Phương pháp lai ghép đường dốc nhất giải bài toán bất đẳng thức biến phân

GV hướng dẫn: PGS.TS. Nguyễn Thị Thu Thủy

SVTH: Vũ Thị Tâm

Viện Toán ứng dụng và Tin học
Trường Đại học Bách khoa Hà Nội

Ngày 5 tháng 3 năm 2022

Nội dung chính

Nội dung

Tóm tắt nội dung chính

1) Giới thiệu bài toán bất đẳng thức biến phân (VIP); bài toán giải phương trình toán tử và bài toán điểm bất động
2) Trình bày phương pháp lai ghép đường dốc nhất xấp xỉ nghiệm bài toán bất đẳng thức biến phân VIP
3) Đưa ra ví dụ số minh họa

Ánh xạ không giãn

Định nghĩa

Cho C *là một tập con khác rỗng của không gian Hilbert thực* H. *Ánh xạ* $T : C \to H$ *được gọi là*

(i) *Ánh xạ* L*-liên tục Lipschitz trên* C *nếu tồn tại hằng số* $L \geq 0$ *sao cho*

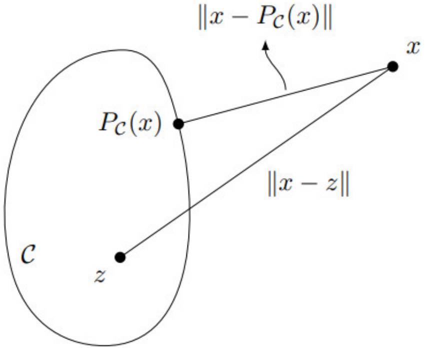
$$\|T(x) - T(y)\| \leq L\|x - y\| \quad \forall x, y \in C. \tag{1}$$

(ii) *Trong* (1), *nếu* $L \in [0, 1)$ *thì* T *được gọi là ánh xạ co; nếu* $L = 1$ *thì* T *được gọi là ánh xạ không giãn.*

Toán tử chiếu trong không gian Hilbert

Cho C là một tập con lồi đóng khác rỗng trong không gian Hilbert thực H. Phép chiếu mêtric chiếu H lên C được định nghĩa như sau: với mỗi $x \in H$ tồn tại duy nhất một phần tử $P_C x \in C$ thỏa mãn

$$\|x - P_C x\| \leq \|x - y\| \quad \forall y \in C. \tag{2}$$



Hình: Minh họa về phép chiếu metric

Bổ đề

Giả sử P_C là phép chiếu mêtric chiếu không gian Hilbert thực H lên một tập con lồi, đóng, khác rỗng C của H. Khi đó các kết luận sau là đúng:

- $z = P_C x$ *nếu và chỉ nếu bất đẳng thức sau thỏa mãn*

$$\langle x - z, y - z \rangle \leq 0, \quad x \in H, \ y \in C \tag{3}$$

- *Với mọi* $x, y \in H$,

$$\|P_C x - P_C y\|^2 \leq \langle P_C x - P_C y, x - y \rangle . \tag{4}$$

Đặc biệt P_C là ánh xạ không giãn, nghĩa là

$$\|P_C x - P_C y\| \leq \|x - y\| \quad \forall x, y \in H \tag{5}$$

- *Với mọi* $x, y \in H$ *ta có*

$$\| (I - P_C x) x - (I - P_C y) y\|^2 \leq \langle (I - P_C) x - (I - P_C) y, \ x - y \rangle . \tag{6}$$

Toán tử đơn điệu

Cho C là một tập con lồi đóng khác rỗng trong không gian Hilbert thực H, $T : C \to H$ là một toán tử từ C vào H. Toán tử A được gọi là:

(i) Toán tử đơn điệu nếu

$$\langle Tx - Ty, x - y\rangle \geq 0 \quad \forall x, y \in C \tag{8}$$

(ii) Toán tử η-đơn điệu mạnh, nếu tồn tại một hằng số $\eta > 0$ sao cho

$$\langle Tx - Ty, x - y\rangle \geq \eta\|x - y\|^2 \quad \forall x, y \in C \tag{9}$$

