

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC SƯ PHẠM TP.HCM
KHOA TOÁN



LUẬN VĂN THẠC SĨ

ĐỀ TÀI:

XẤP XỈ NGHIỆM CỦA PHƯƠNG TRÌNH TOÁN TỬ VÀ PHƯƠNG PHÁP NEWTON

GVHD : TS. NGUYỄN CAM
SVTH : PHAN THÀNH ĐÔNG

TP.HCM, 2007



LỜI CẢM ƠN

Tôi xin chân thành cảm ơn thầy Nguyễn Cam người đã tận tình hướng dẫn và giúp đỡ trong suốt quá trình làm luận văn. Tôi xin chân thành cảm ơn Ban giám hiệu, Phòng tổ chức cán bộ và tổ Toán của trường Cao Đẳng Sư Phạm Long An đã tạo điều kiện thuận lợi cho tôi theo học lớp cao học. Tôi xin chân thành cảm ơn các bạn học viên trong lớp cao học khóa 15 đã hỗ trợ cho tôi trong suốt khóa học.

Tác giả luận văn

Phan Thành Đông

MỞ ĐẦU

1. LÝ DO CHỌN ĐỀ TÀI

Trong thực tế đa phần các bài toán được đưa về bài toán tìm nghiệm của một phương trình hoặc hệ phương trình. Việc tìm nghiệm chính xác của phương trình là một nhiệm vụ vô cùng khó khăn và có khi không thể thực hiện được, nhưng ta có thể tìm lời giải xấp xỉ của các phương trình này đến độ chính xác cần thiết để đáp ứng được nhu cầu thực tế. Từ những nhu cầu thực tế đó, luận văn “ Xấp xỉ nghiệm của phương trình toán tử và phương pháp Newton” nghiên cứu việc xây dựng lời giải xấp xỉ của một số phương trình và hệ phương trình.

2. MỤC ĐÍCH

Bằng các kiến thức cơ bản của giải tích hàm và đại số tuyến tính, luận văn đưa ra lời giải xấp xỉ của một số bài toán với những điều kiện cụ thể.

3. ĐỐI TƯỢNG VÀ PHẠM VI NGHIÊN CỨU

Nội dung của luận văn là giới thiệu và áp dụng phương pháp Newton để xây dựng lời giải xấp xỉ nghiệm của phương trình $f(x)=0$, trong đó f là ánh xạ đi từ E vào E , với $E = \mathbb{R}_n$ hoặc E là các không gian tuyến tính định chuẩn vô hạn chiều. Với những điều kiện thích hợp thì dãy lặp: $x^{k+1} = x^{k+1} - \gamma_k \nabla f(x^k)$; $x_{k+1} = x_k - f'_{-1}(x_k) f(x_k)$; $x^{k+1} = x^k + \gamma_k \phi(x^k)$ $x^{k+1} = x^k - \gamma_k H^{-1}(x^k) \nabla f(x^k)$, với x_0 tùy ý trong E , các dãy lặp này hội tụ về nghiệm của phương trình. Luận văn gồm ba chương:

Chương 1 dành cho việc giới thiệu phương pháp Newton và một số kiến thức cần thiết để trình bày cho các chương sau.

Chương 2 với nội dung áp dụng phương pháp Newton để trình bày cách xây dựng lời giải xấp xỉ của nghiệm của một phương trình hoặc một hệ phương trình trong không gian hữu hạn chiều.

Chương 3 dành cho việc trình bày mở rộng các kết quả trong chương 2 trên không gian định chuẩn tổng quát với các định lý của Kantorovich.

4. PHƯƠNG PHÁP NGHIÊN CỨU

Trên cơ sở nghiên cứu các kết quả trong giáo trình Constructive Real Analysis của giáo sư Allen A. Goldstein và các giáo trình giải tích hàm khác luận văn đã xây dựng được lời giải xấp xỉ của một số phương trình và hệ phương trình.

Chương 1:

GIỚI THIỆU PHƯƠNG PHÁP NEWTON

1.1. ĐẶT VẤN ĐỀ

Chúng ta xét việc tìm căn bậc hai của số dương a bằng phép tính toán lặp đơn giản, được cho bởi công thức như sau: $x_{n+1} = \frac{1}{2} \left(x_n + \frac{a}{x_n} \right)$. Công thức này là kết quả của phương pháp Newton mà ta sẽ giới thiệu ở phần sau.

Nếu x_n xấp xỉ $\sqrt{a}$ thì sai số tương đối của xấp xỉ này được cho bởi công thức $\left| \frac{x_n - \sqrt{a}}{\sqrt{a}} \right|$

Định lý

i) Giả sử a và x_0 là các số dương

ii) Ta xác định dãy $\{x_n\}$ bởi $x_{n+1} = \frac{1}{2} \left(x_n + \frac{a}{x_n} \right)$

iii) Đặt $\delta_n = \frac{x_n - \sqrt{a}}{\sqrt{a}}$. Thì

a) $\delta_{n+1} = \frac{1}{2} \left(\frac{\delta_n^2}{1 + \delta_n} \right) \quad n = 0, 1, 2, \dots$

b) $\delta_n \geq 0 \quad n = 0, 1, 2, \dots$

c) $\forall \varepsilon > 0: \frac{x_n}{a} (x_n - x_{n+1}) < \varepsilon \Rightarrow |\delta_n| < \varepsilon, \forall n \in \mathbb{N}$

Chứng minh

a) Do (iii) $x_n = \sqrt{a}(\delta_n + 1)$, dùng (ii) ta được:

$$x_{n+1} = \frac{1}{2} \left(\sqrt{a}(\delta_n + 1) + \frac{a}{\sqrt{a}(\delta_n + 1)} \right) = \sqrt{a} \left[1 + \frac{1}{2} \frac{\delta_n^2}{1 + \delta_n} \right]$$

$$\text{Cũng do (iii): } \sqrt{a}(1 + \delta_{n+1}) = \sqrt{a} \left(1 + \frac{x_{n+1} - \sqrt{a}}{\sqrt{a}} \right) = \sqrt{a} \left(\frac{x_{n+1}}{\sqrt{a}} \right) = x_{n+1}$$

$$\text{Nên ta có: } \delta_{n+1} = \frac{1}{2} \frac{\delta_n^2}{1 + \delta_n}$$

Vậy a) được chứng minh

$$\text{b) Từ iii) } \Rightarrow \delta_o = \frac{x_o - \sqrt{a}}{\sqrt{a}} \Rightarrow x_o = \sqrt{a}(\delta_o + 1)$$

$$\Rightarrow \delta_o + 1 > 0 \text{ (vì } x_o > 0, a > 0)$$

$$\Rightarrow \delta_1 = \frac{1}{2} \left(\frac{\delta_o^2}{1 + \delta_o} \right) > 0$$

$$\text{Suy ra } \delta_n \geq 0, \forall n \text{ bằng phương pháp quy nạp (vì } \delta_{n+1} = \frac{1}{2} \frac{\delta_n^2}{1 + \delta_n})$$

c) Từ ii) ta có:

$$x_{n+1} = \frac{1}{2}x_n + \frac{a}{2x_n} \Rightarrow x_n - x_{n+1} = \frac{x_n}{2} - \frac{a}{2x_n} = \frac{x_n^2 - a}{2x_n}$$

$$\Rightarrow \frac{x_n}{a}(x_n - x_{n+1}) = \frac{1}{2} \frac{x_n^2 - a}{a}$$

$$\text{Do giả thiết trong c) ta có: } \frac{x_n}{a}(x_n - x_{n+1}) < \varepsilon$$

$$\Rightarrow \frac{x_n^2 - a}{a} < 2\varepsilon < 2\varepsilon + \varepsilon^2$$

$$\text{Do đó } x_n^2 < a(1 + 2\varepsilon + \varepsilon^2) \Rightarrow x_n^2 < a(1 + \varepsilon)^2 \Rightarrow x_n < \sqrt{a}(1 + \varepsilon) \text{ nên}$$

$$\frac{x_n - \sqrt{a}}{\sqrt{a}} = \delta_n = |\delta_n| < \varepsilon \text{ với } n = 1, 2, 3, \dots; (\text{do b) nên } \delta_n = |\delta_n|)$$

1.2. PHÁP LẬP VÀ ĐIỂM BẤT ĐỘNG (iteration and fixed points)

Định nghĩa 1.2.1

Cho $I = [a; b]$ và f là hàm số liên tục trên I lấy giá trị trong I . Ta gọi $x \in I$ là điểm bất động của f nếu $f(x) = x$

Bổ đề 1.2.1

Mọi hàm liên tục f đi từ I vào chính nó luôn có một điểm bất động

Chứng minh

Nếu $a \in I$ không là điểm bất động thì $f(a) > a$ (vì $f(a) \geq a$)

Nếu $b \in I$ không là điểm bất động thì $f(b) < b$ (vì $f(b) \leq b$)

Đặt $h(x) = f(x) - x$, ta có: $h(a) = f(a) - a > 0$, $h(b) = f(b) - b < 0$

mà h liên tục nên có $z \in I$ thỏa $h(z) = 0$ hay $f(z) = z$

Định nghĩa 1.2.2

Một ánh xạ đi từ I vào chính nó gọi là ánh xạ co nếu tồn tại $0 < q < 1$ sao cho với mọi cặp điểm

$$x, y \in I \text{ thì } |f(x) - f(y)| \leq q|x - y|$$

Định lý 1.2.1

Cho f là ánh xạ co trên I. Đặt $x_{n+1} = f(x_n)$ với $x_0 \in I$ thì f có điểm bất động duy nhất z thỏa:

$$\text{dãy } x_n \rightarrow z \text{ và } |x_{n+1} - z| \leq q^{n+1}|x_0 - z|$$

Chứng minh

Do tính chất của ánh xạ co nên f là hàm liên tục từ I vào chính nó

Theo bổ đề 1.2.1 thì f có điểm bất động, ta gọi là z

Ta có:

$$|x_{n+1} - z| = |f(x_n) - f(z)| \leq q|x_n - z| \leq q|f(x_{n-1}) - f(z)| \leq q^2|x_{n-1} - z| \leq \dots \leq q^{n+1}|x_0 - z|$$
 Ta thiết lập được

$$\text{công thức: } |x_{n+1} - z| \leq q^{n+1}|x_0 - z|, \text{ với } n \geq 0$$

hơn nữa do $0 < q < 1$ nên $\lim_{n \rightarrow \infty} x_n = z$

Chứng minh sự duy nhất

Giả sử hàm số đã cho có hai điểm bất động khác nhau là z_1 và z_2

$$\text{Ta có: } 0 < |z_1 - z_2| = |f(z_1) - f(z_2)| \leq q|z_1 - z_2| < |z_1 - z_2| \text{ (mâu thuẫn)}$$

Do đó $z_1 = z_2$.

Bổ đề 1.2.2

Cho hàm số f có đạo hàm liên tục trên I và f là ánh xạ đi từ I vào chính nó.

Nếu $|f'(x)| < 1$ trên I thì f là ánh xạ co.

Chứng minh

Áp dụng định lý giá trị trung bình cho cặp x, y tùy ý thuộc I ta có:

$$|f(x) - f(y)| = |f'(\xi)||x - y| \text{ với } \xi \text{ là số nằm giữa } x \text{ và } y$$

Do $\max_{x \in I} |f'(x)| < 1$ nên f là ánh xạ co.

Giả sử h là hàm đơn điệu trên I, h có đạo hàm dương liên tục, giả sử h có nghiệm z thuộc phần trong (interior) của I thì $h(a) < 0 < h(b)$. Ta định nghĩa hàm: $F(x) = x - \lambda h(x)$ nếu F là ánh xạ đi từ I vào chính nó ta phải có $a \leq F(x) \leq b, \forall x \in I$. Nếu $\lambda > 0$ thì

$F(a) > a$ và $F(b) < b$, do đó với $\lambda > 0$, đủ nhỏ thì $a < F(x) < b, \forall x \in I$, hơn nữa bởi vì $F'(x) = 1 - \lambda h'(x)$ và $h'(x) > 0$ nên với $\lambda > 0$, đủ nhỏ thì $|F'(x)| < 1$

Định lý 1.2.2

Giả sử $h \in C^1[a, b]$, $h(a).h(b) < 0$ và tồn tại hai số μ, γ sao cho $0 < \mu \leq h'(x) \leq \frac{1}{\gamma}, \forall x \in I$

Đặt dãy: $x_{n+1} = x_n - \gamma h(x_n)$ với x_0 tùy ý thuộc I thì $x_n \rightarrow z$ (với z là nghiệm của h) và $|x_{n+1} - z| \leq (1 - \mu\gamma)^{n+1} |x_0 - z|$

Chứng minh

Đặt $F(x) = x - \gamma h(x)$, chú ý rằng z là điểm bất động của F khi và chỉ khi $h(z) = 0$

do $0 < \mu\gamma \leq \gamma h'(x) \leq 1 \Rightarrow 0 \leq 1 - \gamma h'(x) \leq 1 - \mu\gamma < 1$, với mọi $x \in I$

nên $0 \leq F'(x) \leq 1 - \mu\gamma < 1, \forall x \in [a, b]$; F là hàm đơn điệu tăng trên $[a, b]$

Do $h(a)h(b) < 0$ và h đơn điệu tăng trên $[a, b]$ nên $h(a) < 0$ và $h(b) > 0$

từ đây ta có: $F(a) > a$ và $F(b) < b$ (vì $\gamma > 0$)

do F đơn điệu tăng nên $a < F(x) < b, \forall x \in [a, b]$. Hơn nữa $|F'(x)| \leq 1 - \mu\gamma < 1$ áp dụng định lý

1.2.1 và bổ đề 1.2.2 ta được: $x_n \rightarrow z$ và $|x_{n+1} - z| \leq (1 - \mu\gamma)^{n+1} |x_0 - z|$

Chú ý rằng nghiệm z trong định lý là duy nhất bởi vì F có duy nhất điểm bất động. Nếu h là ánh xạ đơn điệu giảm thì $-h$ là ánh xạ đơn điệu tăng và có nghiệm giống như nghiệm của h .

Xét ví dụ

Cho hàm: $h(x) = x^2 - \alpha, \alpha > 0$, giả sử $\alpha \in (a^2; b^2)$ thì $h(a) < 0, h(b) > 0$, và

$$0 < 2a \leq h'(x) < 2b, \forall x \in [a, b]$$

Theo định lý 1.2.2 ở trên,

dãy $x_{n+1} = \left\{ x_n - \frac{1}{2b}(x_n^2 - \alpha) \right\}$ tiến về $\sqrt{\alpha}$ với x_0 tùy ý thuộc $[a, b]$

và ta có: $|x_{n+1} - \sqrt{\alpha}| \leq \left(1 - \frac{a}{b}\right)^{n+1} |x_0 - \sqrt{\alpha}|$.

1.3. PHƯƠNG PHÁP NEWTON

Giả sử h thỏa giả thiết của định lý 1.2.2, đặt $\lambda(x) = \frac{1}{h'(x)}$, $F(x) = x - \lambda(x)h(x)$, hơn nữa giả

sử $h \in C^2[a; b]$ ta có $F'(x) = \frac{h''(x)h(x)}{[h'(x)]^2}$.

Phép lặp $x_{n+1} = F(x_n) = x_n - \frac{h(x_n)}{h'(x_n)}$ được gọi là phương pháp Newton.

Theo định lý 1.2.1 và bổ đề 1.2.2, ta có sự hội tụ của dãy $\{x_n\}$ với điều kiện $|F'(x)| \leq q < 1, \forall x \in [a; b]$ và F là ánh xạ đi từ $[a; b]$ vào chính nó.

Gọi z là điểm bất động của F và viết

$$x_n - z = F(x_{n-1}) - F(z) = F'(\xi_n)(x_{n-1} - z) \text{ tức là}$$

$$|x_n - z| = \frac{|h''(\xi_n)h(\xi_n)|}{[h'(\xi_n)]^2} |x_{n-1} - z|$$

ở đây ξ_n là số nằm giữa x_{n-1} và z

Cho $\{x_n\} \rightarrow z$, khai triển h quanh nghiệm của nó ta nhận được:

$$|h(\xi_n) - h(z)| = |h(\xi_n)| = |h'(\eta_n)| |\xi_n - z| < |h'(\eta_n)| |x_{n-1} - z|$$

$$\Rightarrow |h(\xi_n)| < |h'(\eta_n)| |x_{n-1} - z|$$

ở đây η_n nằm giữa ξ_n và z. Đặt: $B_n = \frac{|h''(\xi_n)h'(\eta_n)|}{[h'(\xi_n)]^2}$ và đặt $B = \sup_n B_n$

$$\text{thì } |x_n - z| < B |x_{n-1} - z|^2. \text{ Quan sát ta thấy khi } n \rightarrow \infty \text{ thì } B_n \rightarrow \left| \frac{h''(z)}{h'(z)} \right|$$

Xét ví dụ sau đây

Nếu áp dụng phương pháp Newton vào hàm số: $h(x) = x^2 - \alpha$, $h'(x) = 2x$

$$\text{thì ta được công thức: } x_{n+1} = x_n - \frac{h(x_n)}{h'(x_n)} = x_n - \frac{x_n^2 - \alpha}{2x_n} = \frac{1}{2} \left(x_n + \frac{\alpha}{x_n} \right)$$

Với α cho trước ta chọn đoạn $[a; b]$ sao cho hàm F của phương pháp Newton là ánh xạ co.

Cách chọn a, b như sau:

Với $0 < a^2 < \alpha < b^2$, $b > \frac{a^2 + \alpha}{2a}$ và $3a^2 > \alpha$, chẳng hạn chọn $a = \frac{\alpha}{2}$, $b > 3a$

thì ta có: Với $x \in [a; b] \Rightarrow a \leq F(x) = x - \frac{h(x)}{h'(x)} = x - \frac{x^2 - \alpha}{2x} = \frac{x^2 + \alpha}{2x} \leq b$ để có được điều này

ta cần chứng minh giá trị max và min của F trên $[a; b]$ thuộc vào $[a; b]$.

Ta có $F'(x) = \frac{1}{2} \left(1 - \frac{\alpha}{x^2} \right)$ và $F''(x) = \frac{\alpha}{x^3} > 0$,

nên $F'(\sqrt{\alpha}) = 0$ và $F(\sqrt{\alpha}) = \sqrt{\alpha} \in [a; b]$.

do đó : max F phải xảy ra tại điểm $x = a$ hoặc $x = b$

bởi vì F' chỉ triệt tiêu tại duy nhất điểm thuộc $[a; b]$

nhưng $F(a) = \frac{(a^2 + \alpha)}{2a} < b$ và $F(b) = \frac{(b^2 + \alpha)}{2b} < b$ (vì $\alpha < b^2$) nên $\max F < b$.

ta còn có $\min F(x) = F(\sqrt{\alpha}) = \sqrt{\alpha} \in [a; b]$.

Vậy $a < \min F(x) < \max F(x) < b$

Từ giả thiết $3a^2 > \alpha$ ta suy ra được $-\frac{\alpha}{a^2} > -3$

nên: $-1 < \frac{1}{2} \left(1 - \frac{\alpha}{a^2} \right) \leq \frac{1}{2} \left(1 - \frac{\alpha}{x^2} \right) \leq \frac{1}{2} \left(1 - \frac{\alpha}{b^2} \right) < \frac{1}{2}$,

do đó $|F'(x)| < 1$ trên $[a, b]$

Vậy F là ánh xạ co trên $[a; b]$

Chú ý rằng nếu chúng ta chọn x_o bởi

$x_o = \alpha + 1 > \sqrt{\alpha}$, thì với $a = \sqrt{\frac{\alpha}{2}}$ và $b = \max\{3a, \alpha + 1\}$ thì $x_o \in [a; b]$

1.4. ÁNH XẠ TỰA CO (subcontractor)

Định nghĩa

Một ánh xạ tựa co là một ánh xạ đi từ khoảng hữu hạn I vào chính nó thỏa:

i) Với $x, y \in I \Rightarrow |f(x) - f(y)| \leq |x - y|$

ii) Nếu $x \neq f(x)$ thì $|f(f(x)) - f(x)| < |f(x) - x|$

Định lý 1.4.1

Giả sử f là một ánh xạ tựa co. Chọn x_o tùy ý trong I, đặt $x_{n+1} = f(x_n)$ thì $\{x_n\}$ có giới hạn là điểm bất động của f

Định lý này sẽ được chứng minh trong phần định lý điểm bất động của ánh xạ tựa co của không gian mêtric tổng quát trong 1.5

Bổ đề 1.4.1

Giả sử $f \in C^1[a; b]$; $0 \leq f'(x) \leq 1$ và $f'(x) \neq 1$ tại một số x thuộc $[a; b]$ thì

$$0 \leq \frac{1}{b-a} \int_a^b f'(t) dt < 1$$

Chứng minh

Do f' liên tục trên $[a, b]$ nên $\exists z \in [a; b]: f'(z) = \min_{[a; b]} f'(x)$

Từ giả thiết $\exists x \in [a, b]: 0 \leq f'(x) < 1$ ta có: $0 \leq f'(z) = q < 1$ Do f' liên tục trên $[a; b]$ nên tồn tại khoảng mở

$$N \subset I: \forall x \in N \Rightarrow f'(x) < \frac{(1+q)}{2}$$

Đặt $\mu(N)$ là độ đo của N thì:

$$\begin{aligned} \int_a^b f'(t) dt &= \int_N f'(t) dt + \int_{I \setminus N} f'(t) dt \leq \mu(N) \frac{1+q}{2} + (b-a) - \mu(N) = \\ &= \mu(N) \left[\frac{1+q}{2} - 1 \right] + (b-a) < b-a \quad (\text{vì } \frac{1+q}{2} - 1 < 0) \end{aligned}$$

$$\text{Vậy: } 0 \leq \frac{1}{b-a} \int_a^b f'(t) dt < 1.$$

Hệ quả

Giả sử $h \in C^1[a; b]$; $h(a)h(b) < 0$, $0 \leq \mu \leq h'(x) \leq \frac{1}{\gamma}$ và với mỗi khoảng con I' của $[a, b]$, tồn

tại x thuộc I' sao cho $h'(x) \neq 0$

đặt $x_{n+1} = x_n - \gamma h(x_n)$ với x_0 tùy ý trong $[a, b]$ thì dãy $\{x_n\}$ hội tụ về nghiệm của h .

