = P_C\left(x_k - \alpha_k g^k\right)$.

- Nếu $\xi_k = \varepsilon_k = 0, \forall k \in N$ thì p_x là phép chiếu vuông góc và khi đó thuật toán trở thành thuật toán chiếu gradient, tiêu chuẩn dừng là $g_k = 0$ ở bước 1, và $x^{k+1} = x^k$ ở bước 2.

Bổ đề 3.1. *Với mọi k, ta có các bất đẳng thức sau*

(i) $\alpha_k \left\|g^k\right\| \le \beta_k$,

(ii) $\beta_k \left\|x^{k+1} - x^k\right\| \le \beta_k^2 + \xi_k$.

Chứng minh.

(i) Theo định nghĩa α_k ta có:

$$\alpha_k \left\|g^k\right\| = \frac{\beta_k \left\|g^k\right\|}{\max\left\{\rho_k, \left\|g^k\right\|\right\}} \le \beta_k$$

(ii) Thay $x = x^k$ vào (3.4) ta có:

$$\left\|x^{k+1} - x^k\right\|^2 \le \langle \alpha_k g^k, x^k - x^{k+1}\rangle + \xi_k$$

$$\le \alpha_k \left\|g^k\right\| \cdot \left\|x^{k+1} - x^k\right\| + \xi_k$$

(Theo Bất đẳng thức Cauchy-Schwarz)

$$\le \beta_k \left\|x^{k+1} - x^k\right\| + \xi_k$$

(Theo chứng minh (i).

Hay $\left\|x^{k+1} - x^k\right\|^2 - \beta_k \left\|x^{k+1} - x^k\right\| - \xi_k \le 0$ (3.5)

Xét phương trình bậc hai $f(\theta) = \theta^2 - \beta\theta - \xi$

$$\theta \le \frac{\beta + \sqrt{\beta^2 + 4\xi}}{2} \qquad (3.6)$$

Nhân cả hai vế của bất đẳng thức (3.6) với β và sử dụng bất đẳng thức Cauchy cho hai số dương a và b, $ab \le \dfrac{a^2 + b^2}{2}$ ta có:

$$\beta\theta \le 2^{-1}\left(\beta^2 + \beta\sqrt{\beta^2 + 4\xi}\right)$$

$$\le 2^{-1}\left(\beta^2 + \frac{\beta^2 + \beta^2 + 4\xi}{2}\right) = \beta^2 + \xi$$

Thay vào (3.5) với

$\theta = \left\|x^{k+1} - x^k\right\|, \beta = \beta_k, \xi = \xi_k$ ta được.

$$\beta_k \left\|x^{k+1} - x^k\right\| \le \beta_k^2 + \xi_k. \qquad \square$$

Mệnh đề 3.1. *Giả sử bài toán (P) có tập nghiệm S(f,C) khác rỗng. Khi đó với mọi* $x^* \in S(f,C)$ *và với mọi k, ta có:*

(i) $\left\|x^{k+1} - x^*\right\|^2 \le \left\|x^k - x^*\right\|^2 + 2\alpha_k\left(f\left(x^*\right) - f\left(x^k\right)\right) + \delta_k$,

trong đó $\delta_k = 2\alpha_k\varepsilon_k + 2\beta_k^2 + 4\xi_k$.

(ii) Dãy $\left\{\left\|x^k - x^*\right\|^2\right\}$ *hội tụ với mọi* $x^* \in S(f,C)$

(iii) Dãy $\left\{x^k\right\}$ *bị chặn.*

Chứng minh. Ta có:

$$\left\|x^k - x^*\right\|^2 = \left\|x^{k+1} - x^k\right\|^2 - 2\langle x^k - x^{k+1}, x^* - x^{k+1}\rangle + \left\|x^{k+1} - x^*\right\|^2.$$