(iii) Toán tử đơn điệu mạnh ngược trên C với hệ số $\beta > 0$ (hay β-đơn điệu mạnh ngược trên C) nếu

$$\langle Tx - Ty, x - y\rangle \geq \beta\|Tx - Ty\|^2 \quad \forall x, y \in C.$$

Bài toán bất đẳng thức biến phân

Cho C là một tập con lồi đóng khác rỗng trong không gian Hilbert thực H và ánh xạ $F : C \to H$. Bài toán bất đẳng thức biến phân xác định bởi miền ràng buộc C và ánh xạ F (thường gọi là ánh xạ giá hay ánh xạ mục tiêu), ký hiệu VI(F, C), là bài toán:

$$\text{Tìm } x^* \in C \quad \text{sao cho} \quad \langle F(x^*), x - x^* \rangle \geq 0 \quad \forall x \in C. \tag{10}$$

Tập nghiệm của bài toán bất đẳng thức biến phân (10) với ánh xạ giá F và tập ràng buộc C được ký hiệu là $\Omega_{(F,C)}$.

Bài toán giải phương trình toán tử

Mệnh đề

Nếu $H = \mathbb{R}^n$, $C = \mathbb{R}^n$ và ánh xạ $F : \mathbb{R}^n \to \mathbb{R}^n$ thì $x^* \in \mathbb{R}^n$ là nghiệm của bài toán bất đẳng thức biến phân $\mathrm{VI}(F, C)$ khi và chỉ khi x^* là nghiệm của phương trình toán tử $F(x^*) = 0$.

Bài toán điểm bất động

Cho C là tập con khác rỗng của không gian Hilbert thực H và ánh xạ $T : C \to C$. Điểm $x \in C$ được gọi là điểm bất động của ánh xạ T nếu $T(x) = x$.

Ký hiệu tập điểm bất động của ánh xạ T là Fix(T), nghĩa là

$$\mathsf{Fix}(T) := \{x \in C : \quad T(x) = x\}.$$

Bổ đề

Cho C là một tập con lồi đóng khác rỗng của không gian Hilbert thực H và ánh xạ $T : C \to C$. Cho $x^* \in C$ và $\lambda > 0$. Khi đó, $x^* \in \Omega_{FC}$ khi và chỉ khi $x^* \in \mathrm{Fix}(T)$, ở đây $T(x) = P_C(x - \lambda F(x))$.

Điều kiện tồn tại duy nhất nghiệm

Điều kiện tồn tại duy nhất nghiệm

Nếu F là ánh xạ η-đơn điệu mạnh và L-liên tục Lipschitz ánh xạ vào C thì VI(F, C) có nghiệm duy nhất.

Phương pháp lai ghép đường dốc nhất

Điều kiện ban đầu

Giả sử C là tập điểm bất động của ánh xạ không giãn $T : H \to H$. Giả sử F là ánh xạ η-đơn điệu mạnh và L-liên tục Lipschitz trên C thỏa mãn:

$$\|Fx - Fy\| \leq L\|x - y\|, \quad x, y \in C, \tag{11}$$

$$\langle Fx - Fy, x - y\rangle \geq \eta\|x - y\|^2, \quad x, y \in C. \tag{12}$$

Lấy một số cố định $\mu \in \left(0, 2\eta/L^2\right)$ và một dãy số thực $\{\lambda_n\}$ trong khoảng $(0, 1)$ thỏa mãn các điều kiện dưới đây:

(L1) $\lim_{n\to\infty} \lambda_n = 0$;

(L2) $\sum_{n\to\infty} \lambda_n = \infty$;

(L3) $\lim_{n\to\infty}(\lambda_n - \lambda_{n+1})/\lambda_{n+1}^2 = 0$.