Chứng minh

Với $F(x) = x - \gamma h(x)$ thì $\forall x \in [a; b]$ ta có $0 \leq F'(x) = 1 - \gamma h'(x) \leq 1$ và ta cũng có $a \leq F(x) \leq b$ do đó: $|F(x) - F(y)| \leq |x - y|, \forall x, y \in [a; b]$

(vì $|F(x) - F(y)| \leq |F'(\xi)| |x - y| \leq |x - y|$)

Chọn $x_0 \in I$,

nếu x_0 là nghiệm của h (hay là điểm bất động của F) thì dãy $\{x_n\}$ hội tụ về x_0 (đã chứng minh trong định lý 1.2.2)

nếu $F(x_o) \neq x_o$ thì do bổ đề 1.4.1 trên ta có:

$$\begin{aligned} |x_2 - x_1| &= |F(F(x_o)) - F(x_o)| = |F(x_1) - F(x_o)| = \\ &= |x_1 - x_o| \left| \frac{1}{x_1 - x_o} \int_{x_o}^{x_1} F'(t) dt \right| < |x_1 - x_o| = |F(x_o) - x_o| \end{aligned}$$

Vậy F thỏa điều kiện của ánh xạ co, áp dụng định lý 1.4.1 trên ta suy ra dãy $\{x_n\}$ hội tụ về z , với z là điểm bất động của F mà điểm bất động của F chính là nghiệm của h . Vậy $\{x_n\}$ hội tụ về z và $h(z) = 0$.

1.5. KHÔNG GIAN MÊ TRÍC

1.5.1. Các định nghĩa

Định nghĩa 1.5.1

Một không gian mê tric là một cặp gồm một tập hợp M và một hàm số thực không âm d , $d : M \times M \rightarrow \mathbb{R}$, hàm d thỏa ba điều kiện sau:

$$i) d(x; y) = 0 \text{ nếu và chỉ nếu } x = y$$

$$ii) d(x; y) = d(y; x), \forall x, y \in M$$

$$iii) d(x; y) + d(y; z) \geq d(x; z), \forall x, y, z \in M$$

Một không gian mê tric được định nghĩa như trên được ký hiệu là (M, d) .

Định nghĩa 1.5.2

Một ánh xạ F đi từ không gian mê tric M vào chính nó được gọi là một ánh xạ co trên M nếu có một số $q < 1$ sao cho với mọi cặp $x, y \in M$ thì $d(Fx, Fy) \leq qd(x, y)$

Để tiện cho việc trình bày sau này ta viết:

$$F(F(x)) = F^2x, F(F(F(x))) = F^3x, \dots$$

1.5.2. Định lý điểm bất động của ánh xạ co

Cho (M, d) là không gian mêtric đầy đủ, và F là ánh xạ co trên M . Chọn x_o là phần tử tùy ý của M . Thì dãy $\{F^n x_o\}$ hội tụ về z , với z là điểm bất động duy nhất của F .

Chứng minh

Đặt $F^n x_o = x_n$, với hai số tự nhiên m, n và $m \geq n$ thì

$$\begin{aligned} d(x_n, x_m) &= d(F^n x_o, F^m x_o) \leq q d(F^{n-1} x_o, F^{m-1} x_o) \leq \dots \leq \\ &\leq q^n d(x_o, F^{m-n} x_o) = q^n d(x_o, x_{m-n}) \end{aligned}$$

Ta có: $d(x_o, x_s) \leq d(x_o, x_1) + d(x_1, x_2) + \dots + d(x_{s-1}, x_s)$

$$\text{hay } d(x_o, x_s) \leq \sum_{i=1}^s d(x_{i-1}, x_i)$$

$$\text{do đó } q^n d(x_o, x_{m-n}) \leq q^n \sum_{i=1}^{m-n} d(x_{i-1}, x_i)$$

Mặt khác chúng ta có: $\forall i \geq 1$ thì

$$d(x_{i-1}, x_i) = d(F^{i-1} x_o, F^i x_o) \leq q d(F^{i-2} x_o, F^{i-1} x_o) \leq \dots \leq q^{i-1} d(x_o, x_1)$$

Do đó: $d(x_n, x_m) \leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \dots + d(x_{m-1}, x_m) \leq$

$$\leq q^n d(x_o, x_1) + q^{n+1} d(x_o, x_1) + \dots + q^{m-1} d(x_o, x_1) = q^n d(x_o, x_1) \sum_{i=1}^{m-n} q^{i-1}$$

$$\text{nhưng } \sum_{i=1}^{\infty} q^{i-1} = \frac{1}{1-q} \text{ nên } d(x_n, x_m) \leq q^n d(x_o, x_1) \frac{1}{1-q}$$

Vậy $\{x_n\}$ là dãy Cauchy trong không gian đầy đủ M nên $\{x_n\} \rightarrow z \in M$

Bởi vì F là ánh xạ co trên M nên nó liên tục trên M do đó: $z = \lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} F(x_n) = F(z)$

Tính duy nhất

Giả sử có hai điểm bất động z_1, z_2 và $z_1 \neq z_2$ khi đó:

$$0 < d(z_1, z_2) = d(F(z_1), F(z_2)) \leq qd(z_1, z_2) < d(z_1, z_2) \text{ (vô lý).}$$

Vậy định lý được chứng minh xong $\square$

Hệ quả

$F: M \rightarrow M$, (M, d) là không gian mê tric đầy đủ. Nếu tồn tại số tự nhiên n sao cho F^n là ánh xạ co trên M thì F có điểm bất động duy nhất.

Chứng minh

Do F^n là ánh xạ co trên M nên theo định lý ánh xạ co F^n có duy nhất điểm bất động, ta gọi là z .

Ta có $Fz = FF^n z = F^n Fz$ nên Fz là điểm bất động của F^n mà điểm bất động của F^n là duy nhất $\Rightarrow Fz = z \Rightarrow z$ là điểm bất động của F

Giả sử có $z_1 \neq z$ thỏa $Fz_1 = z_1$

thì $F^n(z_1) = z_1$ nên z_1 là điểm bất động của F^n

do đó $z_1 = z$. Tính duy nhất đã được chứng minh.

1.5.3. Định lý điểm bất động của ánh xạ tựa co

Đặt $Q: M \rightarrow M$, M Là không gian mê tric thỏa:

$$i) d(Qx, Qy) \leq d(x, y)$$

$$ii) \text{ Nếu } x \neq Qx \text{ thì } d(Qx, Q^2x) < d(x, Qx)$$

iii) Q có miền giá trị là tập compact.

Khi đó với mỗi x thuộc M , dãy $\{Q^n x\}$ hội tụ về điểm bất động của Q

Chứng minh

Do giả thiết i) nên ta có thể viết:

$$d(Q^n x, Q^{n+1} x) = d(QQ^{n-1} x, QQ^n x) \leq d(Q^{n-1} x, Q^n x) \leq \dots \leq d(x, Qx)$$

Do đó $\{d(Q^n x, Q^{n+1} x)\}$ là dãy số thực không tăng, bị chặn dưới bởi 0 nên nó có giới hạn.

Do iii) $\{Q^n(x)\} \in Q(M) : \text{compact} \Rightarrow$ tồn tại dãy con $\{Q^{n_k} x\}$ hội tụ về phần tử $y \in Q(M)$

Do $\{d(Q^n x, Q^{n+1} x)\}$ hội tụ, nên mọi dãy con $\{d(Q^{n_k} x, Q^{n_k+1} x)\}$ và $\{d(Q^{n_k+1} x, Q^{n_k+2} x)\}$ đều hội tụ và có cùng một giới hạn.

Ta có: $\lim_{k \rightarrow \infty} d(Q^{n_k} x, Q^{n_k+1} x) = \lim_{k \rightarrow \infty} d(Q^{n_k} x, Q Q^{n_k} x) = d(y, Qy)$

do đó: $d(y, Qy) = \lim_{k \rightarrow \infty} d(Q^{n_k+1} x, Q^{n_k+2} x) = \lim_{n \rightarrow \infty} d(Q Q^{n_k} x, Q^2 Q^{n_k} x) = d(Qy, Q^2 y)$

(do Q liên tục)

từ ii) ta suy ra $y = Qy$

do $\{Q^{n_k} x\}$ hội tụ về y nên: với $\varepsilon > 0$ cho trước ta chọn $N > 0$ thỏa $d(Q^N x, y) < \varepsilon$ thì

$$\begin{aligned} d(Q^{N+n} x, y) &= d(Q^{N+n} x, Q^{N+n} y) \leq \\ &\leq d(Q^{N+n-1} x, Q^{N+n-1} y) \leq \dots \leq d(Q^N x, y) < \varepsilon \quad (\text{do i}) \end{aligned}$$

do đó $\{Q^n(x)\}$ hội tụ về y .

Vậy định lý đã được chứng minh xong.

Chương 2:

PHƯƠNG PHÁP NEWTON VÀ LỜI GIẢI XẤP XỈ CỦA PHƯƠNG TRÌNH TRONG KHÔNG GIAN HỮU HẠN CHIỀU

Trong chương này chúng ta nghiên cứu việc ứng dụng của phương pháp Newton trong việc xây dựng lời giải xấp xỉ của nghiệm của phương trình trong không gian hữu hạn chiều.

2.1. CÁC ĐỊNH NGHĨA

Một tập con S của không gian metric $\mathbb{R}_n$ có tính chất, mỗi cặp điểm x, y thuộc S thì đoạn thẳng nối giữa hai điểm x, y thuộc S , S được gọi là tập lồi. Nói cách khác tập S gọi là tập lồi nếu x, y thuộc S thì $\lambda x + \mu y$ cũng thuộc S với λ, μ là hai số không âm và $\lambda + \mu = 1$. Bao lồi của tập S là giao của tất cả các tập lồi chứa S .

Một hàm số $f: S \subset \mathbb{R}_n \rightarrow \mathbb{R}$, với S là tập lồi, thỏa:

$\forall x, y \in S: f(\lambda x + \mu y) \leq \lambda f(x) + \mu f(y); \lambda, \mu > 0; \mu + \lambda = 1$, thì f được gọi là một hàm lồi (convex function).

Cho F là ánh xạ đi từ $\mathbb{R}_n$ vào chính nó, mà các thành phần của F thuộc lớp $C^1(\mathbb{R}_n)$. *Jacobian* của ánh xạ F tại z thuộc $\mathbb{R}_n$ là ma trận J với các thành phần là:

$$\left\{ \frac{\partial F_i(z)}{\partial x_j} \right\} (1 \leq i \leq n; 1 \leq j \leq n) \text{ và được ký hiệu là } J(z). \text{ Do đó } J(z)x \text{ là ký hiệu của tích của ma}$$

trận $J(z)$ và véc tơ n chiều x .

Định lý giá trị trung bình

Cho hàm $f \in C^1(S)$ với S là tập lồi của $\mathbb{R}_n$ với phần trong không rỗng, ta có:
 $f(z) - f(y) = [\nabla f(\xi), z - y]; \forall z, y \in S$ trong đó ξ thuộc đoạn thẳng nối giữa z và y , còn

$\nabla f(\xi) = \left(\frac{\partial f}{\partial x_1}(\xi), \frac{\partial f}{\partial x_2}(\xi), \dots, \frac{\partial f}{\partial x_n}(\xi) \right)$ là véc tơ n chiều (gọi là *Gradient* của f tại ξ), và

$$[\nabla f(\xi), z - y] = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\xi)(z_i - y_i).$$

Để cho gọn từ đây trở đi ta ký hiệu $L(x, y)$ là đoạn thẳng mở nối giữa hai điểm x, y .

2.2. CHUẨN

Ta đã có hàm khoảng cách $d(x,y)$ trên $\mathbb{R}_n$, $d^2(x,y) = \sum_{i=1}^n (x_i - y_i)^2$.

Hàm $d(x,0) = \left[\sum_{i=1}^n x_i^2 \right]^{\frac{1}{2}} = \|x\| = [x,x]^{\frac{1}{2}}$ được gọi là một chuẩn

Chuẩn là mê tric thỏa các điều kiện sau đây:

- i) $\|x\| = d(x,0) = 0$ nếu và chỉ nếu $x = 0$
- ii) $\|x\| + \|y\| \geq \|x + y\|$ (bất đẳng thức tam giác)
- iii) $\|\alpha x\| = |\alpha| \|x\|, \forall \alpha \in \mathbb{R}$

Bất kỳ hàm số nào đi từ $\mathbb{R}^n$ vào $\mathbb{R}^+$ thỏa ba tính chất i), ii), iii) được gọi là một chuẩn.

Gọi A là ma trận cấp m.n, và x là véc tơ n chiều. Ma trận A diễn tả một ánh xạ tuyến tính đi từ $\mathbb{R}_n$ vào $\mathbb{R}_m$.

Ta định nghĩa chuẩn $\|A\|$ là số M bé nhất thỏa bất đẳng thức $\|Ax\| \leq M \|x\|$ với mọi x, tất nhiên số $\|A\|$ luôn tồn tại bởi vì tập các số thực M được chọn bị chặn dưới bởi 0.

Do đó: $\|A\| = \inf \{ B : \|Ax\| \leq B \|x\|, \forall x \in \mathbb{R}_n \}$

Ta có kết quả sau:

$$\|A\| = \sup \left\{ \frac{\|Ax\|}{\|x\|} : x \neq 0 \right\} = \sup \{ \|Ax\| : \|x\| = 1 \}.$$

2.3. ĐỊNH LÝ GIÁ TRỊ TRUNG BÌNH MỞ RỘNG

Cho $F : \mathbb{R}_n \rightarrow \mathbb{R}_n$, S là tập lồi trong $\mathbb{R}_n$, giả sử F có Jacobian tại mỗi điểm của S. Thì

$$\|Fz - Fy\| \leq \sup \{ \|J(\xi)\| : \xi \in L(z,y) \} \|z - y\|, \text{ với } \|x\|^2 = [x,x]$$

Chứng minh

Áp dụng định lý giá trị trung bình đối với hàm số thực $\sum_{i=1}^n F_i(y) u_i$ (trong đó u_i là các thành phần của véc tơ của véc tơ đơn vị u).

ta được: $\sum_{i=1}^n (F_i(y) - F_i(z)) u_i = \left[\nabla \left(\sum_{i=1}^n F_i(\xi) \right) u_i, y - z \right]$ trong đó $\xi \in L[y,z]$ và phụ thuộc vào u.

$$\text{Do đó: } \left| \sum_{i=1}^n (F_i(y) - F_i(z)) u_i \right| = \left| \sum_{i=1}^n u_i [\nabla F_i(\xi), y - z] \right| \leq$$

$$\leq \|u\| \left[\sum_{i=1}^n [\nabla F_i(\xi), y-z]^2 \right]^{\frac{1}{2}} = \|J(\xi)(y-z)\| \leq \|J(\xi)\| \|y-z\| \leq \sup \{ \|J(\xi)\|, \xi \in L(y,z) \} \|y-z\|$$

Giả sử $F(y) \neq F(z)$ ta chọn $u = \frac{F(y) - F(z)}{\|F(y) - F(z)\|}$ thì

$$\left| \sum_{i=1}^n (F_i(y) - F_i(z)) u_i \right| = \frac{1}{\|F(y) - F(z)\|} \left| \sum_{i=1}^n (F_i(y) - F_i(z))^2 \right| = \|F(y) - F(z)\|$$

Ta được điều phải chứng minh.

2.4. CHẶN PHỔ

Cho A là ma trận cấp $n \times n$, với thành phần là các số thực, và A^* là ma trận chuyển vị của ma trận A , A^* có được bằng cách đổi chỗ giữa dòng và cột của ma trận A . Ta gọi A là ma trận đối xứng nếu $A = A^*$.

Cho A là ma trận đối xứng, ta có dạng toàn phương của ma trận A là hàm số:

$$x \rightarrow [x, Ax] = x^* Ax = \sum_{i=1}^n \sum_{j=1}^n x_i A_{ij} x_j$$

Các số $\mu = \min \{ [x, Ax] : \|x\| = 1 \}$; $\lambda = \max \{ [x, Ax] : \|x\| = 1 \}$ được gọi là các chặn theo phổ của A .

Do ánh xạ $[., A.]$ là ánh xạ liên tục và tập $S = \{x : \|x\| = 1\}$ là tập compact nên hàm $[., A.]$ sẽ đạt max và min trên S . Nếu $\mu > 0$ thì A được gọi là xác định dương, nếu $\mu = 0$ thì A gọi là bán xác định dương, nếu $\mu < 0$ và $\lambda > 0$ thì A gọi là không xác định. Xác định âm và bán xác định âm được định nghĩa tương tự.

Định lý

Giả sử A là ma trận đối xứng. Khi đó

$$\sup \{ [x, Ax] : \|x\| = 1 \} = \sup \{ \|Ax\| : \|x\| = 1 \} = \|A\|$$

Chứng minh

Áp dụng bất đẳng thức Schwarz ta có: $[x, Ax] \leq \|x\| \|Ax\| \leq \|A\|, \forall x : \|x\| \leq 1$

Đặt $\sup [x, Ax] = \sup \{ |\mu|, |\lambda| \} = \alpha$, thì $\alpha \leq \|A\|$

Ta chứng minh $\alpha \geq \|A\|$

Ta có:

$$\begin{aligned}
\|Ax\|^2 &= \frac{1}{4} [4[Ax, Ax]] = \frac{1}{4} [2[Ax, Ax] + 2[A^2x, x]] \\
(do: [Ax, Ax] &= (Ax)^* (Ax) = (x^* A^*) (Ax) = \\
&= x^* (A^* A) x = x^* (A^2 x) = [x, A^2 x] = [A^2 x, x]) \\
&= \frac{1}{4} \left\{ \gamma^2 [Ax, x] + [Ax, Ax] + [A^2 x, x] + \frac{1}{\gamma^2} [A^2 x, Ax] - \right. \\
&\left. \left[\gamma^2 [Ax, x] - [Ax, Ax] - [A^2 x, x] + \frac{1}{\gamma^2} [A^2 x, Ax] \right] \right\} \\
&= \frac{1}{4} \left\{ [A\gamma x, \gamma x] + [A\gamma x, A\gamma^{-1} x] + [A(\gamma^{-1} Ax), \gamma x] + [A(\gamma^{-1} Ax), A\gamma^{-1} x] - \right. \\
&\left. [[A\gamma x, \gamma x] - [A\gamma x, A\gamma^{-1} x] - [A(\gamma^{-1} Ax), \gamma x] + [A(\gamma^{-1} Ax), A\gamma^{-1} x]] \right\} \\
&= \frac{1}{4} \left\{ [A(\gamma x), \gamma x + A\gamma^{-1} x] + [A(\gamma^{-1} Ax), \gamma x + A\gamma^{-1} x] - \right. \\
&\left. [[A\gamma x, \gamma x - A\gamma^{-1} x] - [A(\gamma^{-1} Ax), \gamma x - A\gamma^{-1} x]] \right\} \\
&= \frac{1}{4} [[A(\gamma x + \gamma^{-1} Ax), \gamma x + A\gamma^{-1} x] - [A(\gamma x - \gamma^{-1} Ax), \gamma x - A\gamma^{-1} x]] \\
&\leq \frac{1}{4} \left(\alpha \|\gamma x + \gamma^{-1} Ax\|^2 + \alpha \|\gamma x - \gamma^{-1} Ax\|^2 \right) \leq \frac{\alpha}{4} [2(\|\gamma x\|^2 + \|\gamma^{-1} Ax\|^2)] \\
&= \frac{\alpha}{2} [\gamma^2 \|x\|^2 + \gamma^{-2} \|Ax\|^2]
\end{aligned}$$

Ở đây γ là số dương tùy ý. Khi $\|Ax\| \neq 0$ thì hàm số cuối cùng đạt min khi $\gamma^2 = \frac{\|Ax\|}{\|x\|}$, nên

$$\gamma^2 \|x\| = \|Ax\|$$

$$\text{từ đây ta có: } \|Ax\|^2 \leq \alpha [\|Ax\| \cdot \|x\|] \Rightarrow \|Ax\| \leq \alpha \|x\| \Rightarrow \alpha \geq \|A\|$$

$$\text{Vậy } \|A\| = \alpha$$

$$\text{Mà theo kết quả ở trên thì } \|A\| = \sup \{ \|Ax\|, \|x\| = 1 \}$$

$$\text{nên ta có: } \sup \{ [x, Ax] : \|x\| = 1 \} = \sup \{ \|Ax\| : \|x\| = 1 \} = \|A\|$$

Bổ đề 2.4.1

Giả sử A là ma trận đối xứng, các véc tơ đơn vị làm cực đại, cực tiểu ánh xạ $[., A.]$ là các véc tơ riêng của A .

