Tạp chí Tin học và Điều khiển học, T. 17, S. 2 (2001), 45–50

XÁC ĐỊNH GRADIENT CỦA MỘT HÀM BẰNG PHƯƠNG PHÁP MONTE-CARLO

TRẦN CẢNH

Abstract. In the work the gradient: grad $f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n}\right)$ of a differentiable function $f(x)$ is determined by random model. The construction of an unbiassed estimator $\varsigma(x) = (\varsigma_1(x), \dots, \varsigma_n(x))$ of grad $f(x)$ is established successfully.

Tóm tắt. Trong công trình này gradient: grad $f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n}\right)$ của một hàm khả vi được xác định bằng một mô hình ngẫu nhiên. Việc thiết lập một ước lượng không chệch $\varsigma(x) = (\varsigma_1(x), \dots, \varsigma_n(x))$ được xác lập thành công.

1. MỞ ĐẦU

Lược đồ dò tìm ngẫu nhiên đã được sử dụng một cách hữu hiệu đối với một loại bài toán điều khiển cỡ lớn để cho lời giải tối ưu toàn cục (xem [1]). Ở đây sự hội tụ của lời giải gần đúng về lời giải đúng (theo quan điểm xác suất) và việc đánh giá "sai số" theo số phép lặp N_0 cũng được chỉ ra.

Tuy nhiên, nhiều bài toán điều khiển loại này, nhất là các bài toán cực trị toàn cục (xem [3]) đòi hỏi một độ chính xác cao hơn, buộc chúng ta phải cải tiến mô hình đã nêu để làm tăng tốc độ hội tụ. Mô hình phối hợp giữa phương pháp dò tìm ngẫu nhiên với phương pháp biến phân địa phương là một hướng đang được nghiên cứu trong việc cải tiến mô hình. Cùng với hướng này chúng tôi sẽ đề nghị một hướng cải tiến khác đó là mô hình phối hợp giữa phương pháp dò tìm ngẫu nhiên với phương pháp gradient ngẫu nhiên.

Nhằm mục đích kể trên, trong bài này một loại ước lượng không chệch của véc tơ gradient được thiết lập trên cơ sở các kết quả của mô hình ngẫu nhiên tính tổng của chuỗi và giới hạn của dãy số.

2. MÔ HÌNH NGẪU NHIÊN TÍNH TỔNG CỦA CHUỖI VÀ GIỚI HẠN CỦA DÃY SỐ

2.1. Xét một chuỗi số hội tụ có tổng là s:

$$\sum_{i=0}^{\infty} s_i = s. \tag{1}$$

Giả sử tồn tại dãy số $\{q_i\}_{i\geq 0}$, sao cho:

$$\sum_{i=0}^{\infty} q_i = 1, \quad q_i > 0 \quad (\forall i \geq 0), \tag{2}$$

$$|s_i| \leq c q_i \ \text{ với } \ c = \text{const} > 0. \tag{3}$$

Với những điều kiện này rõ ràng chuỗi (1) là hội tụ tuyệt đối.

* Công trình được sự hỗ trợ của đề tài KT 04-115 thuộc chương trình Nghiên cứu Cơ bản Nhà nước

Gọi $\nu \in \{0, 1, 2, \dots\}$ là đại lượng ngẫu nhiên rời rạc với phân bố xác suất:

$$P\{\nu = i\} = q_i \quad (\forall i \geq 0). \tag{4}$$

Gọi $\xi \in [0, 2c]$ là đại lượng ngẫu nhiên phân bố đều với mật độ xác suất:

$$p(x) = \frac{1}{2c}\chi_{[0,2c]}(x), \tag{5}$$

trong đó $\chi_{[0,2c]}(x)$ là hàm đặc trưng (chỉ thị) của tập $[0, 2c]$.

Gắn với các đại lương ngẫu nhiên ν, ξ được tạo ra trên máy tính (xem [2]), ta lập đại lượng ngẫu nhiên rời rạc $\eta = \eta(\nu, \xi)$ theo công thức sau:

$$\eta = \begin{cases} c & \text{khi } \xi < \frac{s_\nu}{q_\nu} + c \\ -c & \text{khi } \xi \geq \frac{s_\nu}{q_\nu} + c \end{cases} \tag{6}$$

