

## VỀ MỘT PHƯƠNG PHÁP ĐIỀU KHIỂN CỰC TRỊ HỘI TỤ NHANH

Nguyễn Văn Mạnh

Đại học bách khoa Hà nội

### I. Mở đầu

Một trong những chỉ tiêu chất lượng quan trọng nhất của hệ thống điều khiển cực trị là độ tác động nhanh, tức là thời gian cần thiết để hệ thống đạt đến điểm tối ưu của hàm mục tiêu. Vì vậy vấn đề cải tiến tăng cường độ tác động nhanh của hệ thống điều khiển cực trị xưa nay luân luân có tính thời sự và được nhiều tác giả quan tâm [1,2,3,4]. Ngoài yếu tố trễ và quán tính do bản chất vật lý của đối tượng quyết định thì độ tác động nhanh của hệ thống phụ thuộc chủ yếu vào phương pháp tổ chức tìm kiếm tối ưu hay chính thuật toán tối ưu hoá cơ sở. Về các thuật toán tối ưu hoá có thể xem trong [5,7]. Hiện nay các phương pháp điều khiển cực trị dựa chủ yếu trên cơ sở các thuật toán đơn giản như thuật toán Gauss (tuyệt theo tọa độ), thuật toán Gradient, thuật toán hạ nhanh nhất, ... Những thuật toán kể trên tuy rất đơn giản song có một nhược điểm cơ bản là hội tụ chậm. Đặc biệt đối với các hàm mục tiêu có dạng "khe" rõ rệt (khi mặt đồng mức của hàm mục tiêu bị kéo dài thẳng hoặc cong) tốc độ hội tụ thực tế của các thuật toán này trở nên chậm chạp tới mức không thể chấp nhận được [5,6,7]. Điều đó làm cho chất lượng của quá trình điều khiển cực trị xấu, đôi khi còn có thể làm mất khả năng bám theo điểm tối ưu khi tọa độ của nó thay đổi không ngừng theo thời gian.

Ngày nay khi kỹ thuật của máy tính điện tử và vi xử lý đã phát triển tới mức có khả năng thực hiện mọi thuật toán tối ưu hoá phức tạp dưới dạng chương trình giải cứng trong bộ nhớ EPRCM. chúng tôi đề xuất ứng dụng một thuật toán điều khiển cực trị có độ hội tụ mạnh hơn rất nhiều so với các thuật toán nói trên. Thuật toán có tên gọi là "Vượt khe theo hướng Gradient trong không gian kéo dài" (gọi tắt là "VK-D") [6]. Theo kết quả nghiên cứu và thử nghiệm của chúng tôi đây

cùng là thuật toán hội tụ vào loại nhanh nhất trong số các thuật toán tối ưu hoá phi tuyến đã biết hiện nay đối với các hàm dạng khe nói chung [6].

## 2. Mô tả thuật toán "VK-D"

Giả sử cần cực tiểu hoá hàm mục tiêu  $J(U)$ ,  $J(U)$  có một điểm cực tiểu  $U_*$  và khả vi liên tục trong không gian Oclit  $n$  chiều:  $E^n$ , ( $U \in E^n$ ). Thuật toán tìm kiếm tối ưu VK - D được xây dựng tuân nguyên tắc lặp liên tiếp theo công thức dưới đây với chỉ số  $k$  tăng dần:

$$U_{k+1} = U_k + \alpha_{k+1} S_k, \quad S_k \in E^n, \quad \alpha_{k+1} > 0, \quad k = 1, 2, \dots, \quad (1)$$

trong đó  $U_k, U_{k+1} \in E^n$ , là điểm đầu và điểm cuối của bước lặp thứ  $k+1$ ;  $S_k$  là hướng chuyển động của tọa độ tìm kiếm, được xác định trên cơ sở toán tử kéo dãn không gian [8];  $\alpha_{k+1}$  là độ dài bước chuyển dịch, xác định theo nguyên lý vượt khe [6,9].

Bản chất vượt khe của thuật toán VK-D thể hiện ở chỗ điểm đầu  $U_k$  và điểm cuối  $U_{k+1}$  của mỗi bước lặp thứ  $k+1$  luôn nằm về hai phía của điểm cực tiểu của hàm mục tiêu trên hướng  $S_k$  bắt đầu từ  $U_k$ .

