



A SELF-ADAPTIVE CONJUGATE GRADIENT CQ PROJECTION METHOD FOR SPLIT FEASIBILITY PROBLEMS WITH APPLICATIONS

Nguyen The Vinh*

University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam

ARTICLE INFO

TYPE: Research Article

Received: 21/03/2023

Revised: 20/06/2023

Accepted: 01/07/2023

Published online: 15/08/2023

<https://doi.org/10.47869/tcsj.74.6.5>

* Corresponding author

Email: thevinhbn@utc.edu.vn; Tel: 0982161132

Abstract. The split feasibility problem is a convex minimization problem. It is an active field due to its effective applications in image restoration and signal processing. In this paper, by combining the inertial technique and gradient projection method, we proposed a CQ projection algorithm combined with the conjugate gradient method to solve the split feasibility problem in Hilbert spaces. Polyak's step size is used to improve the running time in the case of large-scale problems. Unlike the existing conjugate gradient projection methods, our new method is designed such that our algorithm only needs to compute one value of the objective function and one value of the gradient in each iteration. With suitable assumptions, we show that the algorithm converges weakly to a solution of the problem. Our method extends and improves on some recent results in this field. As an example application, we examine the performance of our method on the signal processing problem with synthetic data. The results of numerical simulations in the sparse recovery problem and the comparison with known algorithms show the effectiveness of the proposed algorithm.

Keywords: split feasibility problem, conjugate gradient method, Polyak's step size, CQ method, weak convergence, sparse recovery problem



PHƯƠNG PHÁP CHIỀU CQ TỰ THÍCH ỦNG VỚI HƯỚNG GRADIENT LIÊN HỢP GIẢI BÀI TOÁN CHẤP NHẬN TÁCH VÀ ỦNG DỤNG

Nguyễn Thế Vinh*

Trường Đại học Giao thông vận tải, Số 3 Cầu Giấy, Hà Nội, Việt Nam

THÔNG TIN BÀI BÁO

CHUYÊN MỤC: Công trình khoa học

Ngày nhận bài: 21/03/2023

Ngày nhận bài sửa: 20/06/2023

Ngày chấp nhận đăng: 01/07/2023

Ngày xuất bản Online: 15/08/2023

<https://doi.org/10.47869/tcsj.74.6.5>

* Tác giả liên hệ

Email: thevinhbn@utc.edu.vn; Tel: 0982161132

Tóm tắt. Bài toán chấp nhận tách là bài toán tối ưu lồi. Bài toán này đang được nghiên cứu mạnh vì những ứng dụng hiệu quả của nó trong xử lý ảnh và xử lý tín hiệu. Trong bài báo này, chúng tôi đề xuất một thuật toán dạng chiều CQ kết hợp với phương pháp gradient liên hợp để giải bài toán chấp nhận tách trong không gian Hilbert. Cỗ bước Polyak được sử dụng để giảm thời gian chạy trong trường hợp bài toán có số chiều lớn. Thuật toán của chúng tôi chỉ cần tính một lần giá trị của hàm và đạo hàm trong mỗi bước lặp. Với các giả thiết phù hợp, chúng tôi đã chứng minh được rằng thuật toán hội tụ yếu đến một nghiệm của bài toán chấp nhận tách. Phương pháp của chúng tôi mở rộng và cải tiến một số kết quả gần đây theo hướng nghiên cứu này. Các kết quả số trong bài toán khôi phục thưa và sự so sánh với các thuật toán đã biết chỉ ra tính hữu hiệu của thuật toán đề xuất.

Từ khóa: Bài toán chấp nhận tách, phương pháp gradient liên hợp, cỗ bước Polyak, phương pháp CQ, hội tụ yếu, bài toán khôi phục thưa.

© 2023 Trường Đại học Giao thông vận tải

1. ĐẶT VẤN ĐỀ

Cho C, Q lần lượt là các tập con lồi đóng khác rỗng của các không gian Hilbert thực $\mathcal{H}_1, \mathcal{H}_2$. Không mất tính tổng quát, trong các không gian Hilbert này, ta ký hiệu tích vô hướng là $\langle \cdot, \cdot \rangle$, chuẩn cảm sinh là $\|\cdot\|$, toán tử đồng nhất là I , và $A: \mathcal{H}_1 \rightarrow \mathcal{H}_2$ là toán tử tuyến tính bị chặn, A^* là toán tử liên hợp của A . **Bài toán chấp nhận tách** (Split feasibility problems, viết tắt là bài toán SFP) là bài toán

$$\text{tìm điểm } x^* \in C \text{ sao cho } Ax^* \in Q. \quad (1)$$

Ký hiệu tập nghiệm của bài toán là Ω , tức là

$$\Omega = \{x^* \in C : Ax^* \in Q\}.$$

Bài toán chấp nhận tách do Censor và Elfving [1] đưa ra lần đầu tiên vào năm 1994, có nhiều ứng dụng trong xử lý tín hiệu, xử lý ảnh,... (ví dụ xem trong [1-12]). Cụ thể, xét một phép biến đổi tuyến tính như sau: $y = Ax$, với A là ma trận lấy mẫu kích thước $M \times N$, véc tơ x là tín hiệu gốc có kích thước N , và y là véc tơ đo đạc được có kích thước là M . Thông thường, M nhỏ hơn N rất nhiều, làm thế nào để khôi phục lại x khi biết y và A ? Ta biết rằng, trong trường hợp này, ta có M phương trình để tìm N nghiệm, nhưng M nhỏ hơn N nên sẽ không có nghiệm duy nhất. Ta có thể minh họa trường hợp trên bằng phép biến đổi Fourier rời rạc (Discrete Fourier Transform) với y chỉ lấy M tần số trong số N tần số từ phép biến đổi Fourier đầy đủ. Câu hỏi đặt ra là liệu có thể khôi phục lại được tín hiệu x chỉ từ M tần số của tín hiệu y đo được? Điều thú vị là E. Candes, J. Romberg và T. Tao trong công trình [13] đã cho thấy câu trả lời cho câu hỏi trên là khẳng định trong trường hợp x là tín hiệu thưa (sparse), tức là x có độ dài N chỉ chứa l phần tử khác không. Cho tín hiệu thưa x có chiều dài N , trong đó bao gồm l phần tử khác không và $N - l$ phần tử còn lại, được xem như là bằng không. Thông qua phép chiếu với ma trận lấy mẫu A ta có véc tơ đo đạc y :

$$y = A.x + \eta,$$

trong đó η là hàm nhiễu Gauss có trung bình bằng không và phương sai σ^2 .