Bổ đề 1. *Với giả thiết* (2), (3) *đại lượng ngẫu nhiên* η *có kỳ vọng và phương sai hữu hạn:*

$$E\{\eta\} = \sum_{i=0}^{\infty} s_i = s, \tag{7}$$

$$D\{\eta\} = c^2 - s^2. \tag{8}$$

Chứng minh. Từ (4), (6) và công thức tính kỳ vọng có điều kiện ta có:

$$E\{\eta\} = E\{E(\eta/\nu)\} = \sum_{i=0}^{\infty} P\{\nu = i\}E\{\eta/\nu = i\}$$

$$= \sum_{i=0}^{\infty} q_i[cP\{\xi < \frac{s_i}{q_i} + c\} - cP\{\xi \geq \frac{s_i}{q_i} + c\}] \tag{9}$$

Từ (3) ta suy ra $0 \leq \frac{s_i}{q_i} + c \leq 2c$, do đó dựa vào tính phân bố đều của ξ ta thu được:

$$P\{\xi < \frac{s_i}{q_i} + c\} = \int_0^{\frac{s_i}{q_i}+c} \frac{1}{2c}dx = \frac{1}{2c}\Big(\frac{s_i}{q_i} + c\Big), \tag{10}$$

$$P\{\xi \geq \frac{s_i}{q_i} + c\} = \int_{\frac{s_i}{q_i}+c}^{2c} \frac{1}{2c}dx = \frac{1}{2c}\Big(c - \frac{s_i}{q_i}\Big). \tag{11}$$

Thay (10), (11) vào (9) ta thu được (7):

$$E\{\eta\} = \sum_{i=0}^{\infty} q_i\Big[c\frac{1}{2c}\Big(\frac{s_i}{q_i} + c\Big) - c\frac{1}{2c}\Big(c - \frac{s_i}{q_i}\Big)\Big] = \sum_{i=0}^{\infty} s_i = s < \infty$$

Để chứng minh (E8) ta tính:

$$E\{\eta^2\} = E\{E(\eta^2/\nu)\} = \sum_{i=0}^{\infty} q_i E\{\eta^2/\nu = i\}$$

$$= \sum_{i=0}^{\infty} q_i\Big[c^2 P\Big\{\xi < \frac{s_i}{q_i} + c\Big\} + c^2 P\Big\{\xi \geq \frac{s_i}{q_i} + c\Big\}\Big]$$

$$= \sum_{i=0}^{\infty} c^2 q_i\Big[P\Big\{\xi < \frac{s_i}{q_i} + c\Big\} + P\{\xi \geq \frac{s_i}{q_i} + c\}\Big]$$

$$= c^2 \sum_{i=0}^{\infty} q_i = c^2,$$

cho nên

$$D\{\eta\} = E\{\eta^2\} - \big(E\{\eta\}\big)^2 = c^2 - s^2. \qquad \square$$

Ví dụ 1. Nghiệm lại tổng của chuỗi sau:

$$S = \sum_{n=0}^{\infty} (-1)^n \frac{1}{n! 2^n} = e^{-1/2} \approx 0,606$$

Ta chọn ν là đại lượng ngẫu nhiên rời rạc có phân bố Poisson: $q_n = e^{-\lambda} \frac{\lambda^n}{n!}$ và chọn $c = e^{\lambda} \geq \frac{e^{\lambda}}{(2\lambda)^n}$ với $\lambda \geq 0,5$. Bây giờ ta phải so sánh $\xi \in [0,\ 2e^{\lambda}]$ với đại lượng $\frac{s_\nu}{q_\nu} + e^{\lambda}$. Sau khi rút gọn biểu thức ta chỉ phải so sánh $2\xi_1$ với đại lượng $1 + \frac{(-1)^\nu}{(2\lambda)^\nu}$, trong đó ξ_1 phân bố đều trên $[0,\ 1]$. Kết quả tính trên máy với $\lambda = 0,8$. (Xem trong bảng 1, cột "tổng của chuỗi").

Bảng 1. Kết quả tính trên máy

Số lần lặp	Tổng của chuỗi		Giới hạn của dãy		Tổng chuỗi Fourier	
	Kết quả	Sai số	Kết quả	Sai số	Kết quả	Sai số
2560	0,606	0,004	0,448	0,115	0,890	0,010
3840	0,618	0,012	0,311	0,022	0,897	0,003
5120	0,634	0,028	0,398	0,065	0,899	0,001
6400	0,649	0,043	0,307	0,026	0,920	0,020
7680	0,554	0,052	0,384	0,051	0,883	0,017
8960	0,616	0,010	0,435	0,120	0,873	0,027
10240	0,592	0,014	0,386	0,053	0,922	0,022
11520	0,603	0,003	0,348	0,015	0,898	0,002
12800	0,576	0,030	0,349	0,016	0,911	0,011
14080	0,601	0,005	0,315	0,018	0,909	0,009
15360	0,600	0,006	0,256	0,077	0,919	0,019
16640	0,616	0,010	0,348	0,015	0,880	0,020
17920	0,602	0,004	0,319	0,014	0,907	0,007
19200	0,627	0,021	0,386	0,053	0,874	0,026
20480	0,620	0,014	0,336	0,003	0,885	0,015
21760	0,602	0,004	0,353	0,020	0,910	0,010
23040	0,592	0,014	0,338	0,005	0,919	0,019
24320	0,610	0,004	0,323	0,010	0,897	0,003
25600	0,600	0,006	0,340	0,007	0,893	0,007