Cách tính véc tơ chuyển dịch  $S_k$  và độ dài bước  $\alpha_{k+1}$  trong mỗi chu trình lặp có thể tóm tắt dưới đây.

*Xác định bước chuyển dịch  $S_k$*

Ký hiệu  $J'(U_k)$  là Gradient của hàm mục tiêu  $J(U)$  tại điểm  $U_k \in E^n$ . Tại mỗi bước  $k+1$ , ( $k = 0, 1, \dots$ ), xác định ma trận  $B_{k+1}$ . Ở bước đầu tiên ( $k = 0$ ) cho  $B_1 = I$ ,  $I$  là ma trận đơn vị, và hướng  $S_0 = -J'(U_0)$ . Kể từ bước thứ hai trở đi  $S_k$  được tính theo công thức sau

$$r_k = g_k - g_{k-1} = B_k^T J'(U_k) - B_k^T J'(U_{k-1}), \quad k = 1, 2, \dots, \quad (2)$$

ở đây  $T$  là dấu chuyển vị ma trận,  $g_k$  là Gradient trong không gian bị kéo dãn qua phép biến đổi của ma trận  $B_k$

$$R_k = I + (1/\rho_k - 1) \frac{r_k r_k^T}{r_k^T r_k}, \quad \rho_k > 0, \quad (3)$$

$$B_{k+1} = B_k B_k. \quad (4)$$

trong đó  $R_k$  là toán tử kéo dãn không gian với hệ số kéo dãn  $\rho_k$  dọc theo hướng  $r_k$ , còn  $B_k$  thực hiện phép kéo dãn tích lũy trên các bước lặp. Sẽ gán  $B_{k+1} = I$  nếu

$$\|B_{k+1}^T J'(U_k)\| \leq \eta \|B_{k+1}\| \|J'(U_k)\|, \quad 1 > \eta > 0, \quad (5)$$

$\|\cdot\|$  là ký hiệu chuẩn ma trận và véc tơ trong không gian Oclit.

Cuối cùng hướng chuyển dịch ở bước  $k+1$  tính theo công thức

$$S_k = -B_{k+1} B_{k+1}^T J'(U_k). \quad (6)$$

*Xác định độ dài bước vượt khe*

Ta ký hiệu  $h_k(\alpha) = J(U_k + \alpha S_k)$ . Bước vượt khe  $\alpha_{k+1}$  được xác định theo điều kiện sau

$$\frac{\partial h_k(\alpha)}{\partial \alpha} \Big|_{\alpha=\alpha_{k+1}} > 0, \tag{7}$$

$$h_k(\alpha_{k+1}) \leq \min h_k(\alpha) + \delta_k, \tag{8}$$

$$\delta_k \geq 0, \lim_{k \rightarrow \infty} \delta_k = 0, \sum_{k=0}^{\infty} \delta_k = \delta < \infty. \tag{9}$$

Về mặt kỹ thuật bước vượt khe có thể xác định bằng cách cho véc tơ điều khiển  $U$  thay đổi dần theo hướng  $S_k$  với tốc độ không đổi cho tới khi thoả mãn bất đẳng thức (7). Ngay khi đó (8) cũng được thoả mãn. Trên quan điểm logic số bước vượt khe có thể xác định như sau:

*Bước 1:* Cho  $\alpha > 0$ .

Nếu  $h_k(\alpha) \geq h_k(0)$  thì gán  $\alpha = 0$ ,  $\beta = \alpha$ ; trái lại, tăng  $\alpha$  theo quy luật cấp số nhân với công bội  $q > 1$  cho đến khi  $h_k(\alpha) \leq h_k(q^\alpha)$  thì gán  $\beta = q\alpha$ . Sau đó chuyển sang bước 2.

*Bước 2.* Tính các giá trị:  $\theta = \alpha + \gamma(\beta - \alpha)$ ,  $h_k(\theta)$ ,  $0 < \gamma < 1$ .