Để khôi phục tín hiệu gốc \bar{x} , vấn đề được đưa ra là giải quyết bài toán tối ưu hóa cực tiểu:

$$\bar{x} = \min_x \|x\|_1 \quad \text{với} \quad \|y - Ax\|_2^2 \leq \varepsilon. \quad (2)$$

Ta biết rằng bài toán (2) tương đương với bài toán sau:

$$\bar{x} = \min_x \|y - Ax\|_2^2 \quad \text{với} \quad \|x\|_1 \leq t, \quad (3)$$

với t là số thực dương thích hợp.

Nếu ta đặt $C = \{x \in \mathcal{H}_1 : \|x\|_1 \leq t\}$, $Q = \{y\}$ thì việc giải bài toán (3) có thể quy về giải bài toán SFP (1). Bài toán chấp nhận tách ban đầu được Censor và Elfving nghiên cứu trong không gian hữu hạn chiều, sau đó được nhiều nhà toán học khác nghiên cứu trong không gian vô hạn chiều (chẳng hạn xem trong các bài báo cùng các trích dẫn liên quan [2,5,6,8,9,11,12,14-19]).

Phương pháp lặp phổ biến nhất để giải bài toán SFP là thuật toán CQ do Byrne đề xuất trong [3]. Phương pháp này là trường hợp đặc biệt của thuật toán chiết gradient trong bài toán tối ưu lồi với giá trị tối ưu 0:

$$\min_{x \in C} f(x) := \frac{1}{2} \|Ax - P_Q Ax\|^2.$$

Như ta đã biết hàm f là lồi khả vi và $\nabla f(x) = A^*(I - P_Q)Ax$, $x \in \mathcal{H}_1$ là đơn điệu mạnh ngược với hằng số $\frac{1}{\|A\|^2}$, nghĩa là:

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \frac{1}{\|A\|^2} \|\nabla f(x) - \nabla f(y)\|^2 \quad \forall x, y \in \mathcal{H}_1.$$

Tuy nhiên, thuật toán CQ đòi hỏi phải biết chuẩn của toán tử tuyến tính bị chặn, điều này rõ ràng không dễ trong tính toán. Vì vậy, các phương pháp tự thích ứng đã được đề xuất, xem chặng hạn trong [9,18,20,21]. Đáng chú ý là López và cộng sự trong bài báo [9] đã xây dựng

một phương pháp chọn cỡ bước sao cho khi áp dụng thuật toán CQ thì không yêu cầu thông tin về chuẩn của toán tử. Lưu ý rằng phương pháp CQ chỉ thu được hội tụ yếu trong không gian Hilbert. Về các kết quả thú vị khác liên quan đến bài toán chấp nhận tách, có thể xem trong bài khảo cứu của Ansari và Rehan [22] và các kết quả gần đây trong [5,7,11,15-17,19].

Năm 2022, Huang và Liu [7] đã kết hợp thuật toán gradient đa bước của Kesornprom, Pholasa và Cholamjiak [8] và phương pháp gradient liên hợp của Sakurai và Iiduka [23] để đề xuất một thuật toán gradient liên hợp mới giải bài toán SFP và đưa ra một ứng dụng trong lĩnh vực xử lý tín hiệu. Tuy nhiên qua nghiên cứu kỹ bài báo này chúng tôi thấy có một số nhược điểm như sau: Khối lượng tính toán còn khá lớn (tính hai lần giá trị của hàm và gradient); phần chứng minh sự hội tụ của thuật toán có lỗ hổng từ sau công thức (19) (chỉ có thể suy ra dãy lặp là đơn điệu tựa Féjer chứ không thể đơn điệu Féjer như bài báo đã chỉ ra); giả thiết (C4) không đủ để đảm bảo các đại lượng M và \tilde{M} là hữu hạn. Dựa trên những phân tích đó, trong bài báo này chúng tôi đề xuất một thuật toán CQ với hướng gradient liên hợp mới và chứng minh sự hội tụ yếu của thuật toán đó. Thuật toán của chúng tôi khắc phục được các hạn chế của Thuật toán 1 trong [7], đồng thời cho một phương pháp hợp nhất các thuật toán dạng CQ sử dụng cỡ bước Polyak được trình bày trong các nghiên cứu gần đây, như trong [9,18,24].

Phân còn lại của bài báo được tổ chức như sau: Phần đầu của Mục 2 trình bày một số kết quả cơ bản để sử dụng cho việc chứng minh, phân tích hội tụ của thuật toán CQ với hướng gradient liên hợp trong phần sau của mục này. Kết quả số giải bài toán LASSO minh họa tính hiệu quả của thuật toán do chúng tôi đề xuất. Phần 3 là kết luận của bài báo.