Ví dụ 2. Tính tổng của chuỗi

$$S = \frac{8}{\pi^2} \sum_{n=0}^{\infty} (-1)^n \frac{\sin \frac{2n+1}{2}\pi x}{(2n+1)^2} \qquad (0 \leq x \leq 2).$$

Chuỗi này chính là khai triển Fourier của hàm số:

$$f(x) = \begin{cases} x & \text{khi } 0 \le x \le 1 \\ 2 - x & \text{khi } 1 < x \le 2 \end{cases}$$

Khi chọn $q_n = \dfrac{1}{(n+1)(n+2)}$ ta có $\dfrac{|s_n|}{q_n} \le \dfrac{8}{\pi^2 q_n (2n+1)^2} < 2$ $(\forall n \ge 0)$, vậy có thể chọn $c = 2$. Kết quả tính trên máy ứng với $x = 1,1$. (Xem trong bảng 1, cột "tổng của chuỗi Fourier").

2.2. Xét một dãy số hội tụ $\{f_n\}_{n \ge 0}$

$$\lim_{n\to\infty} f_n = f.$$

Giả thiết rằng tồn tại một hằng số $c > 0$ và một dãy số $\{q_i\}_{i \ge 0}$, sao cho:

$$|f_i - f_{i-1}| \le c q_i \ (\forall i \ge 1); \quad |f_0| \le c q_0, \tag{13}$$

$$q_i > 0 \ (\forall i \ge 0); \quad \sum_{i=0}^{\infty} q_i = 1. \tag{14}$$

Gắn với các đại lượng ngẫu nhiên ξ, ν đã nêu, ta lập đại lượng ngẫu nhiên:

$$\varsigma = \begin{cases} c & \text{khi } \xi < \frac{f_\nu - f_{\nu-1}}{q_\nu} + c \\ -c & \text{khi } \xi \ge \frac{f_\nu - f_{\nu-1}}{q_\nu} + c \end{cases} \tag{15}$$

Bổ đề 2. *Giả sử các giả thiết* (3), (14) *được thỏa mãn. Khi đó giới hạn* (12) *tồn tại hữu hạn và đại lượng ngẫu nhiên* ς *có kỳ vọng và phương sai hữu hạn:*

$$E\{\varsigma\} = \lim_{n\to\infty} f_n = f, \tag{16}$$

$$D\{\varsigma\} = c^2 - f^2. \tag{17}$$

Chứng minh. Xét chuỗi $\sum_{n=0}^{\infty} s_n$, trong đó:

$$s_n := f_n - f_{n-1} \ (n \ge 1); \quad s_0 := f_0. \tag{18}$$

Từ các giả thiết (13), (14) ta suy ra các điều kiện dạng (2), (3) đối với chuỗi $\sum_{n=0}^{\infty} s_n$ được thỏa mãn, do đó chuỗi này hội tụ (tuyệt đối). Đồng thời từ (18) ta có:

$$\sum_{n=0}^{\infty} s_n = \lim_{n\to\infty}(s_0 + \cdots + s_n) = \lim_{n\to\infty} f_n = f. \tag{19}$$

Mặt khác, dựa vào (18), (15) ta có:

$$\varsigma = \begin{cases} c & \text{khi } \xi < \frac{s_\nu}{q_\nu} + c \\ -c & \text{khi } \xi \ge \frac{s_\nu}{q_\nu} + c \end{cases}$$

nghĩa là đại lượng ngẫu nhiên ς có dạng η trong (6) còn các điều kiện (13), (14) có dạng của điều kiện (3), (2) trong Bổ đề L1. Sử dụng bổ đề này đối với đại lượng ngẫu nhiên ς và (19) ta thu được (16), (17). □

Ví dụ 3. Nghiệm lại giới hạn của dãy

$$f_n = \frac{1 + 2^2 + 3^2 + \cdots + n^2}{n^3} \to \frac{1}{3} \quad (n \to \infty).$$