Để giảm phức tạp, ta thay điều kiện (L3) và (L4) bởi:

(L3)′ $\lim_{n\to\infty} \dfrac{\lambda_n}{\lambda_{n+1}} = 1$ hoặc tương đương $\lim_{n\to\infty} \dfrac{\lambda_n - \lambda_{n+1}}{\lambda_{n+1}} = 0$

(L4)′ $\lim_{n\to\infty} \dfrac{\lambda_n}{\lambda_{n+N}} = 1$ hoặc tương đương $\lim_{n\to\infty} \dfrac{\lambda_n - \lambda_{n+N}}{\lambda_{n+N}} = 0.$

Công thức lặp xấp xỉ nghiệm

Giả sử rằng C là tập điểm bất động của ánh xạ không giãn $T : H \to H$. Cho $u_0 \in H$ tùy ý và gọi $\{\lambda_n\}$ là một dãy trong $[0,1]$. Phương pháp lai ghép đường dốc nhất được Yamada giới thiệu là công thức

$$u_{n+1} := T^{\lambda_{n+1}} u_n = T u_n - \lambda_{n+1} \mu F\left(T u_n\right), \quad n \geq 0 \tag{13}$$

Trong trường hợp C là giao của các tập hợp điểm bất động của N ánh xạ không giãn T_i: $H \to H$

$$C = \bigcap_{i=1}^{N} \mathrm{Fix}\left(T_i\right),$$

với $N \geq 1$ nguyên, dãy $\{u_n\}$ có thể được viết lại ở dạng

$$u_{n+1} = T_{[n+1]}^{\lambda_{n+1}} u_n = T_{[n+1]} u_n - \lambda_{n+1} \mu F\left(T_{[n+1]} u_n\right), \quad n \geq 0. \tag{14}$$

trong đó $T_{[k]} := T_{k \bmod N}$, trong đó số nguyên $k \geq 1$, với hàm mod nhận các giá trị trong tập $\{1, 2, ..., N\}$

Sự hội tụ nghiệm của phương pháp

Định lý hội tụ

Giả sử rằng $0 < \mu < 2\eta/L^2$ và các điều kiện $(\mathrm{L1})$, $(\mathrm{L2})$, $(\mathrm{L3})'$ đặt lên dãy tham số $\{\lambda_n\}$ được thỏa mãn. Khi đó, dãy $\{u_n\}$ sinh ra bởi thuật toán (13) hội tụ mạnh về nghiệm duy nhất u^* của bài toán $\mathrm{VI}(F, C)$.

Sự hội tụ nghiệm của phương pháp

Sự hội tụ mạnh của dãy lặp (14) được nêu trong định lý sau.

Định lý

Cho $\mu \in \left(0, 2\eta/L^2\right)$ *và các điều kiện* $(\mathrm{L1})$, $(\mathrm{L2})$, $(\mathrm{L4})'$ *thỏa mãn. Giả sử thêm rằng*

$$\begin{aligned} C = \bigcap_{n=1}^{N} \mathrm{Fix}\,(T_i) &= \mathrm{Fix}\,(T_1 T_2 \cdots T_N) \\ &= \mathrm{Fix}\,(T_N T_1 \cdots T_{N-1}) \\ &= \cdots = \mathrm{Fix}\,(T_2 T_3 \cdots T_N T_1)\,. \end{aligned} \tag{15}$$

Khi đó, dãy $\{u_n\}$ *được tạo bởi thuật toán* (14) *hội tụ mạnh (hội tụ theo chuẩn) tới nghiệm duy nhất* u^* *của bài toán* $\mathrm{VI}(F, C)$.

Ví dụ số minh họa

Ví dụ

Xét $H = \mathbb{R}^3$, ánh xạ giá F, ánh xạ không giãn T và tập ràng buộc C được cho như sau:

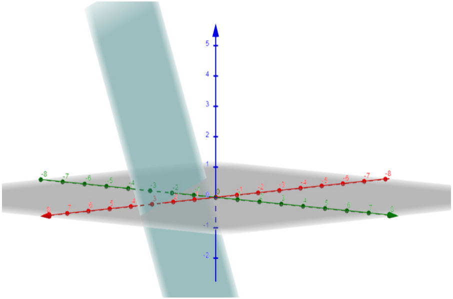
- $F : \mathbb{R}^3 \to \mathbb{R}^3$ xác định bởi $Fx = \dfrac{2}{5}Mx$ với $M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$;
- $T : \mathbb{R}^3 \to \mathbb{R}^3$ xác định bởi $Tx = P_C x$;
- $C = \{x \in \mathbb{R}^3 \mid 2x_1 - 3x_2 - x_3 \leq 5\} \subset \mathbb{R}^3$.

