RESEARCH

Open Access

CrossMark

Strong convergence of a modified proximal algorithm for solving the lasso

Ming Tian[1,2*] and Jun-Ying Gong[1]

*Correspondence:
tianming1963@126.com
[1]College of Science, Civil Aviation
University of China, Tianjin, 300300,
China
[2]Tianjin Key Laboratory for
Advanced Signal Processing, Civil
Aviation University of China, Tianjin,
300300, China

Abstract

As is well known the proximal iterative method can be used to solve the lasso of Tibshirani (J. R. Stat. Soc., Ser. B 58:267-288, 1996). In this paper, we first propose a modified proximal iterative method based on the viscosity approximation method to obtain strong convergence, then we apply this method to solve the lasso.

Keywords: lasso; l_1 regularization; Q-lasso; proximal method; variational inequality

1 Introduction

The lasso of Tibshirani [1] is formulated as the minimization problem

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|_2^2 \quad \text{subject to} \quad \|x\|_1 \le t, \tag{1.1}$$

where A is an $m \times n$ (real) matrix, $b \in \mathbb{R}^m$, $t \ge 0$ is a tuning parameter. The regularization minimization problem which is equivalent to (1.1) is

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|_2^2 + \gamma \|x\|_1, \tag{1.2}$$

where $\gamma > 0$ is a regularization parameter. As the ℓ_1 norm promotes the sparsity phenomenon that occurs in practical problems such as image/signal processing, machine learning and so on, the lasso has received much attention in recent years.
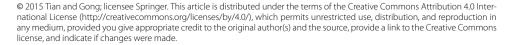
In fact, both (1.1) and (1.2) are equivalent to the basis pursuit (BP) of Chen *et al.* [2]:

$$\min_{x \in \mathbb{R}^n} \|x\|_1 \quad \text{subject to} \quad Ax = b.$$

So we mathematically study the inverse linear system in $\mathbb{R}^n$:

$$Ax = b, \tag{1.3}$$

where A is an $m \times n$ matrix, $b \in \mathbb{R}^m$ is an input, and $x \in \mathbb{R}^n$ stands for the image of interest to be recovered in imaging science. As $m \ll n$, the system (1.3) is underdetermined. Donoho [3], Candes, and others [4–6] pioneered the theory of compressed sensing showing that under certain conditions the underdetermined system (1.3) can determine a

Springer

unique k-sparse solution. Then Cipra [7] pointed out that the ℓ_1 norm is ideal as it ensures not only the parsimony of ℓ_0 but also the computation efficiency of ℓ_2 in sparse recovery.

However, the error of measurements always results in an inaccuracy of the system (1.3):

$$Ax = b + \epsilon.$$

In this case, the BP (1.3) is reformulated as

$$\min_{x \in \mathbb{R}^n} \|x\|_1 \quad \text{subject to} \quad \|Ax - b\|_2 \leq \epsilon, \tag{1.4}$$

where $\epsilon > 0$ is the tolerance level of errors and $\| \cdot \|$ is a norm on $\mathbb{R}^n$. If we let $Q := B_\epsilon(b)$ be the closed ball in $\mathbb{R}^n$ around b and with radius of ϵ, then (1.4) is rewritten as

$$\min_{x \in \mathbb{R}^n} \|x\|_1 \quad \text{subject to} \quad Ax \in Q. \tag{1.5}$$

As Q is a nonempty, closed, and convex subset of $\mathbb{R}^m$, we let P_Q be the projection from $\mathbb{R}^m$ onto Q. The condition $Ax \in Q$ is equivalent to the condition $Ax - P_Q(Ax) = 0$, so the problem (1.5) can be solved via

$$\min_{x \in \mathbb{R}^n} \|x\|_1 \quad \text{subject to} \ (I - P_Q)Ax = 0.$$

Applying the Lagrange method, we obtain the so-called Q-lasso:

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \left\| (I - P_Q)Ax \right\|_2^2 + \gamma \|x\|_1, \tag{1.6}$$

where $\gamma > 0$ is a Lagrangian multiplier.

The Q-lasso is connected with the split feasibility problem (SFP) of Censor and Elfving [8–10]. The SFP is mathematically formulated as the problem of finding a point x with the property:

$$x \in C \quad \text{and} \quad Ax \in Q, \tag{1.7}$$

where C and Q are nonempty, closed, and convex subset of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. An equivalent minimization formulation of the SFP (1.7) is given as

$$\min_{x \in C} \frac{1}{2} \|Ax - P_Q Ax\|_2^2.$$

The ℓ_1 regularization is given as the minimization problem

$$\min_{x \in C} \frac{1}{2} \|Ax - P_Q Ax\|_2^2 + \gamma \|x\|_1, \tag{1.8}$$

where $\gamma > 0$ is a regularization parameter. If the constrained set C is taken to be the entire space $\mathbb{R}^n$, then the problem (1.8) is equivalent to the problem (1.2).

Recently, Xu [11] exploited the following proximal algorithm:

$$x_{n+1} = \left(\text{prox}_{\lambda_n g} \circ (I - \lambda_n \nabla f) \right) x_n \tag{1.9}$$

to solve lasso (1.1), and Alghamdi *et al.* [12] also discussed an iterative algorithm for solving Q-lasso (1.6) via a proximal-gradient method. However, their iterative algorithms only obtain weak convergence.