Ta đặt $f_{-1} = 0$, $f_0 = \alpha$, α là một số tùy ý cho trước, thì giới hạn trên chuyển thành tổng của chuỗi $\sum_{n=0}^{\infty} s_n$ khi ta đặt $s_n = f_n - f_{n-1}$. Như vậy ta có:

$$s_0 = \alpha, \ s_1 = 1 - \alpha, \text{ và với } n \ge 2 \text{ thì } s_n = \frac{-3n^2 + n + 1}{6n^2(n-1)^2}.$$

Ta chọn ν là đại lượng ngẫu nhiên rời rạc có phân bố: $q_n = \dfrac{1}{(n+1)(n+2)}$. Số c cần tìm là max của các số trong tập hợp sau :

$$\Big\{ \frac{|s_0|}{q_0} = 2|\alpha|;\ \frac{|s_1|}{q_1} = 6|1-\alpha|;\ \frac{|-3n^2+n+1|(n+1)(n+2)}{6n^2(n-1)^2},\ (n \geq 2) \Big\}$$

Kết quả tính trên máy ứng với $\alpha = 1$, $c = 4,5$. (Xem trong bảng 1, cột "giới hạn của dãy").

3. MÔ HÌNH NGẪU NHIÊN TÍNH GRADIENT CỦA MỘT HÀM

Xét hàm $f : \mathbf{G}(x)b \to \mathbf{R}^1$, $\mathbf{G}(x) \subset \mathbf{R}^m$ là lân cận lồi và mở của điểm x. Giả sử trên $\mathbf{G}(x)$ hàm f khả vi liên tục theo Lipschitz cấp $\alpha(x)$

$$\left| \frac{\partial f(x^1)}{\partial x_i} - \frac{\partial f(x^2)}{\partial x_i} \right| \leq c(x)\|x^1 - x^2\|^{\alpha(x)} \tag{20}$$

$$(\forall\, x^1, x^2 \in \mathbf{G}(x);\ i = 1 \div m);\ c(x) > 0,\ \alpha(x) > 0.$$

Chọn hai dãy số đơn điệu giảm $\{q_n\}_{n\geq 0}$, $\{\delta_n\}_{n\geq 0}$ thỏa mãn các điều kiện:

$$\sum_{i=0}^{\infty} q_i = 1;\quad 0 < \delta_n \leq \frac{1}{2} q_{n+1}^{\frac{1}{\alpha(x)}}\ (\forall n \geq 0);\quad q_0 > 0. \tag{21}$$

Gọi $\overline{\delta}_n = (-1)^n \delta_n$ và với mỗi $i = 1 \div m$ ta đặt: $f_i^{(-1)}(x) \equiv 0$, còn $f_i^{(0)}$ là số chọn tùy ý sao cho:

$$|f_i^{(0)}(x)| < c(x) q_0, \tag{22}$$

trường hợp còn lại:

$$f_i^{(n)}(x) = \frac{1}{\overline{\delta}_n}\Big[f(x + \overline{\delta}_n e_i) - f(x)\Big] \quad (\forall n \geq 1). \tag{23}$$

Ở đây e_i là véctơ chỉ phương thứ i trong $\mathbf{R}^m$. Lưu ý rằng do tính mở của $\mathbf{G}(x)$ nên ta có thể chọn δ_0 đủ bé sao cho:

$$x \pm \delta_0 e_i \in \mathbf{G}(x) \quad (\forall i = 1 \div m). \tag{24}$$

Dựa vào các dãy $\{f_i^{(n)}(x)\}$ đã xây dựng, ta có thể thiết lập đồng thời các thành phần $\varsigma_i(x)$ $(i = 1 \div m)$ của véc tơ ngẫu nhiên $\varsigma(x) = (\varsigma_1(x), \ldots, \varsigma_m(x))$ theo công thức sau:

$$\varsigma_i(x) = \begin{cases} c(x) & \text{khi } \xi < \frac{1}{q_\nu}\big(f_i^{(\nu)} - f_i^{(\nu-1)}\big) + c(x) \\ -c(x) & \text{khi } \xi \geq \frac{1}{q_\nu}\big(f_i^{(\nu)} - f_i^{(\nu-1)}\big) + c(x) \end{cases} \tag{25}$$

trong đó ν, ξ là hai đại lương ngẫu nhiên độc lập với phân bố xác sất như đã nói ở (4) và (5).