2. NỘI DUNG CHÍNH

2.1. Kiến thức chuẩn bị

Cho dãy $\{x^k\} \subset \mathcal{H}_1$ và C là tập con lồi đóng khác rỗng của \mathcal{H}_1 . Sự hội tụ mạnh (tương ứng, yếu) của dãy $\{x^k\}$ tới x được kí hiệu là $x^k \rightarrow x$ (tương ứng $x^k \rightharpoonup x$). Ta kí hiệu $\omega_w(x^k)$ là tập tất cả các điểm tụ yếu của $\{x^k\}$:

$$\omega_w(x^k) := \{x \in \mathcal{H}_1: x^{k_j} \rightarrow x \text{ với mỗi dãy con } \{k_j\} \text{ của } \{k\}\}.$$

Với mỗi $x \in \mathcal{H}_1$, tồn tại một điểm gần x nhất trong C , kí hiệu $P_C(x)$ sao cho

$$\|x - P_C x\| = \min\{\|x - y\|: y \in C\}.$$

P_C được gọi là phép chiếu metric của \mathcal{H}_1 lên C . Tính chất cơ bản của P_C được cho trong Bô đê 2.1.

Bô đê 2.1. P_C thỏa mãn các bất đẳng thức sau:

- (i) $\langle x - P_C(x), y - P_C(x) \rangle \leq 0, \forall x \in \mathcal{H}_1, y \in C;$
- (ii) $\|P_C(x) - y\|^2 \leq \|x - y\|^2 - \|x - P_C x\|^2, \forall x \in \mathcal{H}_1, y \in C.$

Để chứng minh sự hội tụ của thuật toán đề xuất, ta cần khái niệm và kết quả sau.

Định nghĩa 2.1. ([25]) Dãy $\{x^k\} \subset \mathcal{H}_1$ gọi là đơn điệu tựa Féjer đối với tập S nếu tồn tại số $N_1 \in \mathbb{N}$ sao cho với mọi $z \in S$: $\|x^{k+1} - z\|^2 \leq \|x^k - z\|^2 + \sigma_k$ với mọi $k \geq N_1$, trong đó σ_k là dãy số dương sao cho $\sum_{k=0}^{\infty} \sigma_k < \infty$.

Bô đê 2.2. ([25]) Cho \mathcal{H}_1 là không gian Hilbert thực và $\{x^k\} \subset \mathcal{H}_1$ là đơn điệu tựa Féjer đối với tập lồi đóng khác rỗng $S \subset \mathcal{H}_1$. Khi đó

- (i) Với mỗi $z \in S$, tồn tại giới hạn $\lim_{k \rightarrow \infty} \|x^k - z\|$,
- (ii) Nếu mỗi điểm tụ yếu của $\{x^k\}$ thuộc S thì tồn tại $\bar{x} \in S$ sao cho $\{x^k\}$ hội tụ yếu đến \bar{x} .

2.2. Thuật toán và sự hội tụ

2.2.1. Thuật toán CQ tự thích ứng với hướng gradient liên hợp

Trong phần này chúng tôi sẽ đề xuất thuật toán CQ quán tính với hướng gradient liên hợp để giải bài toán chấp nhận tách (1).

- **Thuật toán 2.1. (Thuật toán CQ tự thích ứng với hướng gradient liên hợp)**

Bước khởi tạo: Lấy $\varepsilon > 0$; $0 \leq \theta < 1$ và các dãy dương $\{\rho_k\}$, $\{\varepsilon_k\}$, $\{\beta_k\}$, $\{\theta_k\}$ sao cho $0 < \rho_k < 4$, $0 < \theta_k, \beta_k < 1$ và $\sum_{k=0}^{\infty} \varepsilon_k < \infty$, $\sum_{k=0}^{\infty} \beta_k < \infty$. Lấy tùy ý x^0, x^1, d^0 trong \mathcal{H}_1 và đặt $k = 1$.

Bước k: Tính

$$\alpha_k = \begin{cases} \min \left\{ \theta, \frac{\varepsilon_k}{\|x^k - x^{k-1}\|} \right\} & \text{nếu } x^k \neq x^{k-1}, \\ \theta & \text{nếu } x^k = x^{k-1}. \end{cases} \quad (4)$$

$$w^k = x^k + \alpha_k(x^k - x^{k-1}),$$

$$\lambda_k = \rho_k \frac{f(w^k)}{\|\nabla f(w^k)\|^2 + \theta_k}, \quad (5)$$

$$\begin{aligned} d^{k+1} &= -\lambda_k \nabla f(w^k) + \beta_k d^k, \\ x^{k+1} &= P_C(w^k + d^{k+1}). \end{aligned} \quad (6)$$

Nếu $f(x^{k+1}) < \varepsilon$ thì dừng thuật toán và x^{k+1} là nghiệm của bài toán. Trái lại, tăng $k := k + 1$ và tính như bước k .

Nhận xét 2.1. (1) Từ cách xác định α_k trong Thuật toán 2.1 ta suy ngay ra rằng

$$\sum_{k=0}^{\infty} \alpha_k \|x^k - x^{k-1}\| \leq \varepsilon_k < \infty.$$

(2) Trường hợp $\alpha_k = 0$ và $d^k = 0$ thì thuật toán trên quy về Thuật toán 3.1 của López và cộng sự [9]. Nếu $d^k = 0$ thì ta nhận được phương pháp CQ quán tính đã nghiên cứu trong [6,24]. Thuật toán của chúng tôi cải tiến của Huang và Liu [7] thể hiện ở hai điểm. Một là khôi lượng tính toán giảm, chỉ phải tính giá trị của hàm và gradient một lần trên mỗi bước lặp. Hai là chúng tôi dùng phương pháp quán tính để tăng tốc thuật toán.