Chứng minh

Gọi f và ϕ là hai hàm số thực thuộc lớp $C^1(\mathbb{R}_n)$

Điều kiện cần để điểm $x \in \{x \in \mathbb{R}_n : \phi(x) = 0\}$ sao cho hàm f đạt max hoặc min tại đó là tồn tại số δ sao cho $\nabla f(x) + \delta \nabla \phi(x) = 0$ (1)

Đặt $f(x) = [x, Ax], \phi(x) = 1 - \|x\|^2$ thì ta có:

$$\frac{\partial}{\partial x_k} f(x) = \frac{\partial}{\partial x_k} \sum_{i=1}^n x_i \sum_{j=1}^n A_{ij} x_j = \sum_{j=1}^n A_{kj} x_j + \sum_{i=1}^n x_i A_{ik} = 2 \sum_{j=1}^n A_{kj} x_j$$

(do A là ma trận đối xứng)

$$\text{do đó } \nabla f(x) = 2Ax, \nabla \phi(x) = -2x$$

từ công thức (1) cho ta: $Ax = \delta x$ với $\|x\| = 1$

bởi vì tồn tại các véc tơ đơn vị mà tại đó $[., A.]$ đạt các cực trị nên kéo theo tồn tại các véc tơ x và số δ thỏa $Ax = \delta x$, với $\|x\| = 1$.

Vậy bổ đề đã được chứng minh $\square$

Bổ đề 2.4.2

Giả sử A là ma trận đối xứng, số $\pm\delta$ là giá trị riêng của A ($Ax = \pm\delta x$) nếu và chỉ nếu δ^2 là giá trị riêng của A^*A ($A^*Az = \delta^2 z$).

Chứng minh

Giả sử $A^*Az - \delta^2 z = (A^2 - \delta^2 I)z = 0$ với $\delta^2 > 0$ thì

$$(A - \delta I)(A + \delta I)z = 0 = (A + \delta I)(A - \delta I)z$$

Có bốn khả năng xảy ra:

$$(A + \delta I)z = 0; (A - \delta I)z = 0; (A - \delta I)[(A + \delta I)z] = 0; (A + \delta I)[(A - \delta I)z] = 0$$

Trong bất kỳ trường hợp nào thì δ hoặc $-\delta$ cũng là giá trị riêng của A . Hơn nữa nếu $\delta^2 = 0$ thì $[z, A^*Az] = 0 \Rightarrow [Az, Az] = 0 \Rightarrow Az = 0$

Ngược lại: Nếu $Ax = \delta x$ thì $A^*Ax = \delta A^*x = \delta Ax = \delta^2 x$.

Bổ đề được chứng minh xong.

Bổ đề 2.4.3

Cho A là ma trận đối xứng với chặn phổ là μ, λ . Thì $\|I - \gamma A\| = \max\{|1 - \gamma\mu|, |1 - \gamma\lambda|\}$ với I là ma trận đơn vị.

Chứng minh

Nếu $\gamma \geq 0$ và $\|x\| = 1$ thì

$$\gamma\mu \leq [x, \gamma Ax] \leq \gamma\lambda \Rightarrow 1 - \gamma\mu \geq [x, Ix] - [x, \gamma Ax] \geq 1 - \gamma\lambda.$$

$$\text{mà } [x, Ix] - [x, \gamma Ax] = [x, (I - \gamma A)x]$$

$$\text{nên: } 1 - \gamma\mu \geq [x, (I - \gamma A)x] \geq 1 - \gamma\lambda \Rightarrow |[x, (I - \gamma A)x]| \leq \max\{|1 - \gamma\mu|, |1 - \gamma\lambda|\}$$

Vì: $1 - \gamma\mu, 1 - \gamma\lambda$ là các chặn phổ của ma trận đối xứng $I - \gamma A$ nên theo định lý trên Ta có:

$$\|I - \gamma A\| = \max\{|1 - \gamma\mu|, |1 - \gamma\lambda|\}$$

Nếu $\gamma < 0$ thì λ, μ đổi thứ tự cho nhau trong bất đẳng thức trên, nghĩa là ta có:

$$1 - \gamma\lambda \geq [x, (I - \gamma A)x] \geq 1 - \gamma\mu$$

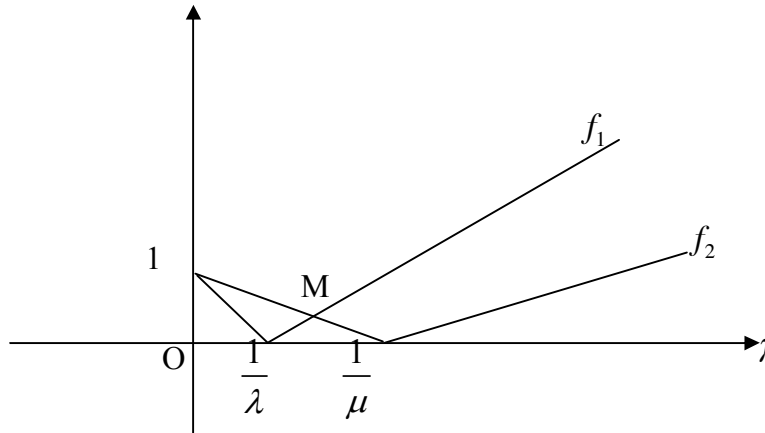
$$\text{Do đó ta cũng có: } \|I - \gamma A\| = \max\{|1 - \gamma\mu|, |1 - \gamma\lambda|\}.$$

Nhận xét: Nếu $\gamma, \lambda, \mu > 0$ thì hàm $f(\gamma) = \max\{|1 - \gamma\mu|, |1 - \gamma\lambda|\}$ có giá trị nhỏ nhất tại $\frac{2}{\mu + \lambda}$.

Chứng minh

$$\text{Đặt } f_1(\gamma) = |1 - \gamma\mu|, f_2(\gamma) = |1 - \gamma\lambda|$$

Khi $\mu > 0$, ta vẽ đồ thị của hai hàm số này trên cùng một hệ trục tọa độ:



Khi đó hàm f đạt min tại M , hoành độ của M là nghiệm của phương trình:

$$1 - \gamma\lambda = -(1 - \gamma\mu) \Leftrightarrow \gamma = \frac{2}{\gamma + \mu}$$

2.5. CỰC TIỂU HÓA HÀM SỐ

Ta gọi *Hessian* của hàm f tại $u \in \mathbb{R}_n$ là ma trận với các thành phần $\frac{\partial^2 f(u)}{\partial x_i \partial x_j}$ với

$$(1 \leq i \leq n, 1 \leq j \leq n)$$

Định lý

Giả sử f là hàm thuộc lớp $C^2(\mathbb{R}_n)$ sao cho phổ của Hessian của f là bị chặn dưới bởi $\mu > 0$ và bị chặn trên bởi λ với mọi x thuộc $\mathbb{R}_n$. Ta định nghĩa tập $I_\delta = \left[\delta; \left(\frac{2}{\lambda} \right) - \delta \right]$ với $0 < \delta \leq \frac{1}{\lambda}$.

Xét dãy $x^{k+1} = x^k - \gamma \nabla f(x^k)$ với $\gamma \in I_\delta$ và x^0 là phần tử tùy ý thuộc $\mathbb{R}_n$. Khi đó dãy $\{x^k\}$ hội tụ về z sao cho $f(z) \leq f(x)$ với mọi x thuộc $\mathbb{R}_n$.

Chứng minh

Ta tìm điểm z sao cho $\nabla f(z) = 0$.

Đặt $G(y) = y - \gamma \nabla f(y)$ thì $G_i(y) = y_i - \gamma \frac{\partial f(y)}{\partial x_i}$

Rõ ràng nếu $\gamma \neq 0$, y là điểm bất động của G khi và chỉ khi y là nghiệm của phương trình:

$$\nabla f(y) = 0$$

Đạo hàm của hàm $G_i(y)$ theo biến x_j được viết: $\frac{\partial G_i(y)}{\partial x_j} = \frac{\partial y_i}{\partial x_j} - \gamma \frac{\partial^2 f(y)}{\partial x_j \partial x_i}$

Hiển nhiên $\frac{\partial y_i}{\partial x_j} = 0$ nếu $i \neq j$ và $\frac{\partial y_i}{\partial x_j} = 1$ nếu $i = j$

Ta có $J_G(y) = I - \gamma H(y)$ với $J_G(y)$ là ma trận Jacobian của G tại y , I là ma trận đơn vị, $H(y)$ là ma trận Hessian của f tại y .

Do đó: $\|J_G(y)\| = \|I - \gamma H(y)\|$, do bổ đề 2.4.3 ta có:

$$\|J_G(y)\| \leq \max\{|1 - \gamma\mu|; |1 - \gamma\lambda|\}$$

Với $\gamma \in I_\delta$ thì $|1 - \gamma\lambda| \leq 1 - \delta\lambda$; $|1 - \gamma\mu| \leq 1 - \delta\mu$

Thật vậy:

$$* \gamma \in I_\delta \Rightarrow \delta \leq \gamma \leq \frac{2}{\lambda} - \delta \Rightarrow \lambda\delta \leq \lambda\gamma \leq 2 - \lambda\delta \Rightarrow -1 + \lambda\delta \leq -1 + \lambda\gamma \leq 1 - \lambda\delta$$

$$\Rightarrow |1 - \lambda\gamma| \leq 1 - \lambda\delta$$

* với $\delta \leq \gamma \leq \frac{2}{\lambda} - \delta \leq \frac{2}{\mu} - \delta$ thì ta cũng có bất đẳng thức thứ hai (làm tương tự).

Nên $\|J_G(y)\| \leq q < 1$ với $q = \max\{1 - \delta\lambda; 1 - \delta\mu\} = 1 - \delta\mu; \forall y \in \mathbb{R}_n$

Theo định lý giá trị trung bình ta có:

$\|Gy - Gz\| \leq \sup\{\|J_G(\xi)\| : \xi \in L(y, z)\} \|y - z\| \leq q \|y - z\|$, nên G là ánh xạ co.

Theo định lý điểm bất động của ánh xạ co ta có: Dãy $x^{n+1} = G(x^n)$ hội tụ về điểm bất động z của G, do đó có $\nabla f(z) = 0$

Theo khai triển Taylor, tồn tại $\xi \in L(x, z)$:

$$\begin{aligned} f(x) - f(z) &= [\nabla f(z), x - z] + \frac{1}{2} [x - z, H(\xi)(x - z)] \\ &= \frac{1}{2} \|x - z\|^2 \left[\frac{x - z}{\|x - z\|}, H(\xi) \cdot \frac{x - z}{\|x - z\|} \right] \geq \frac{1}{2} \mu \|x - z\|^2 \end{aligned}$$

$$\Rightarrow f(x) - f(z) \geq 0 \Rightarrow f(x) \geq f(z)$$

Định lý đã được chứng minh xong $\square$

Hệ quả

Cho F là ánh xạ đi từ $\mathbb{R}_n$ vào chính nó, và có J_F liên tục, đối xứng, xác định dương và có phổ bị chặn trên bởi λ và bị chặn dưới bởi μ . Thì F có nghiệm duy nhất z, và dãy lặp $x^{n+1} = x^n - \gamma F(x^n)$ hội tụ về z với tốc độ của một cấp

số nhân với mọi $\gamma \in \left[\delta; \frac{2}{\lambda} - \delta \right], 0 < \delta \leq \frac{1}{\lambda}$.

Chứng minh

Đặt: $G(y) = y - \gamma F(y)$

$$\begin{aligned} \Rightarrow G_i(y) &= y_i - \gamma F_i(y) \Rightarrow \frac{\partial G_i(y)}{\partial x_j} = \frac{\partial y_i}{\partial x_j} - \gamma \frac{\partial F_i(y)}{\partial x_j} \\ &\Leftrightarrow J_G(y) = I - \gamma J_F(y) \end{aligned}$$

$$\Rightarrow \|J_G(y)\| = \|I - \gamma J_F(y)\|$$

Mà giả thiết cho $J_F(y)$ là ma trận đối xứng xác định dương và bị chặn phổ trên và dưới bởi λ, μ (nên $\mu > 0$)

Theo bổ đề 2.4.3, ta có: $\|I - \gamma J_F(y)\| \leq \max\{|1 - \gamma\mu|, |1 - \gamma\lambda|\}$

Với $\gamma \in I_\delta = \left[\delta, \frac{2}{\lambda} - \delta \right], 0 < \delta \leq \frac{1}{\lambda}$ như trong định lý ta có:

$$|1 - \gamma\lambda| < 1 - \delta\lambda, |1 - \gamma\mu| < 1 - \delta\mu$$

Do đó: $\|J_G(y)\| \leq q < 1$ với $q = \max\{1 - \delta\lambda, 1 - \delta\mu\} = 1 - \delta\mu < 1$

Theo định lý giá trị trung bình ta có:

$$\|G(y) - G(z)\| \leq \sup\{\|J_G(\xi)\| : \xi \in L(y, z)\} \|y - z\| \leq q \|y - z\|$$

nên G là ánh xạ co, theo định lý ánh xạ co ta có G có điểm bất động duy nhất z và dãy $x^{k+1} = G(x^k)$ hội tụ về z

mà do cách đặt G nên điểm bất động z của G chính là nghiệm của F và dãy $x^{k+1} = G(x^k)$ hội tụ với tốc độ là cấp số nhân với công bội $q = 1 - \delta\mu$ (như trong định lý điểm bất động trong không gian mêtric).

Ứng dụng

Cho A là ma trận cấp m, n , $\text{rank} A = n$, do đó $m \geq n$. Gọi y là véc tơ m -thành phần, x là véc tơ n -thành phần. Đặt $F(x) = \|Ax - y\|$, $\|\cdot\|$ là chuẩn Euclide trên $\mathbb{R}_n$. Bây giờ ta tìm min của F .

$$\text{Ta có: } F^2(x) = \|Ax - y\|^2 = [Ax - y, Ax - y] = \sum_{i=1}^m \left(\sum_{j=1}^n A_{ij} x_j - y_i \right)^2$$

Ta chỉ cần tìm min của F^2

$$\text{Ta có: } \frac{\partial F^2}{\partial x_k} = 2 \sum_{i=1}^m \sum_{j=1}^n (A_{ij} x_j - y_i) A_{ik} \text{ và } \frac{\partial^2 F^2}{\partial x_s \partial x_k} = 2 \sum_{i=1}^m A_{ik} A_{is} = (2A^* A)_{sk}$$

mà $(A^* A)^* = A^* (A^*)^* = A^* A$ nên $A^* A$ là ma trận đối xứng

$$\text{mặt khác } [x, A^* A x] = (Ax)^* (Ax) = \sum_{i=1}^m \left(\sum_{j=1}^n A_{ij} x_j \right)^2 \geq 0$$

nếu $x \neq 0$ thì $Ax \neq 0$ (do $\text{rank} A = n$) $\Rightarrow [x, A^* A x] > 0 \Rightarrow A^* A$ xác định dương.

$$\text{Với } M = \left[\sum_i \sum_j A_{ij}^2 \right]^{\frac{1}{2}} \text{ thì ta có } \|Ax\| \leq M \|x\|$$

ta có: với $x \in E_n : \|x\| = 1$ thì:

$$\begin{aligned} [x, A^* A x] &= [x, A^* A x] \leq \|x\| \|A^* A x\| \leq \\ &\leq \left(\sum_{j=1}^m \sum_{i=1}^n (A^* A)_{ij}^2 \right)^{\frac{1}{2}} \|x\| = \left(\sum_{j=1}^m \sum_{i=1}^n (A^* A)_{ij}^2 \right)^{\frac{1}{2}} = \lambda \end{aligned}$$

áp dụng định lý trên: với $\delta \leq \gamma \leq \frac{2}{\lambda} - \delta, 0 < \delta \leq \frac{1}{\lambda}$ thì dãy

$$x^{k+1} = x^k - \gamma \nabla (F^2(x^k)) \text{ với } x_0 \text{ tùy ý, hội tụ về } z \text{ thỏa } \nabla F^2(z) = 0$$

Vậy F đạt min tại z.

Ý nghĩa hình học của phần ứng dụng: Cho trước véc tơ m-chiều y ta luôn tìm được véc tơ n-chiều x để cho khoảng cách giữa Ax và y là ngắn nhất.

2.6. KỸ THUẬT GRADIENT

2.6.1. Đặt vấn đề

Định lý trước đã được chứng minh là một trường hợp đặt biệt của một định lý tổng quát. Các giả thiết có được là quá mạnh và chúng ta thường không có được nó. Nhưng chúng ta có thể làm tốt hơn khi các giả định yếu hơn như sau:

Dãy lặp mà chúng ta thảo luận là $x^{k+1} = x^k - \gamma \nabla f(x^k)$. Bởi vì $\nabla f(x)$ là hướng mà theo hướng này hàm f tăng nhanh trong lân cận của x, nên chúng ta có khi γ đủ nhỏ và $\nabla f(x^k) \neq 0$ thì $f(x^k - \gamma \nabla f(x^k)) < f(x^k)$. Ngoại trừ khi chúng ta chọn γ để f đạt minimize, điều này sẽ tác động đến sự giảm của f với điều kiện $\nabla f(x^k) \neq 0$. Nếu $f(x^k)$ giảm liên tục, chỉ có các x^k xuất hiện trong suốt quá trình của dãy lặp thuộc vào tập $S = \{x \in \mathbb{R}_n : f(x) \leq f(x^0)\}$. Từ các gợi ý này chúng ta chỉ cần đòi hỏi chặn theo phổ của H trên S. Thật vậy, ta thấy rằng nếu $\gamma_k \in \left[\delta, \frac{2}{\lambda_0} - \delta\right]$ và $\|H(x)\| \leq \lambda_0 \forall x \in S$ thì $f(x^k)$ là dãy số giảm, nếu f bị chặn dưới thì $f(x^k)$ hội tụ dưới về một giới hạn và $\{\nabla f(x^k)\}$ hội tụ về 0, hơn nữa H(x) không cần xác định dương trên S.

Bổ đề

Cho f là một hàm số xác định trên một lân cận của điểm z trong $\mathbb{R}_n$. Giả sử ∇f xác định tại z và $\nabla f(z) \neq 0$. Thì khi đó tồn tại z' thuộc $\mathbb{R}_n$ sao cho $f(z') < f(z)$

Chứng minh

Đặt $u(\gamma) = z - \gamma \nabla f(z)$, $g(\gamma) = f(u(\gamma))$

$$\text{Khi đó: } g'(\gamma) = \sum_{i=1}^n \frac{\partial f(u(\gamma))}{\partial x_i} \cdot \frac{du_i}{d\gamma} = \left[\nabla f[u(\gamma)], \frac{du}{d\gamma} \right]$$

$$g'(0) = \sum_{i=1}^n \frac{\partial f(u(0))}{\partial x_i} \cdot \frac{du_i}{d\gamma}(0) = \sum_{i=1}^n \frac{\partial f(z)}{\partial x_i} \left[-\frac{\partial f(z)}{\partial x_i} \right] = -\|\nabla f(z)\|^2 < 0$$

Do g khả vi tại 0 nên:

$$\forall \varepsilon > 0, \exists \delta > 0 : 0 < \gamma < \delta \Rightarrow \frac{g(\gamma) - g(0)}{\gamma} - g'(0) < \varepsilon$$

Chọn $\varepsilon = -\frac{1}{2}g'(0)$, $0 < \gamma < \delta$ và đặt $z' = z - \gamma \nabla f(z)$

$$\text{Thì: } f(z') = g(\gamma) < \varepsilon \gamma + g(0) + \gamma g'(0) = g(0) + \frac{\gamma}{2} g'(0) < g(0) = f(z).$$

Bổ đề đã được chứng minh xong.