Nên $$\left\|x^{k+1} - x^*\right\|^2 = \left\|x^k - x^*\right\|^2 - \left\|x^{k+1} - x^k\right\|^2 + 2\langle x^k - x^{k+1}, x^* - x^{k+1}\rangle$$

$$\le \left\|x^k - x^*\right\|^2 + 2\langle x^k - x^{k+1}, x^* - x^{k+1}\rangle \quad (3.7)$$

Thay $x = x^*$ vào (3.4) ta có:

$$\langle \alpha_k g^k + x^{k+1} - x^k, x^* - x^{k+1}\rangle \ge -\xi_k.$$

Suy ra

$$2\langle x^k - x^{k+1}, x^* - x^{k+1}\rangle \le 2\langle \alpha_k g^k, x^* - x^{k+1}\rangle + 2\xi_k \quad (3.8)$$

Thay (3.8) vào (3.7) ta được.

$$\left\|x^{k+1} - x^*\right\|^2 \le \left\|x^k - x^*\right\|^2 + 2\langle \alpha_k g^k, x^* - x^{k+1}\rangle + 2\xi_k$$

$$= \left\|x^k - x^*\right\|^2 + 2\langle \alpha_k g^k, x^* - x^k\rangle + 2\langle \alpha_k g^k, x^k - x^{k+1}\rangle + 2\xi_k$$

Áp dụng Bất đẳng thức Cauchy-Schwarz và Bổ đề 3.1 (i) ta có bất đẳng thức sau:

$$\left\|x^{k+1} - x^*\right\|^2 \le \left\|x^k - x^*\right\|^2 + 2\alpha_k\langle g^k, x^* - x^k\rangle + 2\beta_k\left\|x^k - x^{k+1}\right\| + 2\xi_k$$

$$\le \left\|x^k - x^*\right\|^2 + 2\alpha_k\langle g^k, x^* - x^k\rangle + 2\beta^2 + 4\xi_k \quad (3.9)$$

(Theo bổ đề 3.1(ii))

Mặt khác, $g^k \in \partial_{\varepsilon_k} f(x^k)$ nên ta có: $\langle g^k, x^* - x^k\rangle \le f(x^*) - f(x^k) + \varepsilon_k$,

Do $\alpha_k > 0$ nhân cả hai vế của bất đẳng thức trên với $2\alpha_k$ ta được:

$$2\alpha_k\langle g^k, x^* - x^k\rangle \le 2\alpha_k f(x^k, x^*) + 2\alpha_k\varepsilon_k \quad (3.10)$$

Thay (3.10) vào (3.9) ta được:

$$\left\|x^{k+1} - x^*\right\|^2 \le \left\|x^k - x^*\right\|^2 + 2\alpha_k\left(f\left(x^*\right) - f\left(x^k\right)\right) + \delta_k \quad (3.11)$$

Trong đó $\delta_k = 2\alpha_k\varepsilon_k + 2\beta_k{}^2 + 4\xi_k$.

(ii) Do $x^* \in S(f, C)$ nên $f\left(x^k\right) \le f\left(x^*\right)$ hay $f\left(x^*\right) - f\left(x^k\right) \le 0$ thay vào (3.11) ta được $\left\|x^{k+1} - x^*\right\|^2 \le \left\|x^k - x^*\right\|^2 + \delta_k$, (3.12)

trong đó $\delta_k = 2\alpha_k\varepsilon_k + 2\beta_k{}^2 + 4\xi_k$.

Theo (3.2), (3.3), (3.4) thì

$$\delta_k > 0 \text{ và } \sum_{k=0}^{+\infty}\delta_k < +\infty \quad (3.13)$$

Từ (3.12) và (3.13) ta được dãy $\left\{\left\|x^k - x^*\right\|^2\right\}$ hội tụ.