Giải

Trước hết ta tìm nghiệm của bài toán bất đẳng thức biến phân VI(F, C). Gọi x^* là nghiệm của bài toán VI(F, C),

$$\begin{aligned}
&\langle Fx^*, x - x^* \rangle \geq 0 \quad \forall x \in C \\
\Leftrightarrow\ &\left\langle \frac{2}{5}x^*, x - x^* \right\rangle \geq 0 \quad \forall x \in C \\
\Leftrightarrow\ &\langle x^*, x - x^* \rangle \geq 0 \quad \forall x \in C \\
\Leftrightarrow\ &\langle x^*, x^* \rangle \leq \langle x^*, x \rangle \quad \forall x \in C \\
\Leftrightarrow\ &\|x^*\|^2 \leq \langle x^*,\ x \rangle \leq \|x^*\| \,.\, \|x\| \ \text{(BĐT Cauchy – Schwart)} \quad \forall x \in C \\
\Leftrightarrow\ &\|x^*\| \leq \|x\| \quad \forall x \in C.
\end{aligned}$$

Vậy nghiệm của bài toán VI(F, C) trong trường hợp này là điểm x^* có chuẩn nhỏ nhất trong C.



Hình: Hình ảnh siêu phẳng $2x_1 - 3x_2 - x_3 = 5$ trong không gian ba chiều

Tập C là nửa không gian chứa điểm gốc tọa độ $(0,0,0)$ có mặt phẳng tựa là mặt phẳng $2x_1 - 3x_2 - x_3 = 5$ như trên hình. Do C chứa điểm $x = (0,0,0)$ nên hiển nhiên $x^* = (0,0,0)$ là nghiệm của bài toán VI(F, C).

Giải nghiệm xấp xỉ bằng phương pháp HSDM

Kiểm tra điều kiện

Trước hết, ta kiểm tra các điều kiện đặt lên ánh xạ giá F và ánh xạ không giãn T:

- F là ánh xạ η-đơn điệu mạnh trên C và L-liên tục Lipschitz C với $\eta = \dfrac{2}{5}$ và $L = 0.692820323027551$ (sử dụng Python với hàm `np.linalg.norm()`);
- Vì $Tx = P_C x$ nên T là ánh xạ không giãn trên C.

Chọn $\lambda_n = \frac{1}{n}$, dễ thấy dãy $\{1/n\}$ thỏa mãn các điều kiện (L1), (L2) và (L3)$'$; với η và L đã được xác định, ta chọn $\mu \in (0, 2\eta/L^2)$ tùy ý.

Thực hiện lặp xấp xỉ nghiệm

Ta tính toán dãy lặp (13) với tiêu chuẩn dừng là sai số giữa nghiệm xấp xỉ và nghiệm đúng nhỏ hơn $\varepsilon > 0$ cho trước với $\varepsilon = 10^{-6}$.

Lần lặp k	x_1^k	x_2^k	x_3^k	$\|x^k - x^*\|$
1	0.2448	0.4896	0.7344	0.9159577282822605
2	0.06932736	0.13865472	0.20798208	0.25939922864953613
3	0.02096459	0.04192919	0.06289378	0.07844232674361971
4	0.00658121	0.01316241	0.01974362	0.024624615211357097
5	0.00211652	0.00423303	0.00634955	0.00791927625197244
...	...	...	...	...
13	$3.51031183 \times 10^{-7}$	$7.02062366 \times 10^{-7}$	$1.05309355 \times 10^{-6}$	$1.313438418153186 8\times 10^{-6}$
14	$1.20979387 \times 10^{-7}$	$2.41958774 \times 10^{-7}$	$3.62938161 \times 10^{-7}$	$4.52663416432314 2 \times 10^{-7}$

Hình: Kết quả tính toán dãy lặp (13) với $x^0 = (1, 2, 3)^\top$, $\lambda = 1/n$, $\mu = 1.6$.