2.6.2. Phương pháp Gradient

Cho hàm f xác định trên $\mathbb{R}_n$, đặt $S = \{x \in \mathbb{R}_n : f(x) \leq f(x^o)\}$. $H(x)$ là ký hiệu Hessian của f tại x , f thuộc $C^2(S)$. Do S là tập không mở nên ta giả sử các đạo hàm riêng tồn tại trên tập mở chứa S .

Định lý

Giả sử f bị chặn dưới, chọn x^o tùy ý trong $\mathbb{R}_n$ và giả sử f thuộc $C^2(S)$. Giả sử có

$\gamma_o > 0 : \|H(x)\| \leq \frac{1}{\gamma_o}, \forall x \in S$. Chọn δ, γ_k thỏa $0 < \delta \leq \gamma_o$ và $\delta \leq \gamma_k \leq 2\gamma_o - \delta$ và đặt

$x^{k+1} = x^k - \gamma_k \nabla f(x^k)$ thì:

i) $\{x^k\} \in S, \{\nabla f(x^k)\}$ hội tụ về 0 và $\{f(x^k)\}$ hội tụ dưới về một giới hạn

ii) Giả sử S bị chặn thì mọi điểm dính (cluster point) z của $\{x^k\}$ đều thỏa $\nabla f(z) = 0$. Nếu $\nabla f(x)$ có nghiệm duy nhất trên S thì $\{x^k\}$ hội tụ về z và f đạt min tại z

iii) Giả sử tồn tại $\mu > 0$ sao cho $[y, H(x)y] \geq \mu \|y\|^2, \forall y \in \mathbb{R}_n, x \in S$, Giả sử S là tập lồi. Thì có duy nhất điểm z để f đạt min tại z , và $\{x^k\}$ hội tụ về z với tốc độ của một cấp số nhân có công bội bé hơn hoặc bằng $1 - \mu\delta$

iv) Giả sử (iii) với $\mu \geq 0$ và S bị chặn. Gọi D là đường kính của khối cầu chứa S . Thì $f(x^k) - \|\nabla f(x^k)\|D \leq f(z) \leq f(x^k)$

Chứng minh

i) Đặt $x(\gamma) = x - \gamma \nabla f(x)$ và $\Delta(x, \gamma) = f(x) - f(x(\gamma))$

Giả sử $x \in S$ và $\nabla f(x) \neq 0$, theo khai triển Taylor, với $x(\gamma) \in S$:

$$\begin{aligned}\Delta(x, \gamma) &= -[f(x(\gamma)) - f(x)] \\ &= -[x(\gamma) - x, \nabla f(x)] - \frac{1}{2}[x(\gamma) - x, H(\xi(\gamma))[x(\gamma) - x]] \\ &= -[-\gamma \nabla f(x), \nabla f(x)] - \frac{1}{2}[-\gamma \nabla f(x), H(\xi(\gamma))[-\gamma \nabla f(x)]] \\ &= \gamma \|\nabla f(x)\|^2 - \frac{\gamma^2}{2} \|\nabla f(x)\|^2 \left[\frac{\nabla f(x)}{\|\nabla f(x)\|}; H(\xi(\gamma)) \frac{\nabla f(x)}{\|\nabla f(x)\|} \right] \\ &= \gamma \|\nabla f(x)\|^2 \left(1 - \frac{\gamma}{2} Q(\gamma) \right)\end{aligned}$$

$$\text{với } Q(\gamma) = \left[\frac{\nabla f(x)}{\|\nabla f(x)\|}; H(\xi(\gamma)) \frac{\nabla f(x)}{\|\nabla f(x)\|} \right], \quad \xi(\gamma) \in L(x, x(\gamma))$$

Bây giờ ta chứng minh khi $\delta \leq \gamma \leq 2\gamma_o - \delta$ thì $x(\gamma) \in S$

Bởi vì $\nabla f(x) \neq 0$ nên theo bổ đề trong 2.6.1 trên ta có $\Delta(x, \gamma) > 0$ khi γ đủ nhỏ.

Đặt $\hat{\gamma}$ là số dương γ bé nhất (nếu có) mà $\Delta(x, \gamma) = 0$

$$* \text{ Nếu } \hat{\gamma} \text{ tồn tại nghĩa là } \Delta(x, \hat{\gamma}) = 0 \text{ thì } \left[1 - \frac{\hat{\gamma}}{2} Q(\hat{\gamma}) \right] = 0$$

$$\text{nên ta có: } Q(\hat{\gamma}) = \frac{2}{\hat{\gamma}}$$

Hơn nữa $x(\hat{\gamma}) \in S$ do $x \in S$ và $f(x) = f(x(\hat{\gamma}))$, suy ra

$\xi(\hat{\gamma}) \in S$ vì $\xi(\hat{\gamma}) \in L(x, x(\hat{\gamma}))$: là đường thẳng nối x với $x(\hat{\gamma})$, với $x \in S$, $x(\hat{\gamma}) \in S$,

$$f(x) = f(x(\hat{\gamma}))$$

Ta còn có: $Q(\hat{\gamma}) \leq |Q(\hat{\gamma})| \leq \|H(\xi(\hat{\gamma}))\| \leq \frac{1}{\gamma_o} \Rightarrow \frac{2}{\hat{\gamma}} \leq \frac{1}{\gamma_o} \Rightarrow \hat{\gamma} \geq 2\gamma_o$

Do đó $x(\gamma) \in S$ khi $\delta \leq \gamma \leq 2\gamma_o - \delta$.

* Nếu $\hat{\gamma}$ không tồn tại, thì $x(\gamma) \in S$ với mọi $\gamma > 0$ và ta cũng có: $Q(\gamma) \leq \frac{1}{\gamma_o}$

$$\text{nên } \Delta(x, \gamma) \geq \gamma \|\nabla f(x)\|^2 \left(1 - \frac{\gamma}{2\gamma_o}\right)$$

$$\text{Hàm } g(\gamma) = \gamma \left(1 - \frac{\gamma}{2\gamma_o}\right), \quad 0 < \delta \leq \gamma \leq 2\gamma_o - \delta$$

$$g'(\gamma) = 1 - \frac{\gamma}{\gamma_o}$$

$$g'(\gamma) = 0 \Leftrightarrow \gamma = \gamma_o$$

Bảng biến thiên của hàm $g(\gamma)$:

γ	δ	γ_o			$2\gamma_o - \delta$
$g'(\gamma)$		+	0	-	
$g(\gamma)$					

Hàm $g(\gamma)$ đạt giá trị lớn nhất tại: $\gamma = \gamma_o$

Hàm $g(\gamma)$ đạt giá trị nhỏ nhất tại $\gamma = \delta$ và tại $\gamma = 2\gamma_o - \delta$ với $g_{\min} = \delta \left(1 - \frac{\delta}{2\gamma_o}\right) > 0$, do đó:

$$\Delta(x^k, \gamma_k) \geq \|\nabla f(x^k)\|^2 \delta \left(1 - \frac{\delta}{2\gamma_o}\right) > 0$$

mà $\Delta(x^k, \gamma_k) = f(x^k) - f(x^k(\gamma_k)) = f(x^k) - f(x^k - \gamma_k \nabla f(x^k)) = f(x^k) - f(x^{k+1})$ nên $\{f(x^k)\}$

là dãy số giảm

Ta lại có f là hàm số bị chặn dưới nên $\{f(x^k)\}$ bị chặn dưới

$$\text{Suy ra tồn tại } L \text{ sao cho: } \begin{cases} f(x^k) \xrightarrow{K \rightarrow \infty} L \\ f(x^k) \geq L, \forall k \end{cases}$$

$$\text{Khi đó: } \nabla f(x^k, \gamma_k) = f(x^k) - f(x^{k+1}) \rightarrow 0 \text{ khi } k \rightarrow \infty$$

$$\Rightarrow \nabla f(x^k) \rightarrow 0 \text{ khi } k \rightarrow \infty.$$

ii) Gọi $\{x^{k_i}\}$ là dãy con của $\{x^k\}$, $x^{k_i} \xrightarrow{i \rightarrow \infty} z$ (vì z là cluster point của $\{x^k\}$)

$$f \in C^2(S) \Rightarrow \nabla f \text{ liên tục tại } z$$

$$\Rightarrow \nabla f(x^{k_i}) \xrightarrow{i \rightarrow \infty} \nabla f(z) = 0$$

Nếu $\{x^k\}$ không hội tụ về z , thì có một dãy con vô hạn $\{x^{k_j}\}$ của $\{x^k\}$ thỏa:

$$\{x^{k_j}\} \xrightarrow{j \rightarrow \infty} z' \neq z$$

lập luận tương tự như trên thì ta có: $\nabla f(z') = 0$ (trái với giả thiết ∇f có nghiệm duy nhất)

Do S bị chặn và S đóng nên S là tập compact, mà f liên tục trên S nên f đạt giá trị nhỏ nhất trên S , ta gọi z' là điểm mà f đạt giá trị nhỏ nhất, suy ra $\nabla f(z') = 0$, suy ra $z = z'$ (do tính duy nhất nghiệm)

Vậy $\lim_{k \rightarrow \infty} x^k = z$ và f đạt cực tiểu tại z

$$\text{iii) Từ giả thiết: } [y, H(x)y] \geq \mu \|y\|^2 \Rightarrow \left[\frac{y}{\|y\|}, H(x) \frac{y}{\|y\|} \right] \geq \mu, y \neq 0, y \in \mathbb{R}_n$$

Nên $\mu > 0$ là chặn phổ của $H(x)$, $x \in S$

$$\text{Đặt: } G(y) = y - \gamma \nabla f(y), \gamma \text{ cố định thuộc } [\delta, 2\gamma_0 - \delta]$$

Theo chứng minh i) thì G là ánh xạ đi từ S vào chính nó

$$\text{Ta có: } G_i(y) = y_i - \gamma \frac{\partial f}{\partial x_i}(y) \text{ và } \frac{\partial G_i}{\partial x_j}(y) = \frac{\partial y_i}{\partial x_j} - \gamma \frac{\partial^2 f}{\partial x_i \partial x_j}(y),$$

$$\text{ở đây: } \frac{\partial y_i}{\partial x_j} = 0 \text{ khi } i \neq j, \frac{\partial y_i}{\partial x_j} = 1 \text{ khi } i = j$$

Do đó: $J_G(y) = I - \gamma H(y)$, $H(y)$ là Hessian của f tại y , $J_G(y)$ là Jacobian của G tại y .

Do $f \in C^2(S)$ nên $H(y)$ là ma trận đối xứng,

$$\|H(y)\| = \sup\{\| [x, H(y)x] \|, \|x\| = 1\} \leq \frac{1}{\gamma_o} \text{ nên } \frac{1}{\gamma_o} \text{ là chặn phổ trên của } H(y)$$

$$\text{Theo bổ đề 2.4.3 ta có: } \|J_G(y)\| = \|I - \gamma H(y)\| = \max\left\{\left|1 - \frac{\gamma}{\gamma_o}\right|, |1 - \gamma\mu|\right\}$$

Ta có: với $\gamma \in [\delta; 2\gamma_o - \delta]$ thì

$$\frac{\delta}{\gamma_o} \leq \frac{\gamma}{\gamma_o} \leq 2 - \frac{\delta}{\gamma_o} \Rightarrow -1 + \frac{\delta}{\gamma_o} \leq -1 + \frac{\gamma}{\gamma_o} \leq 1 - \frac{\delta}{\gamma_o} \Rightarrow \left|1 - \frac{\gamma}{\gamma_o}\right| \leq 1 - \frac{\delta}{\gamma_o}$$

$$\text{và } \delta \leq \gamma \leq 2\gamma_o - \delta = \frac{2}{\frac{1}{\gamma_o}} - \delta \leq \frac{2}{\mu} - \delta \Rightarrow \mu\delta \leq \mu\gamma \leq 2 - \mu\delta$$

$$\Rightarrow -1 + \mu\delta \leq -1 + \mu\gamma \leq 1 - \mu\delta$$

$$\Rightarrow |1 - \mu\gamma| < 1 - \mu\delta$$

$$\text{Nên: } \|J_G(y)\| \leq 1 - \mu\delta < 1, \forall y \in S.$$

Áp dụng định lý giá trị trung bình mở rộng cho hàm G ta được:

$$\|Gy - Gz\| \leq \sup\{\|J_G(\xi)\|, \xi \in L(y, z)\} \|y - z\| \leq q \|y - z\|, \text{ với } q = 1 - \mu\delta < 1$$

Nên G là ánh xạ co trên S , do đó G có điểm bất động z thuộc S và dãy $x^{n+1} = G(x^n)$ hội tụ về

$$z \Rightarrow \nabla f(z) = 0$$

Áp dụng khai triển của Taylor:

$$\exists \xi \in L(x, z): f(x) - f(z) = [\nabla f(z), x - z] + \frac{1}{2} [x - z, H(\xi)(x - z)]$$

$$= \frac{1}{2} \|x - z\|^2 \left[\frac{x - z}{\|x - z\|}, H(\xi) \frac{x - z}{\|x - z\|} \right] \geq \frac{1}{2} \|x - z\|^2 \mu \geq 0, \forall x \in S, \xi \in S \text{ (do } S \text{ lồi)}$$

$$\Rightarrow f(x) \geq f(z)$$

Vậy f đạt minimum tại z thuộc S .

* Ta chứng minh dãy $\{x^k\}$ cũng hội tụ về z khi γ thay đổi trên $[\delta; 2\gamma_o - \delta]$

Áp dụng định lý giá trị trung bình, tồn tại ma trận E với các thành phần tiến về 0 sao cho:

$$\nabla f(x^n) = \nabla f(x^n) - \nabla f(z) = H(z)(x^n - z) + E(x^n - z)$$

$$\text{Ta có: } \|x^{n+1} - z\| = \|x^n - z - \gamma_n \nabla f(x^n)\|$$

$$= \|x^n - z - \gamma_n E(x^n - z) - \gamma_n H(z)(x^n - z)\|$$

$$= \|[I - \gamma_n H(z)](x^n - z) - \gamma_n E(x^n - z)\|$$

$$\leq \|[I - \gamma_n H(z)](x^n - z)\| + \gamma_n \|E(x^n - z)\|$$

$$\leq \|I - \gamma_n H(z)\| \|x^n - z\| + \gamma_n \|E\| \|x^n - z\|$$

$$\leq [(1 - \gamma_n \mu) + \gamma_n \|E\|] \|x^n - z\|$$

$$\Rightarrow \|x^{n+1} - z\| \leq (1 - \gamma_n \mu) \|x^n - z\| \leq (1 - \delta \mu) \|x^n - z\| \quad (\text{do } \gamma_n \text{ bị chặn, } E \text{ là ma trận có các thành phần tiến về } 0)$$

Vậy x^n hội tụ về z với tốc độ của một cấp số nhân có công bội $q = 1 - \delta \mu < 1$

iv) Do $\{x^k\} \subset S, f(z) = \min_{x \in S} f(x)$ nên $f(x^k) \geq f(z)$

Áp dụng khai triển Taylor ta có:

$$f(z) - f(x^k) = [\nabla f(x^k), z - x^k] + \frac{1}{2} [z - x^k, H(\xi)(z - x^k)]$$

với $\xi \in L(z, x^k)$

$$\Rightarrow f(z) = f(x^k) + [\nabla f(x^k), z - x^k] + \frac{1}{2} [z - x^k, H(\xi)(z - x^k)]$$

$$\text{Do } \mu \geq 0 \Rightarrow [z - x^k, H(\xi)(z - x^k)] \geq 0$$

$$\text{nên ta có: } 0 \leq f(x^k) - f(z) \leq [\nabla f(x^k), x^k - z] = \left| [\nabla f(x^k), x^k - z] \right|$$

Áp dụng bất đẳng thức Schwarz ta có:

$$\left| [\nabla f(x^k), x^k - z] \right| \leq \left\| \nabla f(x^k) \right\| \left\| x^k - z \right\|$$

Do S bị chặn và D là đường kính của khối cầu chứa S nên $\left\| x^k - z \right\| \leq D$

Vậy:

$$0 \leq f(x^k) - f(z) \leq \left\| \nabla f(x^k) \right\| \cdot D \Leftrightarrow f(x^k) \geq f(z) \geq f(x^k) - \left\| \nabla f(x^k) \right\| \cdot D \text{ Định lý đã được chứng minh xong.}$$

* **Ghi chú:** Trong trường hợp S bị chặn, giả sử biết được sự tồn tại của chặn $\frac{1}{\gamma_o}$ nhưng ta

không biết được giá trị số của nó là bao nhiêu, thì kết quả sau có thể được dùng: với mọi giá trị

$$m \in (0;1) \text{ thì có tồn tại } \bar{\gamma} > 0 \text{ sao cho: } m = \frac{\Delta(x, \bar{\gamma})}{\bar{\gamma} \left\| \nabla f_k \right\|} \text{ và } \Delta(x, \bar{\gamma}) \geq 2 \left\| \nabla f_k \right\| \gamma_o m (1-m), \text{ ở}$$

$$\text{đây: } \nabla f_k = \nabla f(x^k)$$

Chứng minh

Đặt: $g(\gamma) = \left[1 - \gamma \frac{Q(\gamma)}{2} \right] = \frac{\Delta(x, \gamma)}{\gamma \|\nabla f_k\|^2}$ (do kết quả trong chứng minh định lý trên)

Ta có $g(0) = 1, g(\hat{\gamma}) = 0$

Do g liên tục trên $[0; \hat{\gamma}]$ nên $g([0, \hat{\gamma}]) \supset [0; 1]$

với $m \in (0; 1) \Rightarrow \exists \bar{\gamma} \in (0; \hat{\gamma}) : g(\bar{\gamma}) = m$

$$\Delta(x, \bar{\gamma}) = \bar{\gamma} \cdot m \cdot \|\nabla f_k\|^2 \geq \bar{\gamma} \|\nabla f_k\|^2 \left(1 - \frac{\bar{\gamma}}{2\gamma_o} \right) \quad (\text{do } Q(\hat{\gamma}) \leq \frac{1}{\gamma_o})$$

$$\Rightarrow \bar{\gamma} \geq 2\gamma_o(1 - m)$$

$$\text{Do đó } \Delta(x, \bar{\gamma}) \geq \|\nabla f_k\|^2 \cdot 2\gamma_o m(1 - m)$$

Nhận xét: khi $m = \frac{1}{2}$ là tối ưu. Nhưng bất cứ m nào thỏa $0 < \delta \leq m \leq 1 - \delta$ thì kết quả trên vẫn thỏa mãn.

Để sử dụng kết quả này trong thực tế, ta xét một đoạn cụ thể chẳng hạn: $\left[\frac{1}{4}; \frac{3}{4} \right]$ (ở đây xét

$\delta = \frac{1}{4}$) ta chỉ cần tìm giá trị γ sao cho $\frac{1}{4} \leq g(\gamma) \leq \frac{3}{4}$.

2.7. PHƯƠNG PHÁP STEEPEST DESCENT

Liệu chúng ta có thể không cần giả thiết: tính bị chặn trên theo chuẩn của Hessian và điều kiện liên tục Lipschitz của ∇f .

Giả sử hàm f thuộc lớp $C^1(S)$, thêm vào điều kiện ∇f liên tục đều trên S , tất nhiên nếu ∇f liên tục trên S và S là tập compact thì ∇f sẽ liên tục đều trên S .

Ta định nghĩa: $\Delta(x, \gamma) = f(x) - f(x - \gamma \nabla f(x))$, ($\gamma \geq 0$); $g(x, \gamma) = \frac{\Delta(x, \gamma)}{\gamma \|\nabla f(x)\|^2}$ và

$$S = \{x \in E : f(x) \leq f(x_0)\}, \text{ chọn } \delta, 0 < \delta \leq \frac{1}{2}$$

Định lý 2.7.1

Giả sử ∇f tồn tại, bị chặn và liên tục trên S. chọn $\gamma_k < 1$ sao cho $\delta \leq g(x^k, \gamma_k) \leq 1 - \delta$, $g(x^k, 1) < \delta$; hoặc chọn $\gamma_k = 1$ khi $g(x^k, 1) \geq \delta$. Đặt $x^{k+1} = x^k - \gamma_k \nabla f(x^k)$, khi đó:

i) Nếu f bị chặn dưới và ∇f liên tục đều trên S thì $\nabla f(x^k) \xrightarrow{k \rightarrow \infty} 0$ và $f(x^k)$ hội tụ dưới về L.

ii) Nếu S bị chặn và gọi z là cluster point của $\{x^k\}$ thì $\nabla f(z) = 0$. Nếu tập hợp các cluster point của $\{x^k\}$ là hữu hạn thì $\{x^k\}$ sẽ hội tụ về một trong chúng. Nếu ∇f có nghiệm duy nhất thì f đạt min tại đó.

iii) Nếu f lồi và S là tập lồi thì $L = \inf \{f(x) : x \in \mathbb{R}_n\}$. Nếu S là tập lồi và bị chặn thì mọi cluster point của $\{x^k\}$ đều làm f đạt min tại đó.