(iii) Do dãy $\left\{\left\|x^k - x^*\right\|^2\right\}$ hội tụ ta suy ra dãy $\left\{x^k\right\}$ bị chặn. □

Định lí 3.1. *Giả sử bài toán (P) có tập nghiệm S(f,C) khác rỗng và dãy* $\left\{g^k\right\}$ *bị chặn. Khi đó:*

$$\lim_{k\to+\infty}\sup\left(f\left(x^k\right) - f\left(x^*\right)\right) = 0, \quad \forall x \in S(f, C)$$

Chứng minh. Giả sử $x^* \in S(f, C)$. Theo Mệnh đề 3.1(i) ta có:

$$0 \le 2\alpha_k\left(f(x^k) - f(x^*)\right) \le$$

$$\le \left\|x^k - x^*\right\|^2 - \left\|x^{k+1} - x^*\right\|^2 + \delta_k,$$

trong đó $\delta_k = 2\alpha_k\varepsilon_k + 2\beta_k^2 + 4\xi_k > 0$ và $\sum_{k=0}^{+\infty}\delta_k < +\infty$ theo (3.2), (3.3),(3.4).

Lấy tổng các bất đẳng thức trên với k=0,1...m ta có:

$$0 \le 2\sum_{k=0}^{m}\alpha_k\left(f(x^k) - f(x^*)\right)$$

$$\le \left\|x^0 - x^*\right\|^2 - \left\|x^{m+1} - x^*\right\|^2 + \sum_{k=0}^{m}\delta_k$$

$$\le \left\|x^0 - x^*\right\|^2 + \sum_{k=0}^{m}\delta_k.$$

Cho $m \to +\infty$ ta được:

$$0 \le 2\sum_{k=0}^{+\infty}\alpha_k\left(f(x^k) - f(x^*)\right) + \sum_{k=0}^{+\infty}\delta_k$$

Vì $\sum_{k=0}^{+\infty}\delta_k < +\infty$ nên

$$0 \le 2\sum_{k=0}^{+\infty}\alpha_k\left(f(x^k) - f(x^*)\right) < +\infty \quad (3.14)$$

Theo giả thiết $\{g^k\}$ bị chặn, từ (3.1) và (3.4) ta có tồn tại $L \ge \rho$ sao cho $\left\|g^k\right\| \le L$ với mọi $k \in N$. Do đó:

$$\frac{\gamma_k}{\rho_k} = \max\left\{1, \rho_k^{-1}\left\|g^k\right\|\right\} \ge \frac{L}{\rho}, \quad \forall k \in N$$

Khi đó, theo (3.4) thì:

$$\alpha_k = \frac{\beta_k}{\gamma_k} \ge \frac{\rho}{L}\frac{\beta_k}{\rho_k}, \quad \forall k \in N \quad (3.15)$$

Từ (3.14) và (3.15) ta có:

$$\sum_{k=0}^{+\infty}\frac{\beta_k}{\rho_k}\left(f(x^k)-f(x^*)\right)<+\infty$$

Do $\sum_{k=0}^{+\infty}\frac{\beta_k}{\rho_k}=+\infty$ nên từ đây suy ra

$$\lim_{k\to+\infty}\sup\left(f(x^k)-f(x^*)\right)=0,\ \forall x^*\in S(f,C)\ \square$$

Định lí 3.2. *Giả sử bài toán P có tập nghiệm* $S(f,C)$ *khác rỗng, dãy* $\{g^k\}$ *bị chặn và f nửa liên tục dưới trên tập C. Khi đó dãy* $\{x^k\}$ *hội tụ đến nghiệm của bài toán P*

Chứng minh. Giả sử $x^*\in S(f,C)$. Theo định nghĩa của limsup tồn tại một dãy con $\{x^{k_j}\}$ của $\{x^k\}$ sao cho.

$$\lim_{k\to+\infty}\sup\left(f(x^k)-f(x^*)\right)=\lim_{j\to+\infty}\left(f(x^{k_j})-f(x^*)\right)$$