Định ly 1. *Hàm $f(x)$ với giả thiết* (20) *cùng với các thiết kế* (21), (22), (23), (24) và (25) *ta có:*

$$E\{\varsigma_i(x)\} = \frac{\partial f(x)}{\partial x_i}, \tag{26}$$

$$D\{\varsigma_i(x)\} = c^2(x) - \Big(\frac{\partial f(x)}{\partial x_i}\Big)^2. \tag{27}$$

Chứng minh. Áp dụng công thức số gia giới nội vào (23) ta có:

$$f_i^{(n)}(x) = \frac{1}{\overline{\delta}_n}\Big[f(x + \overline{\delta}_n e_i) - f(x)\Big] = \frac{\partial f\big(x + \theta_i^{(n)}(x)\overline{\delta}_n e_i\big)}{\partial x_i}, \tag{28}$$

trong đó: $0 < \theta_i^{(n)}(x) < 1$. Từ tính không tăng của dãy $\{\delta_n\}_{n\geq 0}$ và tính lồi của $\mathbf{G}(x)$ ta có thể dựa vào (24) suy ra:

$$x + \overline{\delta}_n e_i \in [x - \delta_0 e_i,\ x + \delta_0 e_i] \subset \mathbf{G}(x).$$

Do tính lồi của $\mathbf{G}(x)$ ta còn có: $x + \theta_i^{(n)} \overline{\delta}_n e_i \in \mathbf{G}(x)$.

Trên cơ sở này, từ (28), (20) ta suy ra:

$$\begin{aligned}\left|f_i^{(n)}(x) - f_i^{(n-1)}(x)\right| &\le c(x)\left\|\theta_i^{(n)}(x)\overline{\delta}_n e_i - \theta_i^{(n-1)}(x)\overline{\delta}_{n-1} e_i\right\|^{\alpha(x)}\\ &= c(x)\left|\theta_i^{(n)}(x)\delta_n + \theta_i^{(n-1)}(x)\delta_{n-1}\right|^{\alpha(x)}\\ &= c(x)\left(\theta_i^{(n)}(x)\delta_n + \theta_i^{(n-1)}(x)\delta_{n-1}\right)^{\alpha(x)}.\end{aligned}$$

Khi đó từ (21) ta suy ra:

$$\left|f_i^{(n)}(x) - f_i^{(n-1)}(x)\right| \le c(x)(\delta_n + \delta_{n-1})^{\alpha(x)} \le c(x)(2\delta_{n-1})^{\alpha(x)} \le c(x)q_n.$$

Khi kết hợp điều kiện này với (22) và (2) ta nhận thấy giả thiết của Bổ đề 2 được thỏa mãn đối với bài toán giới hạn của dãy số: $\{f_i^{(n)}(x)\}_{n\ge 0}$ $(i = 1 \div m)$.

Áp dụng Bổ đề 2 ta thu được:

$$E\{\varsigma_i(x)\} = \lim_{n\to\infty} f_i^{(n)}(x), \tag{29}$$

$$D\{\varsigma_i(x)\} = c^2 - \left(\lim_{n\to\infty} f_i^{(n)}(x)\right)^2. \tag{30}$$

Mặt khác từ (21) dễ dàng nhận thấy rằng

$$0 \le \lim_{n\to\infty} \delta_n \le \frac{1}{2}\lim_{n\to\infty} q_n^{\frac{1}{\alpha(x)}} = 0,$$

nghĩa là:

$$\lim_{n\to\infty} \overline{\delta}_n = \lim_{n\to\infty} \delta_n = 0. \tag{31}$$

Do sự tồn tại của các đạo hàm riêng $\frac{\partial f(x)}{\partial x_i}$ $(\forall x \in \mathbf{G}(x),\ 1 \le i \le m)$ nên từ (31), (23) ta suy ra

$$\lim_{n\to\infty} f_i^{(n)}(x) = \lim_{n\to\infty} \frac{1}{\overline{\delta}_n}\Big(f(x + \overline{\delta}_n e_i) - f(x)\Big) = \frac{\partial f(x)}{\partial x_i}. \tag{32}$$

Từ (29) và (32) ta thu được (26) còn (27) cũng thu được từ (30) và (32). □

TÀI LIỆU THAM KHẢO

[1] Trần Cảnh, Phương pháp dò tìm ngẫu nhiên giải một loại bài toán điều khiển, *Tuyển tập Các công trình khoa học* (ngành Toán), HNKH Trường ĐH Khoa học tự nhiên, Hà nội, 1998, tr. 25–40.

[2] Sobol I. M., *Các phương pháp tính toán Monte-Carlo*, FML Moskva, 1973 (tiếng Nga).

[3] Zielinski R., Neumann P., *Stochastysczne Metody Poszukiwania Minimum Funkcij*, WNT Warszawa, 1986.

Nhận bài ngày 10 tháng 4 năm 2000

Nhận bài sau khi sửa ngày 21 tháng 3 năm 2001

Bộ môn Toán

Trường Đại học Xây dựng Hà Nội