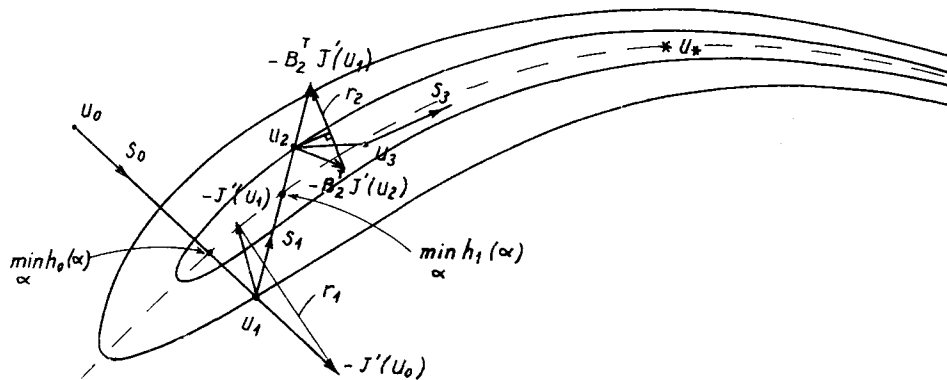
Nếu  $|\beta - \alpha| \leq \varepsilon$ ,  $\varepsilon > 0$ , chuyển sang bước 3.

-  $h_k(\theta) \leq h_k(\alpha)$  thì cho  $\alpha = \theta$  và làm lại bước 2

-  $h_k(\theta) > h_k(\alpha)$  thì cho  $\beta = \alpha$  và làm lại bước 2.

- trong trường hợp còn lại, chuyển sang bước 3.

*Bước 3* Cho  $\alpha = \theta$ . Cho  $\alpha_{k+1} = \theta$ . Bước vượt khe đã tìm xong.



Hình 1.

Điều kiện (7) thể hiện bước vượt khe, đảm bảo cho toán tử kéo dài  $R_k$  thực hiện theo phương hầu như vuông góc với lòng khe [6]. Khi đó các mặt đồng mức được mở ra và hướng chuyển động  $S_k$  của quỹ đạo tìm kiếm sẽ tiến gần về phía điểm tối ưu. Do đó tạo hiệu quả hội tụ rất cao của thuật toán. Điều kiện (8)-(9) đảm bảo cho thuật toán hội tụ đối với lớp rộng các hàm mục tiêu thường gặp trong thực tế. Dãy số  $\delta_k$   $k = 1, 2, \dots$  có thể chọn là một cấp số nhân với công bội nhỏ hơn 1 và số hạng ban đầu  $\delta_0$  là số dương thích hợp.

Quá trình tìm kiếm tối ưu theo thuật toán VK-D có thể giải thích trên hình vẽ (xem hình 1) đối với hàm mục tiêu hai biến  $J(U)$ ,  $U = [u_1, u_2]$ . Với điểm ban đầu  $U_0 = [u_{10}, u_{20}]$  quá trình tìm kiếm bước đầu tiên đi theo hướng  $S_0 = -J'(U_0)$  vượt qua điểm cực tiểu trên hướng  $S_0$  cho tới  $U_1$  sao cho thỏa mãn điều kiện (7)-(9). Từ điểm  $U_1$  bắt đầu thực hiện bước lặp thứ hai theo hướng  $S_1 = -B_2 B_1^T J'(U_1)$ .  $S_1$  là kết quả của hướng antigradien tại điểm  $U_1$  sau khi thực hiện phép co  $\rho_k$  lần thành phần hình chiếu của nó lên hướng  $r_1 = B_1^T (J'(U_1) - J'(U_0))$ . Sự chuyển dịch tiếp tục vượt qua điểm trùng nhất tương ứng với giá trị  $\min h_1(\alpha)$  của hàm mục tiêu cho tới điểm  $U_2$  nào đó thỏa mãn điều kiện vượt khe (7)-(9). Tương tự như vậy quá trình lặp tiếp tục bước thứ ba từ điểm  $U_2$ , ... cho đến khi thỏa mãn điều kiện dừng của thuật toán:

$$\|U_i - U_{i-1}\| < \varepsilon_1, \varepsilon_1 > 0, \quad (10)$$

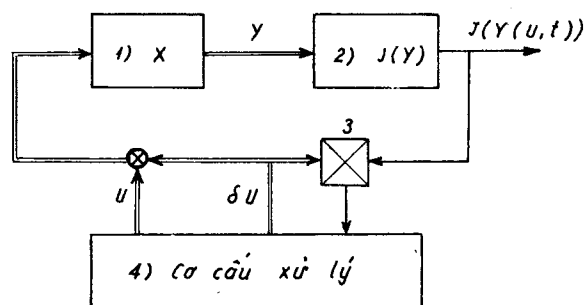
$$\|J'(U_i) - J'(U_{i-1})\| < \varepsilon_2, \varepsilon_2 > 0, \quad (11)$$