Theo Mệnh đề 3.1 (iii), dãy $\{x^{k_j}\}$ bị chặn. Vì vậy, không mất tính tổng quát ta giả sử

$$\lim_{j\to+\infty}x^{k_j}=\overline{x}. \qquad (3.16)$$

Theo giả thiết f nửa liên tục dưới trên tập C và từ Định lý 3.1 ta có.

$$f(x^*)-f(\overline{x})\geq\lim_{j\to+\infty}\inf\left(f(x^*)-f(x^{k_j})\right)$$

$$=-\lim_{j\to+\infty}\sup\left(f(x^{k_j})-f(x^*)\right)$$

$$=-\lim_{j\to+\infty}\sup\left(f(x^k)-f(x^*)\right)=0$$

Do $x^*\in S(f,C)$ nên $f(x^*)-f(\overline{x})\leq 0$ cho nên ta có:

$$f(x^*)-f(\overline{x})=0.$$

hay $f(x^*)=f(\overline{x})$ nên $\overline{x}\in S(f,C)$

Khi đó theo Mệnh đề 3.1 (ii) dãy $\left\{\|x^k-\overline{x}\|^2\right\}$ hội tụ, kết hợp với (3.16) ta được:

$$\lim_{k\to+\infty}x^k=\overline{x},\quad \overline{x}\in S(f,C).\ \square$$

KẾT LUẬN

Bài báo đã giới thiệu thuật toán chiếu gradient xấp xỉ để giải bài toán quy hoạch lồi $\min f(x)$ với $x\in C$ (P) và chứng minh sự hội tụ về nghiệm tối ưu của thuật toán. Thuật toán chiếu dưới gradient xấp xỉ áp dụng tốt cho việc giải bài toán (P) mà hàm mục tiêu f không khả vi.

Bài báo là sản phẩm của đề tài có mã số **T2016-07-16, được** tài trợ bởi kinh phí của trường Đại học CNTT&TT - ĐHTN

TÀI LIỆU THAM KHẢO

1. Stephen Boyd and Lieven Vandenbergh (2004), *Convex Optimization*, Cambridge University Press
2. Lê Dũng Mưu, Nguyễn Văn Hiền, *Nhập môn Giải tích lồi ứng dụng*, Nxb Khoa học tự nhiên và công nghệ, Hà Nội
3. Paulo Santos and Susan Scheimberg (2011), "An inexact subgradient algorithm for equilibrium problems", *Computational &Applied Mathmatics*, 30, pp. 91-107.
4. Trần Vũ Thiệu và Nguyễn Thị Thu Thủy (2011), *Nhập môn Tối ưu phi tuyến*, Nxb Đại học Quốc gia Hà Nội, Hà Nội
5. Hoang Tuy (1997), *Convex Analysis and Global Optimization*, Kluwer Academic Publishers
6. Hoàng Tụy (2006), *Lý thuyết tối ưu*, Viện Toán học, Hà Nội.

SUMMARY
AN INXACT SUBGRADIENT ALGORITHM FOR CONVEX OPTIMIZATION

Quach Thi Mai Lien*, Hoang Phuong Khanh
University of Information and Comunication Technology – TNU

So far, there have been many methods to solve convex programming problems, most of which are those using derivatives. The paper presents an inexactsubgradient projection method, a method to solve convex programming problems. This method is used effectively for $\min_{x\in D} f(x)$ (P) when objective function $f(x)$ is non-differentiable. The paper also prove that conditions for the inexactsubgradient projection algorithm to converge to the solutions of (P) only require ε-subdifferential of function f is bounded and f is lower semicontinuos

Keywords: *Convex optimization, Subgradient, Projected subgradient method; Convergent, Bounded*

Ngày nhận bài:27/11/2016; Ngày phản biện: 12/12/2016; Ngày duyệt đăng: 31/5/2017

* *Tel 0973 252666, Email. qtmlien@ictu.edu.vn*