Kết quả các lần lặp

Lần lặp k	x_1^k	x_2^k	x_3^k	$\|x^k - x^*\|$
1	0.48	0.96	1.44	1.7959955456514918
2	0.2496	0.4992	0.7488	0.9339176837387758
3	0.134784	0.269568	0.404352	0.5043155492189388
4	0.07440077	0.14880154	0.2232023	0.27838218316885427
5	0.04166443	0.08332886	0.12499329	0.15589402257455837
...	...	...	..	...
27	$3.00787358 \times 10^{-7}$	$6.01574715 \times 10^{-7}$	$9.02362073 \times 10^{-7}$	$1.1254432386003892 \times 10^{-6}$
28	1.7798314×10^{-7}	3.5596628×10^{-7}	5.3394942×10^{-7}	$6.65951930151127 \times 10^{-7}$

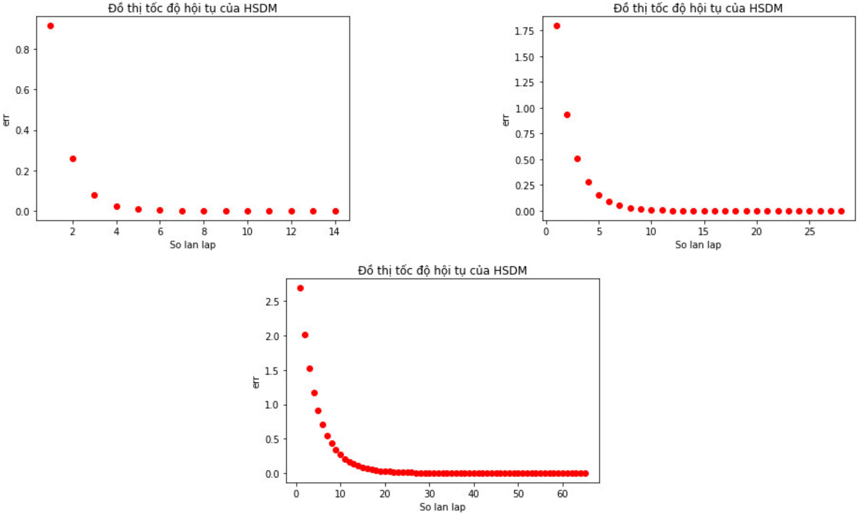
Hình: Kết quả tính toán dãy lặp (13) với $x^0 = (1, 2, 3)^\top$, $\lambda = 1/n$, $\mu = 1$.

Kết quả các lần lặp

Lần lặp k	x_1^k	x_2^k	x_3^k	$\|x^k - x^*\|$
1	0.72	1.44	2.16	2.6939933184772378
2	0.5376	1.0752	1.6128	2.011515011129671
3	0.408576	0.817152	1.225728	1.5287514084585498
4	0.31378637	0.62757274	0.9413591	1.1740810816961662
5	0.24266146	0.48532292	0.72798437	0.9079560365117019
...	...	...	...	...
64	$2.92088285 \times 10^{-7}$	$5.84176571 \times 10^{-7}$	$8.76264856 \times 10^{-7}$	$1.092894290551161 \times 10^{-6}$
65	$2.32962535 \times 10^{-7}$	$4.65925071 \times 10^{-7}$	$6.98887606 \times 10^{-7}$	$8.716659917365623 \times 10^{-7}$

Hình: Kết quả tính toán dãy lặp (13) với $x^0 = (1, 2, 3)^\top$, $\lambda = 1/n$, $\mu = 0.5$.

Sự hội tụ nghiệm của phương pháp



Hình: Đồ thị tốc độ hội tụ của HSDM với lần lượt $\mu = 1.6, \mu = 1, \mu = 0.5$

Nhận xét

Với những giá trị μ khác nhau thì số lần lặp thực hiện thuật toán cũng khác nhau. Cụ thể là với μ càng lớn, càng gần giá trị $2\eta/L^2$ thì số lần lặp càng nhỏ và ngược lại μ càng tiến về 0 thì số lần lặp càng lớn.
Mà μ được tính dựa vào hai hệ số là η và L. Do vậy tính chất của ánh xạ F ảnh hưởng lớn đến tốc độ hội tụ của thuật toán.

Cảm ơn thầy cô và các bạn đã lắng nghe!