$$\|J(U_i) - J(U_{i-1})\| < \varepsilon_3, \varepsilon_3 > 0, \quad (12)$$

trong đó  $\|U\| = (\sum_{j=1}^n u_j^2)^{1/2}$  là chuẩn Oclit hay độ dài véc tơ trong  $L^n$ .

### 3. Mô hình hệ thống

Một phương án điều khiển hệ cực trị xây dựng trên cơ sở thuật toán VK-D dẫn trên hình 2.



Hình 2

Như thông lệ đối tượng điều khiển cực trị chia làm hai phần: phần đối tượng có quán tính và phần đối tượng phi tuyến không có quán tính. Phần đối tượng có quán tính nói chung có thể được mô tả bằng hệ phương trình trạng thái sau:

$$\dot{X} = F(X, U, t), \quad (13)$$

$$Y = H(X, U), \quad (14)$$

trong đó  $X$  là véc tơ trạng thái  $m$  chiều;  $U$  là véc tơ điều khiển  $n$  chiều;  $Y$  là véc tơ đầu ra  $n$  chiều;  $t$  là biến thời gian. Các hàm véc tơ  $F()$  và  $H()$  có thể phi tuyến hoặc tuyến tính. Phần đối tượng phi tuyến không có quán tính có thể là hàm bất kỳ nhưng phải có điểm cực trị.

Hệ thống làm việc theo nguyên tắc lặp liên tiếp. Trong mỗi chu trình lặp xảy ra tuần tự ba công đoạn chính sau

- *Công đoạn 1*: Tính véc tơ Gradient  $J'(U)$  của hàm mục tiêu bằng phương pháp lọc đồng bộ theo công thức

$$g_i = \frac{\partial J(U)}{\partial u_i} = \frac{2}{Ta^2} \int_{t_1}^{t_1+T} J(U + \delta U) a \sin((2i+1)\omega t) dt, \quad a = \text{constant} > 0, \quad (15)$$

$$\delta U = [\delta u_1, \delta u_2, \dots, \delta u_n]^T, \quad \delta u_i = a \sin((2i+1)\omega t), \quad i = 1, 2, \dots, n, \quad (16)$$

$$J'(U) = [g_1, \dots, g_n]^T \quad (17)$$

hoặc lọc kế tiếp (xem trong [3]). Công đoạn này chiếm khoảng 10% thời gian của chu trình.

- *Công đoạn 2*: Xác định véc tơ hướng tìm kiếm  $S_k$  theo các công thức (2)-(6). Công đoạn này chiếm thời gian không đáng kể vì chỉ phụ thuộc vào tốc độ của cơ cấu xử lý trung tâm 4.

- *Công đoạn 3*: Xác định điểm lặp kế tiếp theo nguyên tắc vượt khe (7)-(9). Công đoạn này chiếm phần lớn thời gian trong mỗi chu trình lặp. Đây cũng là công đoạn quan trọng nhất, phụ thuộc rất nhiều vào độ quán tính của đối tượng.

Trong các công đoạn 1 và 3 sẽ thay đổi véc tơ điều khiển  $U$  và nhận thông tin ở đầu ra qua giá trị hàm mục tiêu  $J$ . Dạng hệ thống cho trên hình 2 đã được thử nghiệm bằng số với nhiều dạng đối tượng khác nhau. Dưới đây sẽ dẫn kết quả của hai trường hợp.