Chứng minh

i) Cho γ đủ nhỏ, theo khai triển Taylor tồn tại $\xi \in L(x, x - \gamma \nabla f(x))$:

$$\begin{aligned} 0 < \Delta(x, \gamma) &= f(x) - f(x - \gamma \nabla f(x)) = [\nabla f(\xi), \gamma \nabla f(x)] = \gamma [\nabla f(\xi), \nabla f(x)] \\ &= \gamma \|\nabla f(x)\|^2 + \gamma [\nabla f(\xi) - \nabla f(x), \nabla f(x)] \end{aligned}$$

$$\text{Do đó: } g(x, \gamma) = 1 + \frac{[\nabla f(\xi) - \nabla f(x), \nabla f(x)]}{\|\nabla f(x)\|^2}$$

do ∇f liên tục tại x và $\|\xi - x\| \leq \|x - (x - \gamma \nabla f(x))\| = \gamma \|\nabla f(x)\|$ (*) nên $g(x, 0) = 1$

Nếu $g(x, 1) < \delta$, bởi vì $x - \gamma \nabla f(x) \in S$ nếu $g(x, \gamma) \geq 0$ và $\gamma \geq 0$ thì bởi tính

liên tục của hàm g theo biến thứ hai ta được hàm g sẽ nhận mọi giá trị giữa 1 và δ , do đó

$$\exists \gamma > 0 : \delta \leq g(x, \gamma) \leq 1 - \delta$$

Đặt $\nabla f(x^k) = \nabla f^k$, giả sử $\nabla f^k \neq 0$ và $x^k \in S$ thì

$$\Delta(x^k, \gamma_k) = \gamma_k g(x^k, \gamma_k) \|\nabla f^k\|^2 \geq \gamma_k \delta \|\nabla f^k\|^2 > 0 \text{ nên } f(x^{k+1}) < f(x^k) \text{ và } x^{k+1} \in S$$

Mặt khác theo giả thiết f bị chặn dưới trên S

$$\text{ nên } f(x^k) \xrightarrow{k \rightarrow \infty} L, f(x^k) \geq L, \forall k$$

Ta chứng minh $\nabla f^k \xrightarrow{k \rightarrow \infty} 0$

giả sử $\|\nabla f^k\|^2$ không hội tụ về 0 thì suy ra tồn tại dãy con của $\{x^k\}$ (giả sử là chính nó) và một số

$$\varepsilon > 0 \text{ sao cho } \|\nabla f^k\| \geq \varepsilon, \forall k$$

điều này kéo theo $\{\gamma_k\}$ bị chặn dưới bởi số dương, vì nếu không thì có dãy con của $\{\gamma_k\}$ (giả sử

$$\text{ là chính nó) sao cho } \gamma_k \xrightarrow{k \rightarrow \infty} 0$$

Do ∇f^k bị chặn trên S nên từ (*) ta có $\|x^k - \xi^k\| \xrightarrow{k \rightarrow \infty} 0$

Do ∇f liên tục đều trên S nên $[\nabla f(\xi^k) - \nabla f(x^k), \nabla f(x^k)] \xrightarrow{k \rightarrow \infty} 0$ và suy ra

$$g(x^k, \gamma_k) \xrightarrow{k \rightarrow \infty} 1. \text{ Điều này mâu thuẫn với } g(x^k, \gamma_k) < 1 - \delta \text{ với mọi } k$$

Do đó tồn tại $q > 0 : \gamma_k \geq q \Rightarrow \Delta(x^k, \gamma_k) \geq q\delta\varepsilon$

$$\begin{aligned} \Delta(x^k, \gamma_k) &= f(x^k) - f(x^k - \gamma_k \nabla f(x^k)) \geq q\delta\varepsilon \\ &= f(x^k) - f(x^{k+1}) \geq q\delta\varepsilon \end{aligned}$$

Điều này mâu thuẫn với sự hội tụ của $f(x^k)$.

Vậy ta phải có ∇f^k hội tụ về 0.

ii) Gọi z là cluster point của $\{x^k\}$ thì theo i) ta có $\nabla f(z) = 0$

* Gọi $A = \{z_1; z_2; z_3; \dots; z_n\}$ là tập tất cả các cluster point của $\{x^k\}$.

Do $\|x^{k+1} - x^k\| = \gamma_k \|\nabla f(x^k)\| \rightarrow 0$ và S bị chặn nên $\{x^k\}$ hội tụ về z_i nào đó trong A .

Nếu ∇f có nghiệm duy nhất z và S bị chặn, do số nghiệm của ∇f không ít hơn số cluster points của $\{x^k\}$ nên $\{x^k\} \rightarrow z$, f đạt min tại z (tương tự định lý trong 2.6.2)

iii) Trước tiên ta chứng minh bất đẳng thức: $f(y) \geq f(x) + [\nabla f(x), y - x]$

Ta có: $\frac{f(x + t(y - x)) - f(x)}{t} \leq f(y) - f(x)$, với $t \in (0, 1]$

(do f là hàm lồi nên $f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$)

áp dụng khai triển Taylor:

$$f(x + t(y - x)) - f(x) = t[\nabla f(x + \alpha t(y - x)), y - x], \text{ với } \alpha \in (0; 1)$$

qua giới hạn khi $t \rightarrow 0^+$ (do ∇f liên tục) ta được: $[\nabla f(x), y - x] \leq f(y) - f(x)$ hay $f(y) \geq f(x) + [\nabla f(x), y - x]$.

Ta chứng minh $L = \inf \{f(x) : x \in \mathbb{R}_n\}$

Giả sử $L \neq \inf \{f(x), x \in \mathbb{R}_n\} \Rightarrow \exists z \in S : f(z) < L$

$$\Rightarrow 0 > f(z) - f(x^k) \geq [\nabla f(x^k), z - x^k]$$

Giả sử S bị chặn ta gọi u là cluster point của $\{x^k\}$ thì có dãy con của $\{x^k\}$ hội tụ về u , để đơn giản ta ghi dãy con ấy là $\{x^k\}$, ta có:

$$\{[\nabla f(x^k), z - x^k]\} \xrightarrow{n \rightarrow \infty} 0 \text{ (do } \nabla f(x^k) \rightarrow \nabla f(u) = 0)$$

do đó $0 > \lim [f(z) - f(x^k)] = f(z) - L \geq \lim [\nabla f(x^k), z - x^k] = 0$ (vô lý)

Vậy $L = \inf \{f(x), x \in \mathbb{R}_n\}$

Do S là tập đóng và $\{x^k\}$ hội tụ về u nên $u \in S \Rightarrow f$ đạt min tại u .

Nếu S không bị chặn

$$\begin{aligned} \text{vì: } \|x^{k+1}\|^2 - \|x^k\|^2 &= [x^k - \gamma_k \nabla f^k, x^k - \gamma_k \nabla f^k] - [x^k, x^k] \\ &= -2\gamma_k [x^k, \nabla f^k] + \gamma_k^2 \|\nabla f^k\|^2 \end{aligned}$$

do đó nếu $\|x^k\|$ không bị chặn thì có một dãy con của $\{x^k\}$ (giả sử chính nó) sao cho:

$$\begin{aligned} -[x^k, \nabla f^k] &> \frac{-\gamma_k \|\nabla f^k\|^2}{2} \Rightarrow [\nabla f^k, z - x^k] > [\nabla f^k, z] - \frac{\gamma_k \|\nabla f^k\|^2}{2} \\ \Rightarrow 0 > f(z) - f(x^k) &\geq [\nabla f^k, z - x^k] > [\nabla f^k, z] - \frac{\gamma_k \|\nabla f^k\|^2}{2} \end{aligned}$$

Do $\nabla f^k \rightarrow 0$ (do i))

và γ_k là dãy bị chặn nên dãy $\left\{ [\nabla f^k, z] - \frac{\gamma_k \|\nabla f^k\|^2}{2} \right\}$ hội tụ về 0 nên từ bất đẳng thức trên ta suy

ra điều mâu thuẫn.

Vậy định lý được chứng minh xong $\square$

Hệ quả

Q là ánh xạ gradient của hàm số thực ϕ xác định trên $\mathbb{R}_n$.

Giả sử $\phi(x) \leq q \frac{\|x\|^2}{2}$, $q < 1$. Nếu Q liên tục thì nó có điểm bất động.

Chứng minh

Định nghĩa $f(x) = \frac{[x, x]}{2} - \phi(x)$

$$\Rightarrow f(x) \geq \frac{\|x\|^2}{2} - q \frac{\|x\|^2}{2} = \|x\|^2 \frac{1-q}{2} \Rightarrow f(x) \text{ bị chặn dưới}$$

Ta có: $\nabla f(x) = x - Q(x)$, do Q liên tục nên ∇f liên tục

Lấy x_o tùy ý thuộc $\mathbb{R}_n$ thì tập $S = \{x \in \mathbb{R}_n : f(x) \leq f(x_o)\}$ là tập bị chặn

$$\text{do: } \forall x \in S \text{ theo trên ta có } \|x\|^2 \frac{1-q}{2} \leq f(x) \leq f(x_o) \Rightarrow \|x\|^2 \leq \frac{2f(x_o)}{1-q}$$

Ta có: z là điểm bất động của Q khi và chỉ khi nó là nghiệm của ∇f

Theo định lý trên thì ∇f có nghiệm nên Q có điểm bất động.

Định lý 2.7.2

Giả sử ∇f liên tục trên bao lồi của S và S bị chặn.

Chọn $\gamma_k > 0$ để $f(x^k - \gamma_k \nabla f(x^k))$ đạt giá trị nhỏ nhất, và đặt $x^{k+1} = x^k - \gamma_k \nabla f(x^k)$

Thì $\{\nabla f(x^k)\} \rightarrow 0$ và mỗi điểm dính (cluster point) của $\{x^k\}$ là nghiệm của ∇f

Chứng minh

Gọi z là cluster point của $\{x^k\}$, không mất tính tổng quát nếu ta giả sử dãy con của $\{x^k\}$ hội tụ về z là chính nó.

Ta có $f(x^{k+1}) = f(x^k - \gamma_k \nabla f(x^k)) \leq f(x^k)$ (do $f(x^k - \gamma_k \nabla f(x^k))$ đạt min tại γ_k)

$\Rightarrow \{f(x^k)\}$ là dãy giảm

$f \in C^1, \{x^k\} \rightarrow z \Rightarrow \{f(x^k)\} \rightarrow f(z)$ (hội tụ giảm)

Bây giờ ta chứng minh $\nabla f(z) = 0$.

Giả sử trái lại, gọi $\gamma_o > 0 : f(z - \gamma_o \nabla f(z))$ đạt min tại γ_o

Nên $\exists \eta > 0 : f(z) = f(z - \gamma_o \nabla f(z)) + \eta$

Với mỗi x^k , áp dụng khai triển Taylor ta có:

$$\begin{aligned} f(x^k - \gamma_o \nabla f(x^k)) &= f[z - \gamma_o \nabla f(z) + x^k - z + \gamma_o (\nabla f(z) - \nabla f(x^k))] \\ &= f(z - \gamma_o \nabla f(z)) + [\nabla f(\xi^k), x^k - z + \gamma_o (\nabla f(z) - \nabla f(x^k))] \end{aligned}$$

với $\xi^k \in L(z - \gamma_o \nabla f(z), x^k - \gamma_o \nabla f(x^k))$

do: $\{x^k\} \rightarrow z \Rightarrow \nabla f(\xi^k) \rightarrow \nabla f(z - \gamma_o \nabla f(z))$

và $\{x^k - z + \gamma_o (\nabla f(z) - \nabla f(x^k))\} \rightarrow 0$

vì khi k đủ lớn thì ξ^k thuộc bao lồi của S nên ta có:

$$f(x^k - \gamma_o \nabla f(x^k)) \leq f(z - \gamma_o \nabla f(z)) + \frac{\eta}{2} = f(z) - \frac{\eta}{2}$$

Nhưng $f(z) < f(x^k - \gamma_k \nabla f(x^k)) \leq f(x^k - \gamma_o \nabla f(x^k)) \leq f(z) - \frac{\eta}{2}$ (vô lý)

(do $f(x^k - \gamma \nabla f(x^k))$ đạt min tại γ_k)

Vậy $\nabla f(z) = 0$. Định lý được chứng minh xong $\square$

Định lý 2.7.3

Giả sử ∇f liên tục trên tập bị chặn S và chọn γ_k là zero nhỏ nhất của $\frac{d}{d\gamma}(f(x^k - \gamma \nabla f(x^k)))$

trên nửa đường thẳng $\{x^k - \gamma \nabla f(x^k) : \gamma \geq 0\}$ thì:

Với z là cluster point của $\{x^k\}$ thì $\nabla f(z) = 0$. Nếu tập hợp các cluster point của $\{x^k\}$ là hữu hạn thì $\{x^k\}$ sẽ hội tụ về một trong chúng.

Nếu ∇f có nghiệm duy nhất thì f đạt min tại đó.

Chứng minh

Đặt $\Delta(x, \gamma) = f(x) - f(x - \gamma \nabla f(x)) = \gamma \|\nabla f(x)\|^2 g(x, \gamma)$

Giả sử $\nabla f(x) \neq 0$, ta có: $\gamma g(x, \gamma) = \frac{\Delta(x, \gamma)}{\|\nabla f(x)\|^2}$ và

$$\begin{aligned} \frac{d\Delta(x, \gamma)}{d\gamma} &= \frac{\partial \Delta(x, \gamma)}{\partial x_1} \cdot \frac{\partial x_1}{\partial \gamma} + \frac{\partial \Delta(x, \gamma)}{\partial x_2} \cdot \frac{\partial x_2}{\partial \gamma} + \dots + \frac{\partial \Delta(x, \gamma)}{\partial x_n} \cdot \frac{\partial x_n}{\partial \gamma} \\ &= -\frac{\partial f(x - \gamma \nabla f(x))}{\partial x_1} \left[-\frac{\partial f(x)}{\partial x_1} \right] - \dots - \frac{\partial f(x - \gamma \nabla f(x))}{\partial x_n} \left[-\frac{\partial f(x)}{\partial x_n} \right] \\ &= [\nabla f(x - \gamma \nabla f(x)), \nabla f(x)] \end{aligned}$$

Gọi $\hat{\gamma}$ là nghiệm dương nhỏ nhất của:

$$\frac{d}{d\gamma}(\gamma g(x, \gamma)) = \frac{[\nabla f(x - \gamma \nabla f(x)), \nabla f(x)]}{\|\nabla f(x)\|^2}$$

$$\text{thì } \hat{\gamma} g(x, \hat{\gamma}) = \int_0^{\hat{\gamma}} \frac{[\nabla f(x - \gamma \nabla f(x)), \nabla f(x)]}{\|\nabla f(x)\|^2} d\gamma$$

Giả sử $\{\nabla f(x^k)\}$ không hội tụ về 0

Mỗi $x = x^k$, xem hàm dưới dấu tích phân là hàm của γ , hàm dưới dấu tích phân là 1 khi $\gamma = 0$ và bằng 0 khi $\gamma = \gamma_k = \hat{\gamma}$

Ta chọn γ_o sao cho hàm lấy tích phân $I(x, \gamma) > \frac{1}{2}$ (với

$$I(x, \gamma) = \frac{[\nabla f(x) - \gamma \nabla f(x), f(x)]}{\|\nabla f(x)\|^2}) \text{ khi } \gamma \leq \gamma_o, \text{ cách chọn được thực hiện như sau:}$$

Rút ra dãy con của $\{x^k\}$ (giả sử là chính nó) hội tụ về z. Gọi K là tập hợp xác định bởi

$$K = \{z, x^1, x^2, \dots, x^k, \dots\} \text{ thì K là tập compact con của S}$$

Ta có: $\nabla f(x)$ liên tục đều trên K nên $\exists q > 0, r > 0 : \forall x \in K, q < \|\nabla f(x)\| < r$

do tính liên tục đều của ∇f trên S nên

$$\text{với } \varepsilon = \frac{q}{2}, \exists \delta > 0, \text{ khi } \gamma \|\nabla f(x)\| < \delta, \forall x \in K$$

$$\text{thì } \|\nabla f(x - \gamma \nabla f(x)) - \nabla f(x)\| < \frac{q}{2} < \frac{\|\nabla f(x)\|}{2}$$

$$\text{chọn } \gamma_o = \frac{\delta}{r} \text{ thì khi } \gamma \leq \gamma_o \Rightarrow \gamma \|\nabla f(x)\| < \delta$$

$$\text{Do đó: nếu } \gamma \leq \gamma_o \text{ và } x \in K \text{ ta có: } \|\nabla f(x - \gamma \nabla f(x)) - \nabla f(x)\| < \frac{\|\nabla f(x)\|}{2}$$

$$\Rightarrow \frac{1}{2} > \frac{1}{\|\nabla f(x)\|} \|\nabla f(x - \gamma \nabla f(x)) - \nabla f(x)\|$$

$$\geq \frac{1}{\|\nabla f(x)\|} \left\| \left[\nabla f(x - \gamma \nabla f(x)) - \nabla f(x), \frac{\nabla f(x)}{\|\nabla f(x)\|} \right] \right\|$$

$$= \left| \left[\frac{\nabla f(x - \gamma \nabla f(x))}{\|\nabla f(x)\|}, \frac{\nabla f(x)}{\|\nabla f(x)\|} \right] - 1 \right|$$

$$\text{Suy ra } \frac{[\nabla f(x) - \gamma \nabla f(x), f(x)]}{\|\nabla f(x)\|^2} - 1 > -\frac{1}{2} \text{ nên } I(x, \gamma) > \frac{1}{2}$$

$$\text{Do đó : } \begin{cases} \gamma_o < \hat{\gamma} \\ \hat{\gamma} g(\hat{\gamma}) > \int_0^{\gamma_o} \frac{1}{2} d\gamma = \frac{1}{2} \gamma_o \end{cases}$$

$$\text{Nên } \Delta(x^k, \gamma_k) = \gamma^k g(x, \gamma^k) \left\| \nabla f(x^k) \right\|^2 \geq \frac{\gamma_o q^2}{2} (*)$$

Điều này mâu thuẫn với tính chặn dưới của f trên K. Phần còn lại chứng minh tương tự 2.7.1

Định lý đã được chứng minh xong.

2.8. MỞ RỘNG

Định lý

Giả sử $f \in C^2(\hat{S})$, $\hat{S}$ là tập lồi chứa S, f bị chặn dưới. Giả sử với $\xi \in \hat{S}$ và $x \in \mathbb{R}_n$ thì $\mu \|x\|^2 < [x, H(\xi)x] \leq \lambda \|x\|^2$ với $\mu > 0$.

Đặt $x^{k+1} = x^k - \gamma_k H_k^{-1} \nabla f^k$, ở đây γ_k được chọn sao cho với $\delta < \frac{1}{2}$, $0 < \delta \leq g(\gamma_k) \leq 1 - \delta$,
(chọn $\gamma_k = 1$ nếu được)

Thì ta có các khẳng định sau:

- i) Với x_o tùy ý, dãy $\{x^k\}$ hội tụ về z, với x là điểm mà f đạt min tại đó.
- ii) Tồn tại số N sao cho khi $k > N$ thì $\gamma_k = 1$
- iii) Sự hội tụ của $\{x^k\}$ đến z là *superlinear*, nghĩa là chúng hội tụ với tốc độ nhanh hơn tốc độ của bất cứ cấp số nhân nào.