2.2.2. Phân tích sự hội tụ

Để chứng minh sự hội tụ của Thuật toán 2.1 ta cần các giả thiết sau:

$$(C1) \inf_k \rho_k (4 - \rho_k) > 0,$$

$$(C2) \{x^k\} \text{ và } \{(I - P_Q)Aw^k\} \text{ là hai dãy bị chặn.}$$

Nhận xét 2.2. Nếu tập C là tập bị chặn thì do $\{x^k\} \subset C$ nên về trước của điều kiện (C2) hiển nhiên thỏa mãn. Chú ý nếu dãy $\{x^k\}$ bị chặn thì dẫn đến $\{w^k\}$ bị chặn. Vì thế nếu xét không gian Hilbert hữu hạn chiều thì vì $I - P_Q$ là ánh xạ không giãn nên ta cũng có $\{(I - P_Q)Aw^k\}$ bị chặn. Do đó, điều kiện (C2) không hề khiêm cưỡng vì trong các ví dụ thực tế như bài toán xử lý tín hiệu, hình ảnh thura, ta xét tập bị chặn C là quả cầu chuẩn 1, tức là

$$C = \{x \in \mathcal{H}_1 : \|x\|_1 \leq t\},$$

trong đó t là số thực dương.

Định lý 2.1. Giả sử các điều kiện (C1) và (C2) được thỏa mãn và $\{x^k\}$ là dãy sinh bởi **Thuật toán 2.1.** Khi đó dãy $\{x^k\}$ sẽ hội tụ yếu đến một phần tử của Ω .

Chứng minh:

Từ cách chọn β_k ta suy ra tồn tại k_0 sao cho $\beta_k \leq \frac{1}{2}$ với mọi $k \geq k_0$. Không giảm tổng quát, ta có thể coi $\|\nabla f(w^k)\| \geq \varepsilon$ với mọi $k \geq k_0$. Theo (C2) ta suy ra tồn tại $M_1 > 0$ sao cho

$$f(w^k) = \frac{1}{2} \|(I - P_Q)Aw^k\|^2 \leq M_1.$$

Ta sẽ chỉ ra bằng quy nạp rằng dãy $\{d^k\}$ bị chặn. Thật vậy, giả sử $\|d^k\| \leq M$ với k nào đó ($k \geq k_0$). Từ (5) ta thấy

$$\lambda_k \|\nabla f(w^k)\| = \rho_k \frac{f(w^k) \|\nabla f(w^k)\|}{\|\nabla f(w^k)\|^2 + \theta_k} < \frac{4M_1 \sqrt{2M_1}}{\varepsilon} \|A^*\|.$$

Vì thế $\sup_k \lambda_k \|\nabla f(w^k)\| < \infty$. Đặt

$$M = \max_k \left\{ \|d^k\|, 2 \sup_k \lambda_k \|\nabla f(w^k)\| \right\} < \infty.$$

Khi đó

$$\|d^{k+1}\| = \|-\lambda_k \nabla f(w^k) + \beta_k d^k\| \leq \lambda_k \|\nabla f(w^k)\| + \beta_k \|d^k\| \leq M.$$

Tiếp theo ta chứng minh rằng với mọi $z \in \Omega$ ta có:

$$\|x^{k+1} - z\|^2 \leq \|w^k - z\|^2 - \rho_k (4 - \rho_k) \frac{f^2(w^k)}{\|\nabla f(w^k)\|^2} + \beta_k M_2,$$

trong đó $M_2 = \sup_k 2 \langle w^k - \lambda_k \nabla f(w^k) + \frac{1}{2} \beta_k d^k - z, d^k \rangle$.

Ta có

$$\begin{aligned} \|w^k - \lambda_k \nabla f(w^k) - z\|^2 &= \|w^k - z\|^2 + \lambda_k^2 \|\nabla f(w^k)\|^2 \\ &\quad + 2\lambda_k \langle z - w^k, \nabla f(w^k) \rangle \end{aligned} \quad (7)$$

và

$$\begin{aligned} \langle \nabla f(w^k), w^k - z \rangle &= \langle (I - P_Q)Aw^k, Aw^k - Az \rangle \\ &= \langle (I - P_Q)Aw^k - (I - P_Q)Az, Aw^k - Az \rangle \\ &\geq \|(I - P_Q)Aw^k\|^2 = 2f(w^k). \end{aligned} \quad (8)$$

Từ (7), (8) ta có

$$\begin{aligned} \|w^k - \lambda_k \nabla f(w^k) - z\|^2 &\leq \|w^k - z\|^2 + \lambda_k^2 \|\nabla f(w^k)\|^2 - 4\lambda_k f(w^k) \\ &= \|w^k - z\|^2 + \left(\rho_k \frac{f(w^k)}{\|\nabla f(w^k)\|} \right)^2 - 4\rho_k \frac{f^2(w^k)}{\|\nabla f(w^k)\|^2} \\ &\leq \|w^k - z\|^2 - \rho_k (4 - \rho_k) \frac{f^2(w^k)}{\|\nabla f(w^k)\|^2}. \end{aligned}$$

Từ đó

$$\|x^{k+1} - z\|^2 \leq \|w^k - \lambda_k \nabla f(w^k) + \beta_k d^k - z\|^2$$

$$\begin{aligned}
 &= \|w^k - \lambda_k \nabla f(w^k) - z\|^2 + 2\beta_k \langle w^k - \lambda_k \nabla f(w^k) - z, d^k \rangle + \beta_k^2 \|d^k\|^2 \\
 &\leq \|w^k - z\|^2 - \rho_k(4 - \rho_k) \frac{f^2(w^k)}{\|\nabla f(w^k)\|^2} + 2\beta_k \langle w^k - \lambda_k \nabla f(w^k) - z, d^k \rangle + \beta_k^2 \|d^k\|^2 \\
 &= \|w^k - z\|^2 - \rho_k(4 - \rho_k) \frac{f^2(w^k)}{\|\nabla f(w^k)\|^2} + 2 \langle w^k - \lambda_k \nabla f(w^k) + \frac{1}{2}\beta_k d^k - z, \beta_k d^k \rangle \\
 &= \|w^k - z\|^2 - \rho_k(4 - \rho_k) \frac{f^2(w^k)}{\|\nabla f(w^k)\|^2} + 2\beta_k \langle w^k - \lambda_k \nabla f(w^k) + \frac{1}{2}\beta_k d^k - z, d^k \rangle \\
 &\leq \|w^k - z\|^2 - \rho_k(4 - \rho_k) \frac{f^2(w^k)}{\|\nabla f(w^k)\|^2} + \beta_k M_2,
 \end{aligned}$$

trong đó $M_2 = \sup_k 2 \langle w^k - \lambda_k \nabla f(w^k) + \frac{1}{2}\beta_k d^k - z, d^k \rangle$.