#### 4. Tính toán thử nghiệm

*Ví dụ 1.*

Mô hình đối tượng có dạng tuyến tính với hai biến điều khiển

$$\begin{vmatrix} \dot{x}_1 \\ \dot{x}_2 \end{vmatrix} = \begin{vmatrix} -1 & -0.35 \\ -0.45 & -1 \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \end{vmatrix} + \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \begin{vmatrix} u_1 \\ u_2 \end{vmatrix}, \quad \begin{vmatrix} y_1 \\ y_2 \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \end{vmatrix}$$

$$J(Y(U, t)) = 100(y_1^2 - y_2)^2 + (y_1 + 1)^2.$$

Điểm xuất phát  $U_0 = [0.85, 0.46]^T$ ,  $J_0 = 24.2$ . Có một điểm tối ưu  $U_* = [1.35, 1.45]^T$ ,  $J_* = 0$ . Trạng thái ban đầu  $X(0) = [1, 1]$ . Đặc điểm hàm chỉ tiêu tối ưu  $J$  ở đây có tính chất khe rất rò rệt [7].

*Ví dụ 2.* Đối tượng cũng có dạng tuyến tính nhưng có 3 biến điều khiển

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{pmatrix} = \begin{pmatrix} -1 & 0 & -0.2 & 0 \\ 0 & -1 & -0.7 & -0.2 \\ 0.15 & 0 & -2 & 0 \\ 0 & -0.3 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 0 \\ -1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix},$$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

$$J(Y(U, t)) = 100(y_1^2 - y_2 + y_3)^2 + (y_2 - y_1 + y_3)^2 + (y_1 + y_2 - y_3)^2.$$

Điểm xuất phát  $U_0 = [0.3, 4.35, 2.1]^T$ ,  $J_0 = 215.997$ . Có một điểm tối ưu  $U_* = [0, 0, 0]$ ,  $J_* = 0$ . Trạng thái ban đầu  $X(0) = [1, -2.3, 4]^T$ . Đặc điểm của hàm chỉ tiêu này là hàm toàn phương có tính chất khe ít hơn so với hàm trên [7].

Đối với hai ví dụ trên quá trình tối ưu hoá được bắt đầu bằng việc đưa hệ thống trở về trạng thái xác lập ứng với véc tơ điều khiển ban đầu  $U_0$ . Giá trị hàm mục tiêu ứng với mỗi giá trị véc tơ điều khiển thay đổi  $U$  được xác định sau một thời gian xác lập  $t_x$  kể từ thời điểm đầu mỗi lần thay đổi  $U$  cho đến thời điểm khi tốc độ biến thiên của véc tơ trạng thái  $\|X(t)\|_{t=t_x}$  không lớn hơn một giá trị dương  $\varepsilon$  nào đó. Ở đây chọn  $\varepsilon = 10^{-6}$ . Bằng cách khác có thể xác định thời điểm xác lập theo tốc độ biến thiên đủ nhỏ của hàm mục tiêu  $|dJ/dt|_{t=t_x} \leq \varepsilon$ .

Trên mô hình số véc tơ trạng thái  $X(t)$  được xác định bằng cách giải hệ phương trình vi phân mô tả trạng thái theo phương pháp số (ví dụ phương pháp Ole hoặc phương pháp Runge-Kuta) bắt đầu bằng một trạng thái xác định nào đó đến trạng thái xác lập mới theo định nghĩa trên. Tổng cộng sự thay đổi biến thời gian  $t$  trong mỗi lần tích phân hệ phương trình trạng thái, tương đương với tổng giá trị thời gian thực cần thiết trong quá trình tối ưu hoá.

Kết quả thử nghiệm dẫn trong bảng 1 đối với ví dụ 1 và bảng 2 đối với ví dụ 2. Để dễ so sánh hiệu quả tối ưu hoá bằng thuật toán VK-D và các thuật toán thông dụng trong các bảng 1, 2 đồng thời dẫn ra kết quả tối ưu hoá tốt nhất của các phương án theo thuật toán Gradient thông thường. Các chỉ số để sánh gồm:  $N_b$  bước số lặp tối ưu hoá đã thực hiện;  $t$  thời gian thực hiện tương ứng sau sau bước thứ  $N_b$ ;  $\Delta J(t) = J(U(t)) - J(U_*)$  - độ sai lệch so với giá trị tối ưu thực sự của hàm mục tiêu đạt được sau  $N_b$  bước lặp ở thời điểm  $t$ ;  $\|\Delta U(t)\|$  - khoảng cách tới điểm tối ưu thực sự, đạt được sau bước  $N_b$  tại thời điểm  $t$ .