Ở đây :

$$\nabla f^k = \nabla f(x^k); H_k^{-1} = H^{-1}(x^k); \Delta(x, \gamma) = f(x) - f(x - \gamma H^{-1}(x) \nabla f(x)) \text{ và}$$

$$g(x, \gamma) = \frac{\Delta(x, \gamma)}{\gamma [\nabla f(x), H^{-1}(x) \nabla f(x)]}$$

Chứng minh

i) Trước tiên ta chứng minh $\Delta(x, \gamma) \geq \gamma [\nabla f(x), H^{-1}(x) \nabla f(x)] \left[1 - \frac{\gamma \lambda^2}{2\mu^2} \right]$

Từ giả thiết $\mu \|x\|^2 < [x, H(\xi)x] \leq \lambda \|x\|^2, x \neq 0$

$$\Rightarrow \mu \leq \left[\frac{x}{\|x\|}, H(\xi) \frac{x}{\|x\|} \right] \leq \lambda \Rightarrow \mu \leq \|H(\xi)\| \leq \lambda \Rightarrow \mu \leq \sup \frac{\|H(\xi)y\|}{\|y\|} \leq \lambda, y \in \mathbb{R}_n$$

$$\Rightarrow \mu \|y\| \leq \|H(\xi)y\| \leq \lambda \|y\| \quad (1)$$

Từ $\|H(\xi)y\| \geq \mu \|y\|$ ta có: nếu $H(\xi)y = 0 \Rightarrow y = 0$

$\Rightarrow \ker H(\xi) = \{0\} \Rightarrow H(\xi)$ có nghịch đảo

Từ (1) $\Rightarrow \mu \|H^{-1}(\xi)x\| \leq \|x\| \leq \lambda \|H^{-1}(\xi)x\|$

$$\Rightarrow \begin{cases} \|H^{-1}(\xi)x\| \leq \frac{1}{\mu} \|x\| \\ \|H^{-1}(\xi)x\| \geq \frac{1}{\lambda} \|x\| \end{cases} \Rightarrow \frac{1}{\lambda} \leq \frac{\|H^{-1}(\xi)x\|}{\|x\|} \leq \frac{1}{\mu} \Rightarrow \frac{1}{\lambda} \leq \|H^{-1}(\xi)\| \leq \frac{1}{\mu}$$

$$\Rightarrow \frac{1}{\lambda} \leq \left[\frac{x}{\|x\|}, H^{-1}(\xi) \frac{x}{\|x\|} \right] \leq \frac{1}{\mu} \Rightarrow \frac{\|x\|^2}{\lambda} \leq [x, H^{-1}(\xi)x] \leq \frac{\|x\|^2}{\mu}$$

Áp dụng khai triển Taylor:

$$\begin{aligned} \Delta(x^k, \gamma) &= f(x^k) - f(x^k - \gamma H_k^{-1} \nabla f^k) \\ &= [\gamma H_k^{-1} \nabla f^k, \nabla f^k] - \frac{\gamma^2}{2} [H_k^{-1} \nabla f^k, H(\xi(\gamma)) H_k^{-1} \nabla f^k] \end{aligned}$$

với $\xi(\gamma) \in L(x^k, x^k - \gamma H_k^{-1} \nabla f^k), \nabla f^k \neq 0$

Áp dụng bất đẳng thức Schwarz ta có:

$$0 < [H_k^{-1} \nabla f^k, H(\xi(\gamma)) H_k^{-1} \nabla f^k] \leq \|H_k^{-1} \nabla f^k\|^2 \|H(\xi(\gamma))\|$$

$$\begin{aligned}
\Rightarrow \Delta(x^k, \gamma) &\geq \gamma [\nabla f^k, H_k^{-1} \nabla f^k] - \frac{\gamma^2}{2} \|H_k^{-1} \nabla f^k\|^2 \|H(\xi(\gamma))\| \\
&\geq \gamma [\nabla f^k, H_k^{-1} \nabla f^k] - \frac{\gamma^2}{2} \frac{[\nabla f^k, H_k^{-1} \nabla f^k]}{\frac{\|\nabla f^k\|}{\lambda}} \|H_k^{-1} \nabla f^k\|^2 \|H(\xi(\gamma))\|
\end{aligned}$$

$$(\text{do } \frac{[\nabla f^k, H_k^{-1} \nabla f^k]}{\frac{\|\nabla f^k\|}{\lambda}} \geq 1)$$

$$\begin{aligned}
\Delta(x^k, \gamma) &\geq \gamma [\nabla f^k, H_k^{-1} \nabla f^k] \left(1 - \frac{\gamma \lambda}{2} \|H_k^{-1} \nabla f^k\|^2 \frac{\|H(\xi(\gamma))\|}{\|\nabla f^k\|^2} \right) \\
&\geq \gamma [\nabla f^k, H_k^{-1} \nabla f^k] \left(1 - \frac{\gamma \lambda}{2} \|H_k^{-1}\|^2 \|H(\xi(\gamma))\| \right) \\
&\geq \gamma [\nabla f^k, H_k^{-1} \nabla f^k] \left(1 - \frac{\gamma \lambda^2}{2 \mu^2} \right) \quad (2)
\end{aligned}$$

Bây giờ ta chứng minh S là tập bị chặn

Áp dụng khai triển Taylor, với u, x thuộc S ($u \neq x$) thì ta có:

$$\begin{aligned}
f(x) - f(u) &= [\nabla f(u), x - u] + \frac{1}{2} [x - u, H(C)(x - u)], C \in L(x, u) \\
\Rightarrow \frac{f(x) - f(u)}{\|x - u\|} &= \left[\nabla f(u), \frac{x - u}{\|x - u\|} \right] + \frac{1}{2 \|x - u\|} [x - u, H(C)(x - u)] \\
\Rightarrow \frac{f(x) - f(u)}{\|x - u\|} - \left[\nabla f(u), \frac{x - u}{\|x - u\|} \right] &\geq \frac{\mu}{2} \|x - u\| \quad (3)
\end{aligned}$$

với x thuộc S, do f bị chặn dưới nên tồn tại số B sao cho $f(x) \geq B$

$$\text{giả sử } \|x - u\| \geq \varepsilon > 0 \text{ khi đó } \frac{f(x) - f(u)}{\|x - u\|} \leq \frac{f(x_o) - B}{\varepsilon}$$

$$\Rightarrow \frac{\mu}{2} \|x_o - u\| \leq \frac{f(x_o) - B}{\varepsilon} \Rightarrow \|x_o - u\| \leq \frac{2}{\mu \varepsilon} [f(x_o) - B]$$

Vậy S bị chặn.

Suy ra tồn tại $\hat{\gamma}$ là số dương bé nhất sao cho $g(x, \hat{\gamma}) = 0$, điều này chứng tỏ γ_k được chọn thỏa yêu cầu (tương tự như phần ghi chú trong mục 2.6.2).

$$\begin{aligned}
 (2) &\Rightarrow g(x^k, \gamma_k) \geq \left[1 - \gamma_k \frac{\lambda^2}{2\mu^2} \right] > 0 \\
 &\Rightarrow \begin{cases} \gamma_k \geq \frac{2(1 - g(x^k, \gamma_k))\mu^2}{\lambda^2} \geq \frac{2\delta\mu^2}{\lambda^2} \\ \Delta(x^k, \gamma_k) = \gamma_k g(x^k, \gamma_k) [\nabla f^k, H^{-1}(x_k) \nabla f^k] \geq \frac{2\delta^2\mu^2}{\lambda^2} \frac{1}{\lambda} \|\nabla f^k\|^2 \end{cases} \\
 &\Rightarrow \nabla f(x^k) \rightarrow 0
 \end{aligned}$$

(do $\{f(x^k)\}$ là dãy giảm và bị chặn dưới nên nó hội tụ $\Rightarrow \Delta(x^k, \gamma_k) \rightarrow 0$)

Gọi z là cluster point của $\{x^k\}$ thì $\nabla f(z) = 0$

$$(3) \Rightarrow f(x) - f(z) \geq \mu \frac{\|x - z\|^2}{2} \geq 0 \Rightarrow f(x) \geq f(z), \forall x \in S$$

Vậy f đạt min tại z và z là duy nhất, do đó $\{x^k\} \rightarrow z$.

ii) Ta có: $H(\xi(\gamma)) = H(x) + H(\xi(\gamma)) - H(x)$

Từ đẳng thức

$$\begin{aligned}
 \Delta(x^k, \gamma) &= f(x^k) - f(x^k - \gamma H_k^{-1} \nabla f^k) \\
 &= [\gamma H_k^{-1} \nabla f^k, \nabla f^k] - \frac{\gamma^2}{2} [H_k^{-1} \nabla f^k, H(\xi(\gamma)) H_k^{-1} \nabla f^k] \\
 &\Rightarrow g(x, \gamma) = 1 - \frac{\gamma}{2} - \frac{\gamma [H^{-1}(x) \nabla f(x), (H(\xi(\gamma)) - H(x)) H^{-1}(x) \nabla f(x)]}{2 [\nabla f(x), H^{-1}(x) \nabla f(x)]}
 \end{aligned}$$

$$\Rightarrow \left| g(x, 1) - \frac{1}{2} \right| \leq \frac{\frac{\|H^{-1}(x)\nabla f(x)\|^2}{2\|\nabla f(x)\|^2} \|H(\xi(1)) - H(x)\|}{\lambda} \leq \frac{\|H(\xi(1)) - H(x)\| \lambda}{2\mu^2}$$

$$\Rightarrow \left| g(x^k, 1) - \frac{1}{2} \right| \leq \frac{\lambda}{2\mu^2} \|H(\xi(1)) - H(x^k)\|$$

Ta có:

$$\|H(\xi(1)) - H(x^k)\| \leq \|H(\xi(1)) - H(\xi(\gamma_k))\| + \|H(\xi(\gamma_k)) - H(x^k)\|$$

Do H liên tục trên S , S đóng và bị chặn nên H liên tục đều trên S , ta có:

$$\begin{cases} \|H(\xi(\gamma_k)) - H(x^k)\| \rightarrow 0 \\ \|H(\xi(1)) - H(\xi(\gamma_k))\| \rightarrow 0 \end{cases} \text{ khi } k \rightarrow \infty; \frac{1}{2} \in (\delta, 1 - \delta)$$

$$\Rightarrow g(x^k, 1) \rightarrow \frac{1}{2} \text{ khi } k \rightarrow \infty \Rightarrow \exists N > 0 : k > N \text{ thì } \gamma_k = 1.$$

iii) Giả sử $n > N$

Lập luận tương tự như trong chứng minh iii) của định lý phần 2.6.2), ta có:

$$\nabla f^k = \nabla f(z) + (H_k + E_k)(x^k - z)$$

với E_k là ma trận mà các thành phần của nó tiến về 0 khi $k \rightarrow \infty$

$$\Rightarrow x^{k+1} - z = x^k - z - H_k^{-1}(H_k + E_k)(x^k - z) = -H_k^{-1}E_k(x^k - z)$$

$$\Rightarrow \|x^{k+1} - z\| \leq \frac{1}{\mu} \|E_k\| \|x^k - z\|$$

Do $\|E_k\| \rightarrow 0$ khi $k \rightarrow \infty$, nên $\{x^k\}$ hội tụ với tốc độ *superlinear*.

Vậy định lý đã được chứng minh.

2.9. NGHIỆM CỦA HỆ PHƯƠNG TRÌNH

Xét hệ phương trình $R_i(x) = 0$, $1 \leq i \leq n$, với R_i là các hàm khả vi trên $\mathbb{R}_n$

$$\text{Đặt } g(x) = \sum_{i=1}^n R_i^2(x) = \|R(x)\|^2$$

Giả sử $\nabla g(z) = 0$ và Jacobian J của R tại z không suy biến

$$\text{Từ } \nabla g(z) = 0 \Rightarrow \sum_{i=1}^n R_i(z) \nabla R_i(z) = 0$$

Do $\{\nabla R_i, 1 \leq i \leq n\}$ độc lập tuyến tính nên $R_i(z) = 0$ (với $0 \leq i \leq n$)

Bổ đề 2.9.1

Cho A là ma trận vuông (n dòng n cột) có hạng n.

$$\text{ta có: } 0 \leq \|A^{-1}\|^{-1} \leq \|Au\|, \quad \forall u: \|u\| = 1$$

Chứng minh

$$1 = \|u\| = \|A^{-1}(A(u))\| \leq \|A^{-1}\| \|Au\| \Rightarrow \|Au\| \geq \|A^{-1}\|^{-1}$$

Do A^{-1} là nghịch đảo của A nên $\|A^{-1}\|^{-1} > 0$.

Bổ đề 2.9.2

Cho ánh xạ $R: \mathbb{R}_n \rightarrow \mathbb{R}_n$. Đặt $f(x) = \|R(x)\|$. Gọi z là nghiệm của f và $\{y^k\}$ là dãy mà giới hạn của nó là z. Giả sử $\nabla f(y^k) \neq 0, f(y^k) \neq 0$,

$$\text{đặt } T = \{y^k, k = 1, 2, 3, \dots\}$$

Giả sử J^* xác định và liên tục trên $\overline{T}$, $J^{*-1}(z)$ tồn tại. Thì tồn tại số dương q sao cho

$$\|\nabla f(x)\| \geq q, \quad \forall x \in T$$

Chứng minh

Ta có: $f^2(x) = \sum_{i=1}^n R_i^2(x)$, $f(x)\nabla f(x) = \sum_{i=1}^n R_i(x)\nabla R_i(x)$

$$\text{Nên } \nabla f(x) = \frac{\sum_{i=1}^n R_i(x)\nabla R_i(x)}{\|R(x)\|} = \frac{J^*(x)R(x)}{\|R(x)\|}, \forall x \in T$$

Đặt $g(x) = \|\nabla f(x)\|$, thì g là hàm liên tục tại x thuộc T (do J^* và R liên tục)

$$\text{Đặt } g(z) = \inf \left\{ \frac{\|J^*(z)R(x)\|}{\|R(x)\|}, x \in T \right\}$$

Ta có :

$$\begin{aligned} g(x) - g(z) &= \left\| \frac{J^*(x)R(x)}{\|R(x)\|} \right\| - g(z) = \left\| \frac{(J^*(z) + J^*(x) - J^*(z))R(x)}{\|R(x)\|} \right\| - g(z) \\ &= \left\| J^*(z) \frac{R(x)}{\|R(x)\|} + [J^*(x) - J^*(z)] \frac{R(x)}{\|R(x)\|} \right\| - g(z) \end{aligned}$$

Do J^* liên tục tại $z \Rightarrow (\forall \varepsilon > 0, \exists \delta(z) : x \in \delta(z) \Rightarrow \|J^*(x) - J^*(z)\| < \varepsilon)$, với $\delta(z)$ là quả cầu tâm z bán kính δ

Nên

$$g(x) - g(z) \geq \left\| \frac{J^*(z)R(x)}{\|R(x)\|} \right\| - \|J^*(x) - J^*(z)\| - g(z) \geq -\|J^*(x) - J^*(z)\| > -\varepsilon$$

$\Rightarrow g$ nửa liên tục dưới tại z

Đặt $\|J^{*-1}(z)\| = B$ theo bổ đề 2.9.1 trên ta có :

$$\left\| \frac{J^*(z)R(x)}{\|R(x)\|} \right\| \geq \|J^{*-1}(z)\|^{-1} = B^{-1}$$

$$\Rightarrow g(z) \geq B^{-1} > 0$$

Hàm g có tính chất: dương, nửa liên tục dưới trên $T' = T \cup \{z\}$, T' là tập compact

$$\Rightarrow g \text{ đạt min trên } T' \Rightarrow \exists q > 0 : g(x) \geq q, \forall x \in T'$$

$$\Rightarrow \text{nếu } x \neq z \text{ thì } g(x) \geq q \Rightarrow \|\nabla f(x)\| \geq q.$$

Định lý

Cho $R: \mathbb{R}_n \rightarrow \mathbb{R}_n$, với $R_i \in C^2(S)$, $0 \leq i \leq n$, giả sử S bị chặn.

Nếu $f(z) = 0$, giả sử $J^{*-1}(z)$ tồn tại. Nếu $f(x) \neq 0$, giả sử $\nabla f(z) \neq 0$. Cho $\delta \in \left(0, \frac{1}{2}\right)$, chọn

$$\gamma_k \text{ sao cho } \delta \leq g(x^k, \gamma_k) \leq 1 - \delta$$

Đặt $x^{k+1} = x^k + \gamma_k \phi(x^k)$. Nếu $f(x^k) = 0$ đặt $x^{k+1} = x^k$ thì

i) Mỗi z là cluster point của $\{x^k\}$ là nghiệm của R

ii) $f(x^k)$ hội tụ giảm về 0 với tốc độ của một cấp số nhân.

iii) Nếu J không suy biến trên S thì mỗi nghiệm của R bị cô lập và $\{x^k\}$ hội tụ với tốc độ của một cấp số nhân.

$$\text{Ở đây: } \{S = x \in E_n : f(x) \leq f(x_o)\}, \quad f(x) = \|R(x)\|$$

$$\phi: S \rightarrow \mathbb{R}_n, \text{ với tính chất } [\nabla f(x), \phi(x)] = -f(x) \text{ và } \|\phi(x)\| \leq \frac{f(x)}{\eta \|\nabla f(x)\|}, \quad \eta \in (0; 1)$$

$$\Delta(x, \gamma) = f^2(x) - f^2(x + \gamma \phi(x)); \quad g(x, \gamma) = \frac{\Delta(x, \gamma)}{2\gamma f^2(x)}$$

Chứng minh

i) và ii)

Nếu $f(x^k) \neq 0, k = 0, 1, 2, \dots$ ta luôn có thể chọn được $\phi^k = \phi(x^k)$ chẳng hạn:

Chọn $\phi(x) = \frac{-f(x^k)\nabla f(x^k)}{\|\nabla f(x^k)\|^2}$ thì ϕ thỏa các điều kiện :

$$[\nabla f(x^k), \phi^k] = \left[\nabla f(x^k), -f(x^k) \frac{\nabla f(x^k)}{\|\nabla f(x^k)\|^2} \right] = -f(x^k)$$

$$\|\phi^k\| = \frac{\| -f(x^k)\nabla f(x^k) \|}{\|\nabla f(x^k)\|^2} = \frac{f(x^k)}{\|\nabla f(x^k)\|}$$

Giả sử $f(x^k) \neq 0, k = 1, 2, \dots$

do $\{x^k\} \subset S, S$ bị chặn $\Rightarrow$ trích ra được một dãy con hội tụ của $\{x^k\}$ (không mất tính tổng quát nếu ta giả sử dãy con hội tụ là chính nó).

Áp dụng khai triển Taylor ta được:

$$\begin{aligned} \Delta(x^k, \gamma) &= f^2(x^k) - f^2(x^k + \gamma\phi^k) = -[f^2(x^k + \gamma\phi^k) - f^2(x^k)] \\ &= -[\nabla f^2(\xi^k), \gamma\phi^k] \\ &= -[\nabla f^2(x^k) + \nabla f^2(\xi^k) - \nabla f^2(x^k), \gamma\phi^k] \\ &= -[\nabla f^2(x^k), \phi^k] - \gamma[\nabla f^2(\xi^k) - \nabla f^2(x^k), \phi^k] \end{aligned}$$

Với $\xi^k \in L(x^k, x^k + \gamma\phi^k)$

Do $R_i \in C^2(S)$, S là tập bị chặn nên $R_i, \nabla R_i$ thỏa điều kiện Lipschitz. Ta kiểm tra ∇f^2 cũng thỏa điều kiện Lipschitz

ta có: $\nabla f^2(x) = 2f(x)\nabla f(x) = 2\sum_{i=1}^n R_i(x)\nabla R_i(x)$

$$\Rightarrow \nabla f^2(x) - \nabla f^2(y) = 2 \sum_{i=1}^n [R_i(x) \nabla R_i(x) - R_i(y) \nabla R_i(y)]$$

$$\Rightarrow \|\nabla f^2(x) - \nabla f^2(y)\| \leq 2 \sum_{i=1}^n \|R_i(x) \nabla R_i(x) - R_i(y) \nabla R_i(y)\|$$

$$\begin{aligned} \|R_i(x) \nabla R_i(x) - R_i(y) \nabla R_i(y)\| &\leq \|R_i(x) [\nabla R_i(x) - \nabla R_i(y)]\| + \\ &\quad + \|\nabla R_i(y) [R_i(x) - R_i(y)]\| \\ &= \|R_i(x)\| \|\nabla R_i(x) - \nabla R_i(y)\| + \|\nabla R_i(y)\| \|R_i(x) - R_i(y)\| \\ &\leq MN \|x - y\| + KL \|x - y\| = (MN + KL) \|x - y\| \end{aligned} \Rightarrow \nabla f \quad \text{thỏa}$$

điều kiện Lipschitz

Do đó tồn tại số dương G sao cho: $\|\nabla f^2(x) - \nabla f^2(y)\| \leq G \|x - y\|, \forall x, y \in S$

Ta thấy rằng nếu $\xi^k \in S$ thì $\|\xi^k - x^k\| \leq \|x^k + \gamma \phi^k - x^k\| = \gamma \|\phi^k\|$

Trở lại với việc tính toán $\Delta(x^k, \gamma)$

$$\begin{aligned} \Delta(x^k, \gamma) &= -\gamma [2f(x^k) \nabla f(x^k), \phi^k] - \gamma [\nabla f^2(\xi^k) - \nabla f^2(x^k), \phi^k] \\ &\geq -2\gamma f(x^k) [\nabla f(x^k), \phi^k] - \gamma G \|\phi^k\| \|\xi^k - x^k\| \\ &\geq 2\gamma f^2(x^k) - \gamma^2 G \|\phi^k\|^2 \end{aligned}$$

$$\begin{aligned} &\geq 2\gamma f^2(x^k) - \gamma^2 G \frac{f^2(x^k)}{\eta^2 \|\nabla f(x^k)\|^2} \\ &\geq 2\gamma f^2(x^k) - \gamma^2 G \frac{f^2(x^k)}{\eta^2 q^2} = 2\gamma f^2(x^k) \left(1 - \frac{\gamma G}{2\eta^2 q^2}\right) \end{aligned}$$

Với q là số thỏa $0 < q < \|\nabla f(x)\|$ (do bổ đề 2.9.2)

Gọi $\hat{\gamma}$ là số dương nhỏ nhất sao cho $\Delta(x^k, \gamma) = 0$

Ta có thể viết:

$$\Delta(x^k, \gamma) = 2\gamma f^2(x^k) \left[1 - \frac{[\nabla f^2(\xi^k) - \nabla f^2(x^k), \phi^k]}{2f^2(x^k)} \right]$$

Do $g(x^k, \gamma) = \frac{\Delta(x^k, \gamma)}{2\gamma f^2(x^k)}$, (tương tự như phần ghi chú trong mục 2.6.2), thì γ_k được chọn thỏa

yêu cầu giả thiết, đó là: $g(x^k, 0) = 1$, $g(x^k, \hat{\gamma}) = 0$.