Do vậy

$$\begin{aligned}
 \|x^{k+1} - z\|^2 &\leq \|w^k - z\|^2 - \rho_k(4 - \rho_k) \frac{f^2(w^k)}{\|\nabla f(w^k)\|^2} + \beta_k M_2, \\
 &= (\|x^k - z\| + \alpha_k \|x^k - x^{k-1}\|)^2 - \rho_k(4 - \rho_k) \frac{f^2(w^k)}{\|\nabla f(w^k)\|^2} + \beta_k M_2.
 \end{aligned}$$

Do đó

$$\begin{aligned}
 \|x^{k+1} - z\|^2 &\leq \|x^k - z\|^2 + 2\|x^k - z\| \alpha_k \|x^k - x^{k-1}\| + (\alpha_k \|x^k - x^{k-1}\|)^2 \\
 &\quad - \rho_k(4 - \rho_k) \frac{f^2(w^k)}{\|\nabla f(w^k)\|^2} + \beta_k M_2, \tag{9}
 \end{aligned}$$

và

$$\begin{aligned}
 \rho_k(4 - \rho_k) \frac{f^2(w^k)}{\|\nabla f(w^k)\|^2} &\leq \|x^k - z\|^2 - \|x^{k+1} - z\|^2 + 2\|x^k - z\| \alpha_k \|x^k - x^{k-1}\| \\
 &\quad + (\alpha_k \|x^k - x^{k-1}\|)^2 + \beta_k M_2. \tag{10}
 \end{aligned}$$

Vì dãy $\{x^k\}$ bị chặn nên kết hợp với Nhận xét 2.1, bất đẳng thức (9) ta thấy rằng dãy này là dãy đơn điệu tựa Féjer đối với tập nghiệm. Khi đó theo Bố đề 2.2 với

$\sigma_k = 2\|x^k - z\| \alpha_k \|x^k - x^{k-1}\| + (\alpha_k \|x^k - x^{k-1}\|)^2 + \beta_k M_2$,
ta suy ra tồn tại $\lim_{k \rightarrow \infty} \|x^k - z\|$ và khi đó chuyển qua giới hạn trong (10) ta được

$$\lim_{k \rightarrow \infty} \rho_k(4 - \rho_k) \frac{f^2(w^k)}{\|\nabla f(w^k)\|^2} = 0.$$

Với giả thiết (C1) ta suy ra

$$\lim_{k \rightarrow \infty} \frac{f^2(w^k)}{\|\nabla f(w^k)\|^2} = 0. \tag{11}$$

Vì ∇f là Lipschitz nên ta có

$$\|\nabla f(w^k)\| = \|\nabla f(w^k) - \nabla f(z)\| \leq \|A\|^2 \|w^k - z\| \quad \forall z \in \Omega.$$

Từ đây ta suy ra dãy $\{\nabla f(w^k)\}$ bị chặn (hoặc có thể lập luận như phần đầu của chứng minh để suy ra $\|\nabla f(w^k)\| \leq \sqrt{2M_1}||A^*||$). Từ đó kết hợp với (11) ta đi đến

$$\lim_{k \rightarrow \infty} f(w^k) = 0. \quad (12)$$

Ta sẽ chỉ ra rằng $\omega_w(x^k) \subset \Omega$. Lấy tùy ý $\bar{x} \in \omega_w(x^k)$. Do dãy $\{x^k\}$ bị chặn nên tồn tại một dãy con $\{x^{k_l}\}$ hội tụ yếu đến \bar{x} . Mặt khác theo thuật toán thì $w^k = x^k + \alpha_k(x^k - x^{k-1})$ nên

$$\|w^k - x^k\| = \|\alpha_k(x^k - x^{k-1})\| \leq \varepsilon_k \rightarrow 0. \quad (13)$$

Từ (13) suy ra $w^{k_l} \rightharpoonup \bar{x} \in C$. Do $f(x)$ là hàm lồi nửa liên tục dưới yếu và $w^{k_l} \rightharpoonup \bar{x}$, từ (12) suy ra

$$f(\bar{x}) \leq \liminf_{k \rightarrow \infty} f(w^{k_l}) = 0.$$

Vậy ta có $A\bar{x} \in Q$. Điều này dẫn đến $\bar{x} \in \Omega$. Vì \bar{x} được chọn bất kỳ thuộc $\omega_w(x^k)$ nên $\omega_w(x^k) \subset \Omega$. Theo Bố đề 2.2 ta có điều cần chứng minh.

Nhận xét 2.3. Kết quả hội tụ trên của chúng tôi đã mở rộng và cải tiến các kết quả trong [6-9,24]. Chứng minh sự hội tụ của thuật toán đề xuất của chúng tôi cũng khác với các kết quả khác có sử dụng kỹ thuật quán tính [6,16-19].

Nhận xét 2.4. Nếu $\beta_k = 0$ với mọi k thì trong chứng minh của Định lý 2.1 sẽ không còn xuất hiện số hạng $\beta_k \langle w^k - \lambda_k \nabla f(w^k) + \frac{1}{2} \beta_k d^k - z, d^k \rangle$, vì thế điều kiện (C2) là không cần thiết nữa.