Bảng 1. Kết quả tối ưu hoá ví dụ 1

Theo thuật toán VK-D				Theo thuật toán Gradient			
$N_b$	$t$	$\Delta J(t)$	$\ \Delta V(t)\ $	$N_b$	$t$	$\Delta J(t)$	$\ \Delta U(t)\ $
0	13	24.2	3.2	0	13	21.2	3.2
5	116	2.51	1.83	200	184	2.67	1.67
10	205	0.76	1.26	70	526	0.88	1.37
15	276	0.24	0.85	1400	937	0.23	0.878
20	352	0.023	0.275	700	3587	0.023	0.316
25	410	0.00038	0.029	1000	4778	0.010	0.216

Bảng 2. Kết quả tối ưu hoá ví dụ 2

Theo thuật toán VK-D				Theo thuật toán Gradient			
$N_b$	$t$	$\Delta J(t)$	$\ \Delta U(t)\ $	$N_b$	$t$	$\Delta J(t)$	$\ \Delta U(t)\ $
0	12	215.997	5.196	0	12	215.977	5.196
1	71	28.47	3.135	1	202	27.33	3.127
2	120	0.03	0.123	20	1046	0.0496	0.235
3	130	0.0098	0.108	30	1432	0.0185	0.140
4	139	0.0065	0.118	150	1708	0.0185	0.139
5	141	0.00005	0.0025	500	2571	0.0183	0.128

Từ kết quả thí nghiệm ta thấy để đạt cùng một độ chính xác về giá trị hàm mục tiêu trong cả hai ví dụ trên, thuật toán VK-D cần số bước lặp ít hơn trên 30 lần thời gian tối ưu hoá ít hơn trên 10 lần so với thuật toán Gradient. Ngoài ra thuật toán VK-D hội tụ nhanh dần tới điểm tối ưu và luân luân đạt nghiệm với độ chính xác cao với thời gian ngắn trong cả trường hợp hàm mục tiêu có đặc tính khe rò nét (ví dụ 1) cũng như khi tính chất khe yếu hơn (ví dụ 2). Ngược lại, thuật toán Gradient thường bị tắc ở khe dù nó thể hiện ít hay nhiều. Thực tế qua quá trình tìm kiếm tối ưu theo thuật toán Gradient xảy ra quá lâu, và nói chung không cho phép nhận được nghiệm tối ưu với độ chính xác cao.

## 5. Kết luận

1- Về bản chất thuật toán VK-D là một thuật toán tối ưu hoá phi tuyến hội tụ mạnh. Có thể ứng dụng được trong các trường hợp hàm mục tiêu liên tục bất kỳ cho dưới dạng giải tích ẩn (như trong hai ví dụ trên) hoặc giải tích hiện.

2- Việc xác định hướng chuyển động  $S_k$  tuy có phức tạp hơn so với thuật toán thông dùng nhưng cách xác định bước vượt khe lại rất đơn giản và dễ dàng thực hiện về mặt kỹ thuật.

3- Đối với thuật toán VK-D đòi hỏi bộ nhớ cho ma trận kéo dài tích lũy  $B_{k+1}$  nhưng điều đó không làm giảm tính hấp dẫn của nó trong các bài toán điều khiển

cực trị, vì nói chung các bài toán điều khiển cực trị thực tế thường có số biến nhỏ. Bộ nhớ cần thiết dành cho ma trận  $B_{k+1}$  sẽ không đáng ngại.

4- Đối tượng điều khiển cực trị càng có độ trễ và quán tính lớn thì tính ưu việt của thuật toán VK-D thể hiện càng rõ rệt, đặc biệt trong các trường hợp hàm mục tiêu có dạng khe.