Chọn $m \in (\delta, 1 - \delta)$, sao cho $g(x^k, \gamma_k) = m$

$$\text{nên } \Delta(x^k, \gamma_k) = 2\gamma_k m f^2(x^k) \geq 2\gamma_k f^2(x^k) \left[1 - \frac{\gamma_k G}{2\eta^2 q^2} \right]$$

$$\Rightarrow \gamma_k \geq \frac{2(1-m)\eta^2 q^2}{G} \text{ và}$$

$$\Delta(x^k, \gamma_k) \geq \frac{4(1-m)\eta^2 q^2}{G} m f^2(x^k) \geq \frac{4\delta(1-\delta)\eta^2 q^2}{G} f^2(x^k) > 0$$

Khi $\{x^k\} \rightarrow z \Rightarrow \Delta(x^k, \gamma_k) \rightarrow 0 \Rightarrow \{f^2(x^k)\} \rightarrow 0 \Rightarrow f(x^k) \rightarrow 0 \Rightarrow f(z) = 0$

Vậy mỗi z là cluster point của dãy $\{x^k\}$ thì z là nghiệm của R .

* Ta lại có:

$$f^2(x^k) - f^2(x^{k+1}) = g(x^k, \gamma_k) \cdot 2\gamma_k f^2(x^k)$$

$$\Rightarrow f(x^{k+1}) = f(x^k) \sqrt{1 - 2\gamma_k g(x^k, \gamma_k)} \Rightarrow f(x^{k+1}) = f(x^0) \prod_{i=0}^k (1 - 2\gamma_i g(x^i, \gamma_i))^{\frac{1}{2}} \text{ mà}$$

$$1 - 2\gamma_k g(x^k, \gamma_k) = \frac{f^2(x^{k+1})}{f^2(x^k)} = 1 - \frac{\Delta(x^k, \gamma_k)}{f^2(x^k)} \leq 1 - \frac{4\delta(1-\delta)\eta^2 q^2}{M} := \alpha^2 < 1$$

$$\text{nên } f(x^{k+1}) < f_0 \alpha^k$$

Vậy $\{f(x^k)\}$ hội tụ với tốc độ của một cấp số nhân.

iii) Gọi z là cluster point của $\{x^k\}$

Trước tiên ta chứng minh: nếu J_R không suy biến trên S thì $Q(x)$ (là Hessian của f^2 tại x) là xác định dương với mọi x thuộc lân cận của z .

$$\text{Đặt } A(x) = 2 \sum_{i=1}^n \frac{\partial^2 R_i(x)}{\partial x_i \partial x_k} R_i(x); \quad B(x) = 2 \sum_{i=1}^n \frac{\partial R_i(x)}{\partial x_j} \cdot \frac{\partial R_i(x)}{\partial x_k}$$

$$\text{Thì } Q(x) = A(x) + B(x) \text{ và } [B(x)u, u] = 2\|J(x)u\|^2$$

Do Hessian của R_i liên tục trên $S \Rightarrow A(x)$ liên tục và $A(z) = 0$

Do J không suy biến, gọi y là minimize $\|J^{-1}(\cdot)\|^{-1}$ trên S , đặt $\|J^{-1}(y)\|^{-2} = \mu > 0$.

Chọn lân cận $N(z)$ của z sao cho với mọi $x \in N(z) \Rightarrow \|A(x)\| \leq \mu$

Với mọi véc tơ đơn vị u , ta có:

$$\begin{aligned} [Q(x)u, u] &= [A(x)u, u] + [B(x)u, u] \\ &\geq -\mu + 2\|J(x)u\|^2 \geq -\mu + 2\|J^{-1}(x)\|^{-2} \geq \mu > 0 \end{aligned}$$

(do $[A(x)u, u] \leq \|A(x)\| \leq \mu \Rightarrow [A(x)u, u] \geq -\mu$ và bổ đề 2.9.1)

Vậy $Q(x)$ xác định dương.

Do $\nabla f^2(z) = 0$

$$\begin{aligned} \Rightarrow f^2(x) - f^2(z) &= [\nabla f^2(z), x - z] + \frac{1}{2}[Q(\xi)(x - z), x - z] \\ &\geq \frac{\mu}{2}\|x - z\|^2; \quad \forall x \in N(z) \end{aligned}$$

Vậy nghiệm z là cô lập.

Do Hessian của f^2 là xác định dương và có chuẩn lớn hơn $\mu > 0$ nên áp dụng định lý trong phần 2.6.2 ta được dãy $\{x^k\}$ hội tụ về z với tốc độ của một cấp số nhân.

Vậy định lý được chứng minh $\square$

Sự lựa chọn hướng

Giả sử rằng tại x^k , $J^{-1}(x^k)$ tồn tại và $\|J^{-1}(x^k)\| \leq \frac{1}{\eta} \|\nabla f(x^k)\|$ thì ϕ^k được chọn sao cho:

$J(x^k)\phi^k = -R(x^k)$ (gọi là phương thức Newtonian)

Khi đó $\phi^k = J^{-1}(x^k)R(x^k) \Rightarrow \|\phi^k\| \leq \|J^{-1}(x^k)\| \|R(x^k)\| \leq \frac{f(x^k)}{\eta} \|\nabla f(x^k)\|$ và

$$\begin{aligned} [\nabla f(x^k), \phi^k] &= - \left[\frac{J^*(x^k)R(x^k)}{\|R(x^k)\|}, J^{-1}(x^k)R(x^k) \right] \\ &= -\|R(x^k)\| \left[\frac{J^*(x^k)R(x^k)}{\|R(x^k)\|}, \frac{J^{-1}(x^k)R(x^k)}{\|R(x^k)\|} \right] \\ &= -\|R(x^k)\| [J^*(x^k)y, J^{-1}(x^k)y]; y = \frac{R(x^k)}{\|R(x^k)\|} \\ &= -\|R(x^k)\| y^* J(x^k) J^{-1}(x^k) y \\ &= -\|R(x^k)\| = -f(x^k) \end{aligned}$$

Nếu J không suy biến trên S , tồn tại hằng số M, N sao cho $\|J^{-1}(x)\| \leq M$ và

$\|\nabla f(x)\| \leq N, \forall x \in S$. Giả sử η được chọn sao cho $\eta \leq \frac{1}{MN}$ thì

$\|\phi(x)\| \leq Mf(x) \leq \frac{f(x)}{\eta N} \leq \frac{f(x)}{\eta \|\nabla f(x)\|}$. Do đó nếu J không suy biến trên S , ϕ luôn luôn được chọn

bằng phương thức Newtonian.

Hệ quả

Giả sử các giả thiết của định lý trên vẫn còn đúng. Giả sử J không suy biến

trên S và ϕ được chọn thỏa phương thức Newtonian, nghĩa là $\phi^k = -J_k^{-1}R_k$.

Chọn $\gamma_k = 1$ nếu $g(x^k, 1) > 1 - \delta$; chọn γ_k sao cho $\delta \leq g(x^k, \gamma_k) \leq 1 - \delta$, với $\gamma_k = 1$ (nếu được).

Thì:

i) Tồn tại số dương N sao cho nếu $k > N$ thì $\gamma_k = 1$

ii) $\{x^k\}$ hội tụ với tốc độ superlinear tới nghiệm của R .

Chứng minh

i) Theo định lý trên ta có $\{x^k\} \rightarrow z, R(z) = 0$. Giả sử $f(x^k) > 0, k = 1, 2, 3, \dots$

Gọi Q là Hessian của f^2 tại x và đặt $Q(x^n) = Q_n$.

$$\text{Ta chứng minh } \lim_{n \rightarrow \infty} \frac{[\phi^n, Q_n \phi^n]}{2f_n^2} = \lim_{n \rightarrow \infty} \theta_n = 1 \text{ với } \theta_n = \frac{[\phi^n, Q_n \phi^n]}{2f_n^2}$$

Như trong định lý trước ta có: $Q_n = A(x^n) + B(x^n)$

$$\lim_{n \rightarrow \infty} [u, Q_n u] = \lim_{n \rightarrow \infty} [u, A(x^n) u] + \lim_{n \rightarrow \infty} [u, B(x^n) u] = 2 \lim_{n \rightarrow \infty} \|J_n u\|^2, J_n = J(x_n)$$

Do $\{x^n\}$ hội tụ và ϕ liên tục $\Rightarrow \phi(x^n) := \phi^n$ hội tụ

$$\Rightarrow \lim_{n \rightarrow \infty} [\phi^n, Q_n \phi^n] = 2 \lim_{n \rightarrow \infty} \|J_n \phi^n\|^2$$

$$\text{Do đó: } \lim_{n \rightarrow \infty} \theta_n = \lim_{n \rightarrow \infty} \frac{\|J_n \phi^n\|^2}{f_n^2} = \lim_{n \rightarrow \infty} \frac{\|J_n J_n^{-1} R_n\|^2}{f_n^2} = 1$$

Áp dụng khai triển Taylor:

$$\begin{aligned} \Delta(x^n, \gamma) &= f^2(x^n) - f^2(x^n + \gamma \phi^n) = -\gamma [2f(x^n) \nabla f(x^n), \phi^n] - \frac{\gamma^2}{2} [Q(\xi^n) \phi^n, \phi^n] \\ &= -2\gamma f(x^n) [\nabla f(x^n), \phi^n] - \frac{\gamma^2}{2} [Q_n \phi^n, \phi^n] - \frac{\gamma^2}{2} [(Q(\xi^n) - Q_n) \phi^n, \phi^n] \end{aligned}$$

Với $\xi^n \in L(x^n, x^n + \gamma \phi^n)$. Nếu γ đủ nhỏ sao cho $\xi^n \in S$ thì

$$\Delta(x^n, \gamma) \geq f_n^2 \left[2\gamma - \gamma^2 \theta_n - \frac{\gamma^2 \|Q(\xi^n) - Q_n\|}{2\eta^2 \|\nabla f(x^n)\|^2} \right]$$

Theo bổ đề 2.9.2, $\exists q > 0: \|\nabla f(x)\| \geq q \Rightarrow \|\phi^n\| \leq \frac{f(x^n)}{\eta q}$

Do $g(x^n, \gamma) > 0, \xi^n \in S$, đặt $\theta_n = 1 + \alpha_n$, $\frac{\|Q(\xi^n) - Q_n\|}{2\eta^2 \|\nabla f(x^n)\|^2} = \beta_n, \alpha_n + \beta_n = \lambda_n$

Thì $g(x^n, \gamma) = 1 - \frac{\gamma}{2} - \frac{\gamma \lambda_n}{2}$.

Do $\{\lambda_n\} \rightarrow 0 \Rightarrow \{\exists N > 0: \forall n > N \Rightarrow \delta < g(x^n, 1) < 1 - \delta\}$

Do đó với $n \geq N$ thì $\gamma_n = 1$

Do đó i) được chứng minh.

Chứng minh ii)

Nếu $n = N$ thì $f_n^2 - f_{n+1}^2 = 2f_n^2 \left(\frac{1}{2} - \frac{\lambda_n}{2} \right) = f_n^2 (1 - \lambda_n)$, $f_{n+1}^2 = \lambda_n f_n^2$ với $\{\lambda_n\} \rightarrow 0$

Ta thấy khi $0 \leq \lambda_n < 1$ thì $\|x^n - z\|^2 \leq \frac{2}{\mu} f^2(x^n)$ (như trong định lý)

$$\Rightarrow \|x_n - z\| \leq \sqrt{\frac{2}{\mu}} f_n$$

$$\Rightarrow \|x_o - z\| \leq \sqrt{\frac{2}{\mu}} f_o; \|x^1 - z\| \leq \sqrt{\frac{2}{\mu}} \sqrt{\lambda_1} f_o; \dots; \|x^n - z\| \leq \sqrt{\frac{2}{\mu}} \prod_{i=1}^{n-1} \sqrt{\lambda_i} f_o$$

$$\Rightarrow \|x_n - z\| \leq \sqrt{\frac{2}{\mu}} \prod_{i=1}^{n-1} \sqrt{\lambda_i} f_o \leq \sqrt{\frac{2}{\mu}} \sqrt{\lambda_n} f_o \rightarrow 0 \text{ khi } n \rightarrow \infty$$

Vậy $\{x^n\}$ hội tụ về nghiệm của R với tốc độ superlinear.

2.10. ỨNG DỤNG

Cho f là hàm số bị chặn xác định trên tập T . Giả sử ta biết giá trị của f tại m điểm thuộc T là $f(t_i), i = 1, \dots, m$. Gọi $g_1, g_2, \dots, g_n$ là n hàm số bị chặn trên T , chúng được cho giá trị tại các $t_i, i = 1, \dots, m$

Chúng ta quan tâm tới việc xấp xỉ các giá trị của f bởi tổ hợp tuyến tính của các giá trị của $g_i, i = 1, \dots, n$. Gọi A là ma trận với các thành phần $A_{ij} = g_j(t_i), i = \overline{1, m}; j = \overline{1, n}$ và gọi b là véc tơ với m thành phần $b_i = f(t_i), i = \overline{1, m}$

Ta có:
$$\sum_{j=1}^n x_j g_j(t_i) - f(t_i) = \sum_{j=1}^n A_{ij} x_j - b_i; i = \overline{1, m}$$

Ta ký hiệu chuẩn trên $\mathbb{R}_m$ là $\|\cdot\|$. Ta định nghĩa một xấp xỉ tốt của f bởi tổ hợp

tuyến tính của $g_j (j = \overline{1, n})$ với chuẩn $\|\cdot\|$ là hàm $\sum_{j=1}^n x_j g_j(t_i), i = \overline{1, m}$ nếu x là

min của hàm $\|Ax - b\|$.

Trên $\mathbb{R}_m$ xét chuẩn mới được định nghĩa như sau: $\|x\|_p = \left[\sum_{j=1}^m |x_j|^p \right]^{\frac{1}{p}}, p \geq 1$

Định lý

Tồn tại duy nhất x để hàm $\|Ax - b\|$ đạt giá trị nhỏ nhất, trong đó A là $m.n$ ma trận có $\text{rank} A = n$ ($m \geq n$)

Chứng minh

Đặt $L = \{x \in \mathbb{R}_n : \|Ax - b\|_p \leq K\}$, ta chứng minh tập L này là tập bị chặn.

Lấy $\{x^k\} \in L, \|x^k\| \xrightarrow{k \rightarrow \infty} \infty$, chọn x^o là min của hàm $\frac{\|Ax\|_p}{\|x\|_p}$.

Đặt $\mu^o = \frac{\|Ax^o\|_p}{\|x^o\|_p}$ thì $\mu^o > 0$, do đó: $\|Ax - b\|_p \geq \|Ax\|_p - \|b\|_p \geq \mu^o \|x\|_p - \|b\|_p$ từ bất đẳng thức này

ta suy ra được nếu $\{\|x^k\|_p\} \rightarrow \infty$ thì $\{\|Ax^k - b\|_p\} \rightarrow \infty$ (mâu thuẫn), vậy L bị chặn

Do $\|\cdot\|_p$ là chuẩn và hàm $x \rightarrow \|Ax - b\|_p$ là hàm liên tục, L là tập compact nên nó đạt min trên L.

Ta chứng minh $\|Ax - b\|_p$ đạt min tại điểm duy nhất thuộc $\mathbb{R}_n$

Gọi m là giá trị nhỏ nhất của $\|Ax - b\|_p$, giả sử có $x, y \in L : \|Ax - b\|_p = \|Ay - b\|_p = m$ thì

$$\left\| A \frac{x+y}{2} - b \right\|_p = \left\| \frac{1}{2}(Ax - b) + \frac{1}{2}(Ay - b) \right\|_p \leq \frac{1}{2}\|Ax - b\|_p + \frac{1}{2}\|Ay - b\|_p = m$$

$\Rightarrow \left\| A \frac{x+y}{2} - b \right\|_p = m$, do đó bất đẳng thức trên phải trở thành đẳng thức, nghĩa là phải có:

$$Ax - b = k(Ay - b) \Rightarrow k = \pm 1$$

Nếu $k = 1 \Rightarrow x = y$

Nếu $k = -1 \Rightarrow A \frac{x+y}{2} - b = 0 \Rightarrow m = 0$, do A không suy biến nên phương trình $Ax - b = 0$ có nghiệm duy nhất $\Rightarrow x = y$

Vậy $\|Ax - b\|_p$ đạt min tại duy nhất một điểm thuộc $\mathbb{R}_n$

Chương 3:

CÁC ĐỊNH LÝ KANTOROVICH

Trong chương này chúng tôi nghiên cứu việc mở rộng phương pháp Newton đã nêu trong chương I, bởi các định lý của Kantorovich.

Trong các định lý sau sử dụng chung các giả thiết và ký hiệu sau: f là ánh xạ phi tuyến từ không gian Banach E vào không gian Banach F , $f'(x, \cdot)$ là ký hiệu của đạo hàm theo nghĩa Frechet, $f'(x, h)$ được ký hiệu $f'(x)h$, toán tử ngược $[f'(x)]^{-1}$ được ký hiệu $f'_{-1}(x)$.

Một số kết quả được dùng:

Bổ đề Banach: cho A là toán tử tuyến tính từ không gian Banach E vào chính nó, giả sử $\|A\| < 1$ thì $\|(I - A)^{-1}\| \leq (1 - \|A\|)^{-1}$

Kết quả: Tồn tại phiếm hàm tuyến tính f xác định trên không gian tuyến tính định chuẩn E , sao cho với $y \in E$ thì $f(y) = \|y\|$ và $\|f\| = 1$

Định lý 3.1

Cho x_o thuộc E , tại x_o tồn tại $f'_{-1}(x_o)$

Đặt $\eta_o = \|f'_{-1}(x_o)f(x_o)\|$, $S = \{x \in E : \|x - x_o\| \leq 2\eta_o\}$, $\beta_o = \|f'_{-1}(x_o)\|$

Nếu tồn tại số K sao cho $\|f'(x) - f'(y)\| \leq \frac{K\|x - y\|}{2}$, $\forall x, y \in S$ và

$\beta_o \eta_o K := h_o \leq \frac{1}{2}$ thì f có nghiệm trong S và dãy lặp $x_{k+1} = x_k - f'_{-1}(x_k)f(x_k)$

hội tụ về nghiệm x^* của f , hơn nữa $\|x_n - x^*\| \leq 2^{1-n} (2h_o)^{2^n - 1} \eta_o$

Chứng minh

* Trước tiên ta kiểm tra dãy $\{x_k\}$ được định nghĩa chứa trong S

$$\text{Rõ ràng là } x_1 \in S \text{ nên } \|f'(x_o) - f'(x_1)\| \leq \frac{K}{2} \|x_o - x_1\|$$

$$\text{mà } x_o - x_1 = f'_{-1}(x_o)f(x_o) \Rightarrow \|x - x_o\| = \eta_o \text{ nên } \|f'(x_o) - f'(x_1)\| \leq \frac{K}{2} \eta_o$$

$$\text{và } \|f'_{-1}(x_o)(f'(x_o) - f'(x_1))\| \leq \frac{\beta_o \eta_o K}{2} = \frac{h_o}{2} \leq \frac{1}{4}$$

Áp dụng bổ đề Banach thì $H := I - f'_{-1}(x_o)[f'(x_o) - f'(x_1)]$ là toán tử tuyến tính bị chặn từ E vào E và H có đảo thỏa:

$$\|H^{-1}\| \leq \left(1 - \|f'_{-1}(x_o)[f'(x_o) - f'(x_1)]\|\right)^{-1} \leq \left(1 - \frac{h_o}{2}\right)^{-1}$$

Ta có:

$$\begin{cases} *f'(x_o)H = f'(x_o)[I - f'_{-1}(x_o)(f'(x_o) - f'(x_1))] \\ \quad = f'(x_o) - (f'(x_o) - f'(x_1)) = f'(x_1) \\ *H^{-1}f'_{-1}(x_o) = f'_{-1}(x_1) \end{cases}$$

$$\Rightarrow \|f'_{-1}(x_1)\| \leq \|H^{-1}\| \|f'_{-1}(x_o)\| \leq \left(1 - \frac{h_o}{2}\right)^{-1} \beta_o < (1 - h_o)^{-1} \beta_o := \beta_1$$

Đặt $F_o(x) = x - f'_{-1}(x_o)f(x)$, tính đạo hàm hàm F_o ta nhận được

$$F'_o(x_o) = 0$$

$$\text{mà } f'_{-1}(x_o)f(x_1) = x_1 - F_o(x_1) \text{ và } x_1 = x_o - f'_{-1}(x_o)f(x_o) = F_o(x_o)$$

$$\begin{aligned} \Rightarrow f'_{-1}(x_o)f(x_1) &= F_o(x_o) - F_o(x_1) \\ &= F_o(x_o) - F_o(x_1) - F'(x_o)(x_o - x_1) := z \end{aligned}$$