2.3. Kết quả số

Trong phần này, chúng tôi xét bài toán LASSO sau đây:

$$\min_{x \in R^n} \frac{1}{2} \|Ax - b\|_2^2 \text{ sao cho } \|x\|_1 \leq t, \quad (14)$$

trong đó $b \in R^m$ là tín hiệu quan sát được với nhiễu η , $x \in R^n$ là tín hiệu cần khôi phục, $A (m \times n)$ là ma trận lấy mẫu. Nói chung A là ma trận điều kiện xấu. Đặt

$$C := \{x: \|x\|_1 \leq t\} \text{ và } Q = \{b\}$$

thì bài toán tối ưu có ràng buộc (14) quy về bài toán chấp nhận tách (1).

Ta chọn dữ liệu như sau. Véc tơ $x \in R^n$ là một l -tín hiệu thưa được sinh ra từ phân phối Gauss với l phần tử khác 0 được chọn ngẫu nhiên. Ma trận $A (m \times n)$ được sinh ra từ phân phối chuẩn với giá trị kỳ vọng bằng 0, phương sai bằng 1. Chúng ta sẽ so sánh Thuật toán 2.1 (CGCQ) với thuật toán CQ [9] (Lopez et al. algorithm), thuật toán gia tốc của Gibali-Mai-Vinh (iCQ) [6], Thuật toán 1 của Kesornprom et al. [8] (CQ-gradient) và Thuật toán 1 của Huang-Liu [7] (CGCQ-gradient). Véc tơ quan sát được $b = Ax + \eta$, với nhiễu Gauss trắng η có phương sai bằng 1. Ta lấy

$m = 2048, n = 4096, l = 160$ (số thành phần khác 0),

$$\epsilon_k = \beta_k = \frac{1}{k^{1.1}}, \rho_k = 3.9, \theta = 0.6$$

và cùng điểm khởi tạo $x^0 = (0, \dots, 0)$ cho cả 3 thuật toán.

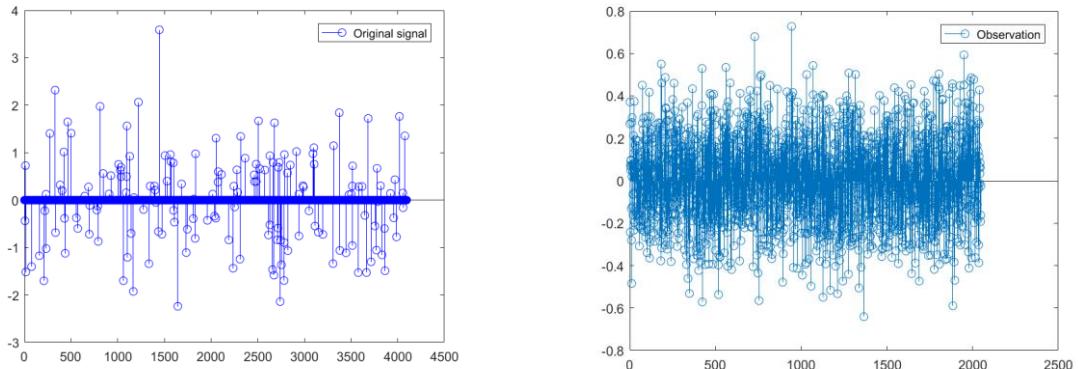
Tham số quan tính trong Thuật toán 2.1 và Thuật toán 1 của Gibali-Mai-Vinh được chọn như sau:

$$\alpha_k = \begin{cases} \theta, & x^k = x^{k-1}, \\ \min \left\{ \theta, \frac{1}{k^{1.1} \|x^k - x^{k-1}\|^2} \right\}, & x^k \neq x^{k-1}. \end{cases}$$

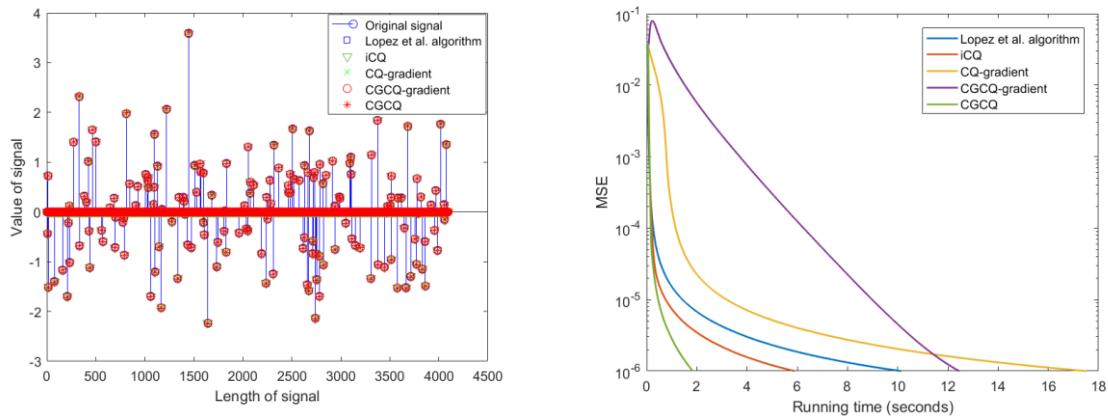
Các thuật toán dừng khi sai số bình phương trung bình