### Tài liệu tham khảo

1. Kazakevich V. V. & Rodov A. B., *Hệ thống tối ưu hoá động*, Nxb "Nenergia", Moxkva 1977 (tiếng Nga).
2. Alecxandrovskeyi N. M. & Egorov X. B., *Hệ thống điều khiển thích nghi các quá trình công nghệ phức tạp*. Nxb "Energia", Moxkva 1973 (tiếng Nga).
3. Kraxovskiy A. A., *Động học các hệ tự chỉnh liên tục*, Nxb "Nauka", Moxkva, 1963 (tiếng Nga).
4. Raxtryghyn L. A., *Các hệ thống điều khiển cực trị*, Nxb "Nauka", Moxkva 1974 (tiếng Nga).
5. Vaxylyev F. P., *Các phương pháp số giải bài toán cực trị*, Nxb "Nauka", Moxkva 1980 (tiếng Nga).
6. Nguyễn Văn Mạnh, *Xây dựng các phương pháp tối ưu hoá các hàm mục tiêu dạng khe trong AXU các quá trình năng lượng nhiệt*, Luận án PTS trường năng lượng Moxkva 1988 (tiếng Nga).
7. Rekhlaitis G. V., Ravindran A. & Ragsdell K. M., *Engineering Optimization, Methods and applications*, A Wiley-Interscience Publication, New-York - Chichester - Brisban - Toronto - Singapor. 1983.
8. Shor N. Z., *Các phương pháp cực tiểu hoá hàm không khả vi và ứng dụng*, Nxb "Naukova Dumka" Kiev 1979 (tiếng Nga).

### Abstract

On a fast converging algorithm for extremal control

*A very fast converging algorithm for extremal control systems is presented in this paper. The algorithm is called "Cleftover stepping in sequentially stretched space". The paper describes the optimization algorithm as well as, a model of extremal control systems and results of experimental computation.*

### Phụ lục

*chứng minh sự hội tụ của thuật toán VK-D*

**Bổ đề 1** *Giả sử  $J(U)$  là hàm cực tiểu hoá bị chặn dưới và khả vi liên tục thì hướng  $S_k$  xây dựng theo các công thức (2)-(6) luôn là hướng giảm của hàm mục tiêu, tức*



là luân có

$$J'(U_k)^T < 0, \quad \forall J'(U_k) \neq 0.$$

*Chứng minh.* Chú ý đến điều kiện (2)-(6) ta có

$$\begin{aligned} J'(U_k)^T S_k &= -J'(U_k)^T B_{k+1} B_{k+1}^T J'(U_k) \\ &= -(B_{k+1}^T J'(U_k))^T (B_{k+1}^T J'(U_k)) = -\|(B_{k+1}^T J'(U_k))^T\|^2 < 0. \end{aligned}$$

Bổ đề được chứng minh.

**Bổ đề 2.** Cùng theo giả thiết của đề 1 ta có bất đẳng thức

$$\|S_k\|^2 / [-S_k^T J'(U_k)] \leq \|B_{k+1}\|^2.$$

*Chứng minh.* Ta biến đổi vế trái của bất đẳng thức

$$\begin{aligned} \frac{\|S_k\|^2}{-S_k^T J'(U_k)} &= \frac{\|B_{k+1} B_{k+1}^T J'(U_k)\|^2}{(B_{k+1} B_{k+1}^T J'(U_k))^T J'(U_k)} \\ &= \frac{J'(U_k)^T B_{k+1} B_{k+1}^T J'(U_k)}{J'(U_k) B_{k+1} B_{k+1}^T J'(U_k)} = \frac{(B_{k+1}^T J'(U_k))^T B_{k+1}^T B_{k+1} (B_{k+1}^T J'(U_k))}{\|B_{k+1}^T J'(U_k)\|^2} \\ &\leq \frac{\|B_{k+1}^T J'(U_k)\| \|B_{k+1}^T B_{k+1}\| \|B_{k+1}^T J'(U_k)\|}{\|B_{k+1}^T J'(U_k)\|^2} = \|B_{k+1} B_{k+1}^T\| \leq \|B_{k+1}\|^2. \end{aligned}$$

Bổ đề được chứng minh.

**Định lý.** Giả  $J(U)$  bị chặn dưới, khả vi liên tục và gradient của nó thoả mãn điều kiện Lipschitz:  $\|J'(U) - J'(V)\| \leq L\|U - V\|$ ,  $U, V \in E^n$ ,  $L$  là hằng số. Khi đó với mọi điểm xuất phát  $U_0 \in E^n$  dãy điểm  $\{U_k\}$ ,  $k = 0, 1, \dots$  xây dựng theo thuật toán VK-D (xem các công thức (1)-(9)) hội tụ về điểm dừng, tức là

$$\lim_{k \rightarrow \infty} \|J'(U_k)\| = 0.$$

Ngoài ra nếu  $T(U)$  là hàm lồi chỉ có một điểm cực tiểu là  $U_*$ , tập hợp điểm

$$M(U_0) = \{U | J(U) \leq J(U_0) + \delta\}, \delta$$

$\delta$  chọn theo (9) giới nội thì dãy điểm  $\{U_k\}$ ,  $k = 1, 2, \dots$  là dãy cực tiểu hoá hội tụ về điểm đó.