Áp dụng kết quả trên đối với phân tử z ta được:

$$\exists g \in E^* (E^* = B(E, R)): g(z) = \|z\|, \|g\| = 1$$

f khả vi frechet trên $S \Rightarrow F_o$ khả vi Frechet trên S

Theo công thức Taylor ta có:

$$g(F_o(x_o) - F_o(x_1)) = (gF_o)'(\bar{x})(x_o - x_1), \text{ với } \bar{x} = \theta x_o + (1 - \theta)x_1, \theta \in (0;1)$$

$$\begin{aligned} \Rightarrow g(z) = \|z\| &= \left((gF_o)'(\bar{x}) - (gF_o)'(x_o) \right)(x_o - x_1) \\ &= g\left((F_o'(\bar{x}) - F_o'(x_o))(x_o - x_1) \right) \\ &\leq \|g\| \|F_o'(\bar{x}) - F_o'(x_o)\| \|x_o - x_1\| \end{aligned}$$

$$\text{Do } F_o'(\bar{x}) - F_o'(x_o) = f_{-1}'(x_o) [f'(x_o) - f'(\bar{x})]$$

$$\text{Nên } \|F_o'(\bar{x}) - F_o'(x_o)\| \leq \frac{\beta_o K \|x_o - \bar{x}\|}{2} \leq \frac{\beta_o K \|x_o - x_1\|}{2}$$

$$\text{Do đó: } g(z) = \|f_{-1}'(x_o) f(x_1)\| \leq \frac{\beta_o K \eta_o^2}{2} = \frac{h_o \eta_o}{2}$$

$$\text{Do } f_{-1}'(x_1) = H^{-1} f_{-1}'(x_o) \text{ nên } f_{-1}'(x_1) f(x_1) = H^{-1} f_{-1}'(x_o) f(x_1)$$

$$\Rightarrow \|f_{-1}'(x_1) f(x_1)\| \leq \left(1 - \frac{h_o}{2}\right)^{-1} \frac{h_o \eta_o}{2} \leq \frac{1}{2} \cdot \frac{h_o \eta_o}{1 - h_o} := \eta_1 \leq \eta_o$$

$$\text{Tương tự, đặt } h_1 = \beta_1 \eta_1 K = \frac{\beta_o}{1 - h_o} \cdot \frac{1}{2} \cdot \frac{h_o \eta_o}{1 - h_o} K = \frac{1}{2} \frac{h_o^2}{(1 - h_o)^2} \leq 2h_o^2 \leq \frac{1}{2}$$

Tổng quát ta xây dựng được các dãy số như sau:

$$\beta_n = \frac{\beta_{n-1}}{1 - h_{n-1}}; \quad \eta_n = \frac{h_{n-1} \eta_{n-1}}{2(1 - h_{n-1})}; \quad h_n = \frac{h_{n-1}^2}{2(1 - h_{n-1})^2}$$

$$\text{Và } F_n(x) = x - f_{-1}'(x_n) f(x)$$

Giả sử rằng mỗi $n = 1, 2, \dots, N$ các bất đẳng thức sau là đúng:

$$\|f_{-1}'(x_n) f(x_n)\| \leq \eta_n, \quad \|f_{-1}'(x_n)\| \leq \beta_n \text{ và } h_n = \beta_n \eta_n K \leq \frac{1}{2} \text{ và giả sử } x_n \in S$$

Ta kiểm tra $x_{N+1} \in S$

Do $x_n \in S$ và $h_n \leq \frac{1}{2} (n = \overline{1, N})$, theo chứng minh trên ta có:

$$\text{Từ } \eta_1 \leq \frac{1}{2} \eta_o \text{ (vì } \frac{h_o}{1-h_o} \leq 1) \Rightarrow \eta_n \leq \frac{1}{2} \eta_{n-1} (n = \overline{1, N})$$

$$\begin{aligned} \Rightarrow \|x_{n+1} - x_o\| &\leq \|x_{N+1} - x_N\| + \|x_N - x_{N-1}\| + \dots + \|x_1 - x_o\| \\ &\leq \eta_n + \eta_{N-1} + \dots + \eta_o = \sum_{i=0}^N \eta_i \leq \eta_o \sum_{i=0}^N \frac{1}{2^i} \leq 2\eta_o \end{aligned}$$

$$\Rightarrow x_{N+1} \in S$$

Lập lại chứng minh như ở trên, ta có: $h_{N+1} \leq \frac{1}{2}$, $\beta_{N+1} = (1 - h_N)^{-1} \beta_N$ và

$\eta_{N+1} \leq \frac{1}{2} \eta_N$ thì ta kết thúc phần lý luận quy nạp.

Ước lượng tương tự như trên ta được:

$$\|x_{n+p} - x_n\| = \eta_n \sum_{j=0}^{p-1} 2^{-j} \leq 2\eta_n, \quad \forall p = 1, 2, \dots$$

Do $\{\eta_n\} \rightarrow 0 \Rightarrow \{x_n\}$ là dãy cauchy, do E là không gian Banach

$$\Rightarrow \{x_n\} \rightarrow x^* \text{ trong E.}$$

* Ta chứng minh $f(x^*) = 0$

Ta có:

$$\|f'(x_n)\| \leq \|f'(x_o)\| + \|f'(x_n) - f'(x_o)\| \leq \|f'(x_o)\| + \frac{K}{2} \|x_n - x_o\| \leq \|f'(x_o)\| + K\eta_o \Rightarrow \{f'(x_n)\} \text{ là}$$

dãy bị chặn

$$\text{Mặt khác } f'(x_n)(x_{n+1} - x_n) = f(x_n)$$

$$\Rightarrow \{f(x_n)\} \rightarrow 0 \Rightarrow f(x^*) = 0$$

* Ước lượng tốc độ của giới hạn

Ta có:

$$h_1 = \frac{1}{2} \left(\frac{h_o}{1-h_o} \right)^2 \leq 2h_o^2; \quad h_2 \leq 2h_1^2 \leq 8h_o^4; \dots; \quad h_n \leq \frac{1}{2} (2h_o)^{2^n}$$

$$\begin{aligned} \eta_n &= \frac{1}{2} h_{n-1} (1-h_{n-1})^{-1} \eta_{n-1} \\ &\leq h_{n-1} \eta_{n-1} \leq h_{n-1} (h_{n-2} \eta_{n-2}) \leq h_{n-1} (h_{n-2} (h_{n-3} \eta_{n-3})) \\ &\leq \dots \leq \frac{1}{2} (2h_o)^{2^{n-1}} \frac{1}{2} (2h_o)^{2^{n-2}} \dots (2h_o) \eta_o = \frac{1}{2^n} (2h_o)^{2^n-1} \eta_o \end{aligned}$$

$$\text{Nên từ } \|x_{n+p} - x_n\| \leq 2\eta_n \Rightarrow \|x^* - x_n\| \leq 2^{-n+1} (2h_o)^{2^n-1} \eta_o$$

Vậy định lý được chứng minh xong $\square$

Nhận xét

Kết quả định lý 3.1 vẫn đúng khi thay quả cầu S bởi quả cầu S' có bán kính nhỏ hơn,

$$S' = \{x \in E : \|x - x_o\| \leq \eta_o N(h_o)\}, \quad N(h) = \frac{1 - \sqrt{1-2h}}{h}$$

Chứng minh

Ta chứng minh hệ thức $\eta_n = \eta_n N(h_n) - \eta_{n+1} N(h_{n+1})$

$$\text{Ta có: } \eta_{n+1} N(h_{n+1}) = \eta_{n+1} \frac{1 - \sqrt{1-2h_{n+1}}}{h_{n+1}} = \frac{1}{2} \frac{h_n \eta_n}{(1-h_n)} \left(\frac{1 - \sqrt{\frac{1-2h_n}{(1-h_n)^2}}}{\frac{1}{2} \frac{h_n^2}{(1-h_n)^2}} \right)$$

$$= \frac{\eta_n (1-h_n)}{h_n} \left(1 - \sqrt{\frac{1-2h_n}{(1-h_n)^2}} \right) = \frac{\eta_n}{h_n} (1-h_n - \sqrt{1-2h_n})$$

$$\begin{aligned}
&= \eta_n \frac{1 - \sqrt{1 - 2h_n}}{h_n} - \eta_n = \eta_n N(h_n) - \eta_n \\
&\Rightarrow \eta_n = \eta_n N(h_n) - \eta_{n+1} N(h_{n+1})
\end{aligned}$$

Ta kiểm tra dãy $\{x_n\}$ thuộc quả cầu tâm x_o bán kính $\eta_o N(h_o)$

như trong định lý 3.1 ta có:

$$\begin{aligned}
\|x_o - x_n\| &\leq \eta_o + \eta_1 + \eta_2 + \dots + \eta_{n-1} = \eta_o N(h_o) - \eta_n N(h_n) \leq \eta_o N(h_o) \\
&\Rightarrow \{x_n\} \in S'.
\end{aligned}$$

Định lý 3.2

Giả sử có các giả thiết trong định lý 3.1, nhưng quả cầu S được thay bởi quả

cầu S' có bán kính $\eta_o L(h_o)$ (với $L(h) = \frac{1 + \sqrt{1 - 2h}}{h}$) nhỏ hơn bán kính của quả cầu S,

$$S' = \{x \in: \|x_o - x\| \leq \eta_o L(h_o)\}$$

Thì phương trình $f(x) = 0$ có nghiệm duy nhất trong quả cầu này

Chứng minh

Tính toán như trong chứng minh của nhận xét 2, ta có hệ thức:

$$\eta_n = \eta_n L(h_n) - \eta_{n+1} L(h_{n+1})$$

$$\text{do đó: } L^2(h_n) = \left(\frac{1 + \sqrt{1 - 2h_n}}{h_n} \right)^2 = \frac{2(1 + \sqrt{1 - 2h_n} - h_n)}{h_n^2} = \frac{2}{h_n} [L(h_n) - 1]$$

$$= \frac{2}{h_n} \left[\frac{\eta_n + \eta_{n+1} L(h_{n+1})}{\eta_n} - 1 \right] = \frac{2\eta_{n+1}}{h_n \eta_n} L(h_{n+1})$$

$$\Rightarrow L^2(h_n) = \frac{2\eta_{n+1}}{h_n \eta_n} L(h_{n+1})$$

Gọi $\bar{x}$ là nghiệm của f trong S',

$$\text{do } \bar{x} \in S' \Rightarrow \|\bar{x} - x_o\| < \eta_o L(h_o)$$

$$\Rightarrow \exists \theta \in (0;1): \|\bar{x} - x_o\| = \theta \eta_o L(h_o)$$

Như trong định lý 3.1 ta có: $F_o(\bar{x}) = \bar{x}$

$$\begin{aligned} \Rightarrow \|\bar{x} - x_1\| &= \|F_o(\bar{x}) - F_o(x_o)\| \\ &= \|F_o(\bar{x}) - F_o(x_o) - F'_o(x_o)(\bar{x} - x_o)\| \end{aligned}$$

$$(\text{vì } F'_o(x_o) = 0)$$

Áp dụng bổ đề Banach:

$$\text{tồn tại } g \in E^*: g(z) = \|z\|, \|g\| = 1, z := F_o(\bar{x}) - F_o(x_o) - F'_o(x_o)(\bar{x} - x_o)$$

Áp dụng công thức Taylor:

$$g(F_o(\bar{x}) - F_o(x_o)) = (gF_o)'(\bar{x})(\bar{x} - x_o), \bar{x} = \theta x_o + (1 - \theta)\bar{x}, \theta \in (0;1)$$

$$\begin{aligned} \Rightarrow g(z) = \|z\| &= \left((gF_o)'(\bar{x}) - gF'_o(x_o) \right)(\bar{x} - x_o) \\ &= g(F'_o(\bar{x}) - F'_o(x_o))(\bar{x} - x_o) \Rightarrow g(z) \leq \|g\| \|F'_o(\bar{x}) - F'_o(x_o)\| \|\bar{x} - x_o\| \end{aligned}$$

$$F'_o(\bar{x}) - F'_o(x_o) = f'_{-1}(x_o)(f'(x_o) - f'(\bar{x}))$$

$$\Rightarrow \|F'_o(\bar{x}) - F'_o(x_o)\| \leq \frac{\beta_o K \|\bar{x} - x_o\|}{2} \leq \frac{\beta_o K \|\bar{x} - x_o\|}{2}$$

$$\Rightarrow g(z) \leq \frac{\beta_o K \|\bar{x} - x_o\|^2}{2} = \frac{\beta_o K \theta^2 L^2(h_o)}{2} = \theta^2 L(h_1) \eta_1$$

$$\text{Ta có: } \|\bar{x} - x_o\| = \theta L(h_o) \eta_o \text{ trong khi đó } \|\bar{x} - x_1\| \leq \theta^2 L(h_1) \eta_1$$

$$\text{Do quy nạp ta có: } \|\bar{x} - x_n\| \leq \theta^{2^n} L(h_n) \eta_n (**)$$

Từ định nghĩa của $L(h)$ ta có: $L(h_n)\eta_n < \frac{2\eta_n}{h_n} = \frac{2}{\beta_n K} < \frac{2}{\beta_o K}$

Nên từ $(**)\Rightarrow \{x_n\} \rightarrow \bar{x}$

Theo kết quả của định lý 3.1 ta có $\{x_n\} \rightarrow x^* \Rightarrow x^* = \bar{x}$

Vậy f có nghiệm duy nhất trong S' .

Định lý được chứng minh xong.

Bổ đề

Cho $x^*, x_1, x_o, N(h), F_o(x)$ được định nghĩa như trong định lý 3.1 và trong nhận xét ở trên, giả sử có các giả thiết của định lý 3.1, với $S = \{x \in E : \|x - x_o\| \leq \eta_o N(h_o)\}$. Nếu $x \in E : \|x - x^*\| \leq \|x_1 - x^*\|$ và $\|x - x_o\| \leq N(h_o)\eta_o$ thì $x' = F_o(x)$ thỏa

$$\|x' - x^*\| \leq \frac{h_o N(h_o)}{2} \|x - x^*\| \text{ và } \|x' - x_o\| \leq N(h_o)\eta_o$$

Chứng minh

* Ta chứng minh bất đẳng thức $\|x' - x^*\| \leq \frac{h_o N(h_o)}{2} \|x - x^*\|$

Ta có: $\|x' - x^*\| = \|F_o(x) - F_o(x^*)\|$ (do x^* là nghiệm của f)

$$\leq \sup_{\theta} \left\{ \|F_o'(\bar{x})\| \|x - x^*\|, \bar{x} = x + \theta(x^* - x), \theta \in (0; 1) \right\}$$

(định lý giá trị trung bình)

$$= \sup_{\theta} \left\{ \|F_o'(\bar{x}) - F_o'(x_o)\| \|x - x^*\| \right\} \quad (\text{do } F_o'(x_o) = 0)$$

$$= \sup_{\theta} \left\{ \|f'_{-1}(x_o)(f'(x_o) - f'(\bar{x}))\| \|x - x^*\| \right\}$$

$$\leq \frac{\beta_o K}{2} \|x - x^*\| \sup_{\theta} \left\{ \|x_o - \bar{x}\| \right\}$$

$$\text{Do: } \|\bar{x} - x_o\| = \|(1-\theta)x + \theta x^* - (1-\theta)x_o - \theta x_o\|$$

$$\begin{aligned} &\leq (1-\theta)\|x - x_o\| + \theta\|x^* - x_o\| \\ &\leq \max\{\|x - x_o\|, \|x^* - x_o\|\} \leq N(h_o)\eta_o \end{aligned}$$

$$\Rightarrow \|x' - x^*\| \leq \beta_o \frac{K}{2} N(h_o)\eta_o \|x - x^*\| = h_o N(h_o) \frac{\|x - x^*\|}{2}$$

* Ta chứng minh bất đẳng thức $\|x' - x_o\| \leq N(h_o)\eta_o$

$$\text{Ta có: } \|x' - x_o\| \leq \|x' - x_1\| + \|x_1 - x_o\|$$

$$\leq \|F_o(x) - F_o(x_o) - F_o'(x_o)(x - x_o)\| + \eta_o$$

như trong chứng minh của định lý 3.1 ta có:

$$\|x' - x_o\| \leq \frac{1}{2}\beta_o K \|x - x_o\|^2 + \eta_o \leq \frac{1}{2}\beta_o K [N(h_o)\eta_o]^2 + \eta_o$$

tương tự như trong chứng minh định lý 3.2 ta có:

$$N^2(h_o) = \frac{2}{h_o}(N(h_o) - 1) \quad (\text{như trong chứng minh 3.2})$$

$$\Rightarrow \frac{1}{2}\beta_o K N^2(h_o)\eta_o^2 = \eta_o N(h_o) - \eta_o \quad (\text{do } \frac{\beta_o K}{h_o} = \frac{1}{\eta_o})$$

$$\Rightarrow \|x' - x_o\| \leq N(h_o)\eta_o$$

Định lý 3.3

Giả sử có các giả thiết của bổ đề trên và định nghĩa dãy lặp: $x'_{n+1} = F_o(x'_n)$

thì $\{x'_n\} \rightarrow x^*$ với tốc độ của một cấp số nhân.

Chứng minh

Đặt $x := x_1$ thì x thỏa điều kiện của bổ đề trên, do đó

$$F_o(x) = F_o(x_1) = x' \text{ thỏa } \|x' - x^*\| \leq \frac{h_o N(h_o)}{2} \|x - x^*\| \text{ và } \|x' - x_o\| \leq N(h_o) \eta_o$$

Do $h_o N(h_o) < 1 \Rightarrow x'$ thỏa điều kiện của bổ đề

Đặt $x'_1 = x_1, x'_1 = x'_2 \Rightarrow x'_2$ thỏa điều kiện của bổ đề, cứ thế cho $x'_3, \dots$

Ta có: $F_o(x_1) = F_o(x'_1) = x'_2$ thỏa

$$\begin{cases} \|x'_2 - x^*\| \leq h_o N(h_o) \|x'_1 - x^*\| \leq \frac{(h_o N(h_o))^2}{2^2} \|x_1 - x^*\| \leq \left(\frac{1}{2}\right)^2 \|x_1 - x^*\| \\ \|x'_2 - x_o\| \leq N(h_o) \eta_o \end{cases}$$

tiếp tục quá trình ta được:

$$\|x'_{n+1} - x^*\| \leq \left(\frac{h_o N(h_o)}{2}\right)^n \|x_1 - x^*\| \leq \left(\frac{1}{2}\right)^n \|x_1 - x^*\|$$

$$\Rightarrow \{x'_n\} \rightarrow x^* \text{ với tốc độ của một cấp số nhân với công bội } q = \frac{1}{2}$$

Vậy định lý được chứng minh xong.

KẾT LUẬN

Bằng việc tìm hiểu phương pháp Newton chúng tôi đã trình bày cách xây dựng lời giải xấp xỉ của một phương trình (trong chương 2 có tìm hiểu nghiệm của một hệ phương trình).

Trong chương 2 vấn đề được giải quyết trên không gian hữu hạn chiều và trong chương 3 đã mở rộng trên không gian Banach với việc trình bày các định lý Kantorovich. Ở đây ngoài việc trình bày cách xây dựng lời giải xấp xỉ chúng tôi còn tìm hiểu việc đánh giá sai số giữa lời giải xấp xỉ và lời giải chính xác của phương trình.

TÀI LIỆU THAM KHẢO

Tiếng Việt

1. Dương Minh Đức (2005), *Giải tích hàm*, Nhà xuất bản đại học Quốc gia TP Hồ Chí Minh, Thành Phố Hồ Chí Minh.
2. Hoàng Tụy (1978), *Giải tích hiện đại*, tập 2, Nhà xuất bản Giáo dục, Thành Phố Hồ Chí Minh.
3. Lê Hoàn Hóa (1999), *Phép tính vi phân trên không gian Banach*, Giáo trình của lớp cao học ngành giải tích Đại học sư phạm TP Hồ Chí Minh, Thành Phố Hồ Chí Minh.
4. Phan Đức Chính (1978), *Giải tích hàm*, Nhà xuất bản đại học và trung học chuyên nghiệp, Hà Nội.
5. Haim Brezis (2002), *Giải tích hàm, Lý thuyết và ứng dụng*, Nguyễn Hội Nghĩa và Nguyễn Thành Long dịch, Nhà xuất bản đại học Quốc gia TP Hồ Chí Minh, Thành Phố Hồ Chí Minh.
6. Sze – Tsen Hu (1978), *Cơ sở giải tích toán học*, Phan Đức Chính dịch, Nhà xuất bản đại học và trung học chuyên nghiệp, Hà Nội.

Tiếng Anh

7. Allen A. Goldstein (1967), *Constructive Real Analysis*, Harper & Row, Publishers, New York.