$$\text{MSE} = \frac{1}{n} \|x^k - x\|^2 < 10^{-6}$$

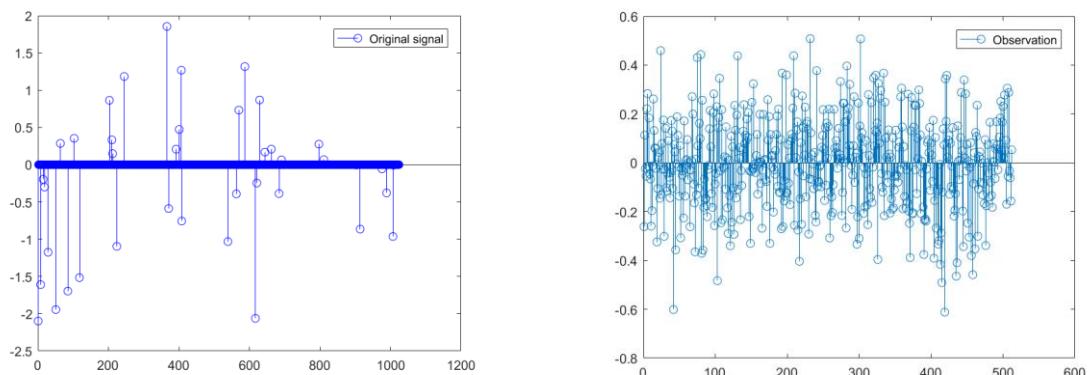
hoặc khi số bước lặp đạt đến 1000. Tín hiệu x được khôi phục bằng cách giải bài toán LASSO (14). Các kết quả số được tính toán trên ngôn ngữ Matlab 2016a. Dưới đây, Hình 1 và Hình 2 minh họa kết quả số cho trường hợp $m = 2048, n = 4096, l = 160$; Hình 3 và Hình 4 minh họa thêm kết quả số cho trường hợp $m = 512, n = 1024, l = 40$.



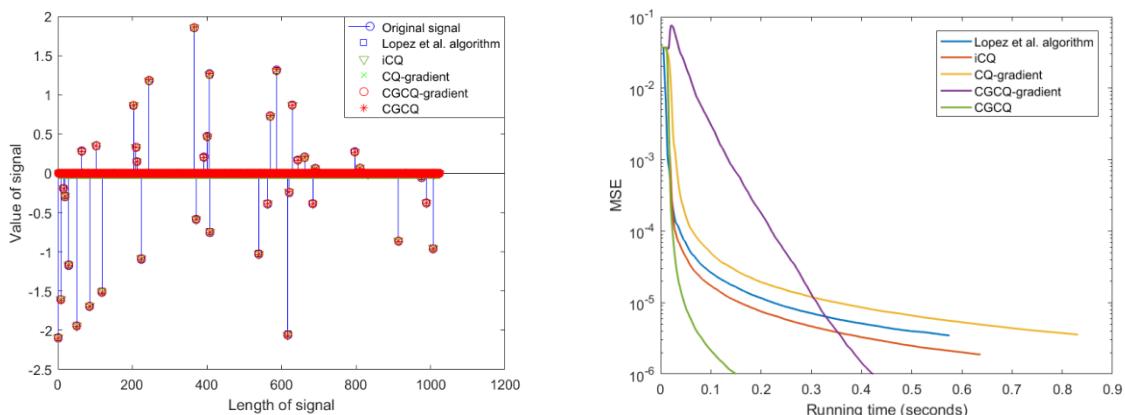
Hình 1. Tín hiệu gốc và tín hiệu quan sát được với $m = 2048, n = 4096, l = 160$.



Hình 2. Tín hiệu khôi phục và so sánh MSE (thời gian tính toán (giây)) trong trường hợp 1.



Hình 3. Tín hiệu gốc và tín hiệu quan sát được với $n = 1024, m = 512, l = 40$.



Hình 4. Tín hiệu khôi phục và so sánh MSE (thời gian tính toán (giây)) trong trường hợp 2.

Quan sát các kết quả chúng tôi thấy rằng các thuật toán đều cho kết quả khôi phục khá chính xác tín hiệu thừa ban đầu. Với cùng một mức sai số $\epsilon = 10^{-6}$, thuật toán của chúng tôi cho thời gian tính toán nhanh nhất, điều này cho thấy tín hiệu quả của việc xây dựng dãy véc tơ $\{d^k\}$. Từ Hình 4 (trường hợp 2), trong sự so sánh MSE chúng ta thấy ngoài thuật toán CGCQ và CGCQ-gradient, các thuật toán còn lại chưa đạt đến mức sai số mặc dù đã chạy hết 1000 bước lặp. Chúng ta cũng thấy rằng thuật toán CGCQ-gradient hiệu quả hơn khi số chiều nhỏ. Điều này cũng cho phép ta nhận định rằng với các thuật toán có nhiều bước tính toán cũng như phải tính giá trị hàm, đạo hàm nhiều lần thì thuật toán chạy tương đối chậm.

3. KẾT LUẬN

Trong bài báo này chúng tôi đã đề xuất một thuật toán CQ quán tính với hướng gradient liên hợp để giải bài toán chấp nhận tách (1) trong không gian Hilbert. Cột bước được tính bằng công thức Polyak, một phương pháp rất hiệu quả khi biết giá trị tối ưu. Chúng tôi chỉ ra rằng dãy sinh bởi thuật toán đề xuất hội tụ yếu đến nghiệm của bài toán chấp nhận tách. Thuật toán của chúng tôi đã cải tiến và mở rộng một số phương pháp CQ với cột bước Polyak được trình bày trong các nghiên cứu gần đây. Các kết quả số cho thấy tín hiệu quả khi so sánh với các thuật toán đã biết.

LỜI CẢM ƠN

Nghiên cứu này được tài trợ bởi Trường Đại học Giao thông Vận tải trong đề tài mã số T2023-CB-001.