*Chứng minh.* Từ điều kiện (8) ta có

$$h_k(0) - h_k(\alpha_{k+1}) \geq h_k(0) - \min_{\alpha} h_k(\alpha) - \delta_k,$$

hay là

$$h_k(0) - h_k(\alpha_{k+1}) + \delta_k \geq h_k(0) - h_k(\alpha).$$

Đổi lại ký hiệu ta có

$$J(U_k) - J(U_{k+1}) + \delta_k \geq J(U_k) - J(U_k + \alpha S_k)^T(\alpha S_k), \quad \forall \alpha > 0.$$

Ta biến đổi vế phải của bất đẳng thức này được tương tự như trong [5]

$$\begin{aligned} J(U_k) - J(U_k + \alpha S_k) &= - \int_0^1 (J'(U_k + \alpha S_k))^T(\alpha S_k) d\tau \\ &= \alpha S_k^T J'(U_k) + \alpha S_k^T J'(U_k) - \int_0^1 (J'(U_k + \alpha S_k))^T(\alpha S_k) d\tau \\ &= -\alpha S_k^T J'(U_k) - \int_0^1 (J'(U_k + \alpha S_k))^T(\alpha S_k) - J'(U_k))^T(\alpha S_k) d\tau \\ &\geq -\alpha S_k^T J'(U_k) - \int_0^1 L \|\tau \alpha S_k\| \|\alpha S_k\| d\tau = -\alpha S_k^T J'(U_k) - \alpha^2 L \|S_k\|^2 \int_0^1 \tau d\tau \\ &= -\alpha S_k^T J'(U_k) - \alpha L \|S_k\|^2 / 2 = -\alpha S_k^T J'(U_k) [1 - 0.5 \alpha L \|S_k\|^2 / (-S_k^T J'(U_k))] \\ &\geq -S_k^T J'(U_k) \alpha (1 - 0.5 \alpha L \|B_{k+1}\|^2). \end{aligned}$$

Vì bất đẳng thức trên đây đúng với mọi  $\alpha$  nên nếu thay  $\alpha = 1/(L\|B_{k+1}\|^2)$  ta sẽ được

$$\begin{aligned} J(U_k) - J(U_k + \alpha S_k) &= -0.5 S_k^T J'(U_k) / (L \|B_{k+1}\|^2) \\ &= 0.5 J'(U_k)^T B_{k+1} B_{k+1}^T J'(U_k) / (L \|B_{k+1}\|^2) \\ &\geq (0.5 \eta^2 / L) \|J'(U_k)\|^2. \end{aligned}$$

xem [5]. So sánh vế trái của bất đẳng thức trên ta được

$$J(U_k) - J(U_{k+1}) + \delta_k \geq (0.5 \eta^2 / L) \|J'(U_k)\|^2 \geq 0,$$

$$J(U_{k+1}) \leq J(U_k) + \delta_k, \quad \delta_k \text{ chọn theo (9).}$$

Vậy theo bổ đề 2 trang 101 trong [5] thì dãy số  $\{J(U_k)\}$ ,  $k \rightarrow \infty$  phải có một giới hạn hữu hạn:  $\lim J(U_k) \geq -\infty$ . Cũng như trong [5] ta được

$$\lim_{k \rightarrow \infty} [J(U_k) - J(U_{k+1})] = 0.$$

Kết hợp với bất đẳng thức vừa nhận được trên đây và cho  $k \rightarrow \infty$  (khi đó  $\delta_k \rightarrow 0$ ) ta được

$$\lim_{k \rightarrow \infty} [J(U_k) - J(U_{k+1})] \lim_{k \rightarrow \infty} (0.5 \eta^2 / L) \|J'(U_k)\|^2 = \lim_{k \rightarrow \infty} \|J'(U_k)\|^2 = 0.$$

Phần còn lại của định lý có thể suy ra một cách dễ dàng tương tự như trong [5].