TÀI LIỆU THAM KHẢO

- [1]. Y. Censor, T. Elfving, A multiprojection algorithm using Bregman projection in a product space, Numer. Algor., 8 (1994) 221-239. <https://doi.org/10.1007/BF02142692>
- [2]. P.K. Anh, N.T. Vinh, V.T. Dung, A new self-adaptive CQ algorithm with an application to the LASSO problem, J. Fixed Point Theory Appl., 20 (2018) 142. <https://doi.org/10.1007/s11784-018-0620-8>
- [3]. C. Byrne, Iterative oblique projection onto convex sets and the split feasibility problem, Inverse Probl., 18 (2002) 441–453. <https://doi.org/10.1088/0266-5611/18/2/310>
- [4]. Y. Censor, T. Bortfeld, B. Martin, A. Trofimov, A unified approach for inversion problems in intensity modulated radiation therapy, Phys. Med. Biol., 51 (2003) 2353-2365.
<https://doi.org/10.1088/0031-9155/51/10/001>
- [5]. Q.L. Dong, S. He, M.T. Rassias, General splitting methods with linearization for the split feasibility problem, J. Glob. Optim., 79 (2021) 813–836. <https://doi.org/10.1007/s10898-020-00963-3>
- [6]. A. Gibali, D.T. Mai, N.T. Vinh, A new relaxed CQ algorithm for solving split feasibility problems in Hilbert spaces and its applications, J. Ind. Manag. Optim., 15 (2019) 963-984.
<https://doi.org/10.3934/jimo.2018080>
- [7]. P. Huang, K. Liu, A new conjugate gradient algorithm for noise reduction in signal processing and image restoration, Front. Phys., 10 (2022) 1053353. <https://doi.org/10.3389/fphy.2022.1053353>
- [8]. S. Kesornprom, N. Pholasa, P. Cholamjiak, On the convergence analysis of the gradient-CQ algorithms for the split feasibility problem, Numer. Algor., 84 (2020) 997-1017. <https://doi.org/10.1007/s11075-019-00790-y>
- [9]. G. Lopez, V. Martin-Marquez, F. Wang, H.-K. Xu, Solving the split feasibility problem without prior knowledge of matrix norms, Inverse Problems, 28(2012) 085004. <https://doi.org/10.1088/0266-5611/28/8/085004>
- [10]. A. Moudaf, A. Gibali, $l_1 - l_2$ regularization of split feasibility problems, Numer. Algor., 78 (2018) 739–757. <https://doi.org/10.1007/s11075-017-0398-6>
- [11]. S. Suantai, B. Panyanak, S. Kesornprom, P. Cholamjiak, Inertial projection and contraction methods for split feasibility problem applied to compressed sensing and image restoration, Optim. Lett., 16 (2022) 1725–1744. <https://doi.org/10.1007/s11590-021-01798-x>
- [12]. F. Wang, Polyak's gradient method for split feasibility problem constrained by level sets, Numer. Algor., 77 (2018) 925-938. <https://doi.org/10.1007/s11075-017-0347-4>
- [13]. E. Candes, J. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Comm. Pure Appl. Math., 59 (2006) 1207-1223. <https://doi.org/10.1002/cpa.20124>
- [14]. N. Buong, Iterative algorithms for the multiple-sets split feasibility problem in Hilbert spaces, Numer. Algor., 76 (2017) 783–798. <https://doi.org/10.1007/s11075-017-0282-4>
- [15]. Q.L. Dong, L. Liu, X. Qin, J.-C. Yao, An alternated inertial general splitting method with linearization for the split feasibility problem, Optimization, (2022).
<https://doi.org/10.1080/02331934.2022.2069567>
- [16]. X. Ma, H. Liu, Inertial relaxed CQ algorithm for split feasibility problems with non-Lipschitz gradient operators, Optimization, 72 (2023) 1239-1260.
<https://doi.org/10.1080/02331934.2021.2010077>

- [17]. X. Ma, H. Liu, An inertial Halpern-type CQ algorithm for solving split feasibility problems in Hilbert spaces, *J. Appl. Math. Comput.*, 68 (2022) 1699-1717. <https://doi.org/10.1007/s12190-021-01585-y>
- [18]. N.T. Vinh, P.T. Hoai, Some subgradient extragradient type algorithms for solving split feasibility and fixed point problems, *Math. Meth. Appl. Sci.*, 39 (2016) 3808–3823.
<https://doi.org/10.1002/mma.3826>
- [19]. Z. Yao, S. Vong, Two inertial-type algorithms for solving the split feasibility problem, *Optimization*, (2022). <https://doi.org/10.1080/02331934.2022.2070066>
- [20]. B. Qu, N. Xiu, A note on the CQ algorithm for the split feasibility problem, *Inverse Prob.*, 21 (2005) 1655–1665. <https://doi.org/10.1088/0266-5611/21/5/009>
- [21]. Q. Yang, On variable-set relaxed projection algorithm for variational inequalities, *J. Math. Anal. Appl.*, 302 (2005) 166–179. <https://doi.org/10.1016/j.jmaa.2004.07.048>
- [22]. Q. H. Ansari, A. Rehan, Split feasibility and fixed point problems, in: *Nonlinear Analysis, Trends in Mathematics*, Springer, (2014) 281-322. https://doi.org/10.1007/978-81-322-1883-8_9
- [23]. K. Sakurai, H. Iiduka, Acceleration of the Halpern algorithm to search for a fixed point of a nonexpansive mapping, *Fixed Point Theory Appl.*, (2014) 202. <https://doi.org/10.1186/1687-1812-2014-202>
- [24]. F. Wang, H. Yu, An inertial relaxed CQ algorithm with an application to the LASSO and elastic net, *Optimization*, 70 (2021) 1101-1119. <https://doi.org/10.1080/02331934.2020.1763989>
- [25] Ya.I. Alber, A.N. Iusem, M.V. Solodov, On the projected subgradient method for nonsmooth convex optimization in a Hilbert space, *Math. Program.*, 81 (1998) 23-35.
<https://doi.org/10.1007/BF01584842>