Recall that Moudafi [13] proposed the viscosity iterative method in 2000 as follows:

$$x_{n+1} = \alpha_n f(x_n) + (1 - \alpha_n) T x_n, \quad n \geq 0, \tag{1.10}$$

where f is a contraction on a real Hilbert space, $\{\alpha_n\}$ is a sequence in $(0, 1)$. In 2004, Xu [14] proved that if $\{\alpha_n\}$ satisfies certain conditions, the sequence generated by (1.10) can converge strongly to a fixed point x^* of T, which is also the unique solution of the variational inequality

$$\langle (I - f) x^*, x - x^* \rangle \geq 0, \quad \text{for } x \in \text{Fix}(T).$$

In this paper, based on the viscosity iterative algorithm (1.10), we propose a modified formulation of the proximal algorithm (1.9). It is proved that the algorithm we propose can obtain strong convergence. Then we also apply this algorithm to solve the lasso and Q-lasso.

2 Preliminaries

Let H be a Hilbert space and let Γ_0 be the space of convex functions in H that are proper, lower semicontinuous, and convex.

Definition 2.1 The proximal operator of $\varphi \in \Gamma_0(H)$ is defined by

$$\text{prox}_\varphi(x) = \arg \min_{v \in H} \left\{ \varphi(v) + \frac{1}{2} \| v - x \|^2 \right\}, \quad x \in H.$$

The proximal operator of order $\lambda > 0$ is defined as the proximal operator of $\lambda \varphi$, that is,

$$\text{prox}_{\lambda \varphi}(x) = \arg \min_{v \in H} \left\{ \varphi(v) + \frac{1}{2\lambda} \| v - x \|^2 \right\}, \quad x \in H.$$

Proposition 2.2 *Let $\varphi \in \Gamma_0(H)$ and $\lambda \in (0, \infty)$.*

(i) $\text{prox}_{\lambda \varphi}$ *is firmly nonexpansive* (*hence nonexpansive*). *Recall that a mapping* $T : H \to H$ *is firmly nonexpansive if*

$$\| Tx - Ty \|^2 \leq \langle Tx - Ty, x - y \rangle, \quad x, y \in H.$$

(ii) $\text{prox}_{\lambda \varphi} = (I + \lambda \partial \varphi)^{-1} = J_\lambda^{\partial \varphi}$, *the resolvent of the subdifferential $\partial \varphi$ of φ.*

Combettes and Wajs [15] shows that the proximal operator $\text{prox}_{\lambda \varphi}$ can have a closed-form expression in some important cases, for example, if we take φ to be the norm of H, then

$$\text{prox}_{\lambda \| \cdot \|}(x) = \begin{cases} (1 - \frac{\lambda}{\|x\|})x, & \text{if } \|x\| > \lambda, \\ 0, & \text{if } \|x\| \leq \lambda. \end{cases}$$

In particular, if $H = \mathbb{R}$, then the above operator is reduced to the scalar soft-thresholding operator:

$$\text{prox}_{\lambda|\cdot|}(x) = \text{sgn}(x)\max\{|x| - \lambda, 0\}.$$

Lemma 2.3 [11] *The proximal identity*

$$\text{prox}_{\lambda\varphi} x = \text{prox}_{\mu\varphi}\left(\frac{\mu}{\lambda}x + \left(1 - \frac{\mu}{\lambda}\right)\text{prox}_{\lambda\varphi} x\right)$$

holds for $\varphi \in \Gamma_0(H)$, $x \in H$, $\lambda > 0$, and $\mu > 0$.

Recall that the function $H \to H$ is convex if

$$f\big((1 - \lambda)x + \lambda y\big) \le (1 - \lambda)f(x) + \lambda f(y)$$

for all $\lambda \in (0, 1)$ and $x, y \in H$. (Note that we consider finite-valued functions.)

The subdifferential of a convex function f is defined as the operator ∂f given by

$$\partial f(x) = \big\{\xi \in H : f(y) \ge f(x) + \langle \xi, y - x \rangle, y \in H\big\}. \tag{2.1}$$

The inequality in (2.1) is referred to as the subdifferential inequality of f at x. We say that f is subdifferentiable at x if $\partial f(x)$ is nonempty. We know that for an everywhere finite-valued convex function f on H, f is everywhere subdifferentiable.

Example
  (i) If $f(x) = |x|$ for $x \in \mathbb{R}$, then $\partial f(0) = [-1, 1]$;
  (ii) if $f(x) = \|x\|_1$, for $x \in \mathbb{R}^n$, then $\partial f(x)$ is given componentwise by

$$\big(\partial f(x)\big)_j = \begin{cases} \text{sgn}(x_j), & \text{if } x_j \ne 0, \\ \xi_j \in [-1, 1], & \text{if } x_j = 0, \end{cases}$$

  for $1 \le j \le n$.

Consider the unconstrained minimization problem:

$$\min_{x \in H} f(x). \tag{2.2}$$

Proposition 2.4 *Let f be everywhere finite-valued convex on H and $z \in H$. Support f is bounded below (i.e., $\inf\{f(x) : x \in H\} > -\infty$). Then z is a solution to minimization (2.2) if and only if it satisfies the first-order optimality condition*:

$$0 \in \partial f(z).$$

Lemma 2.5 [16] *Assume that $\{a_n\}_{n=0}^{\infty}$ is a sequence of nonnegative real numbers such that*

$$a_{n+1} \le (1 - \gamma_n)a_n + \gamma_n \delta_n + \beta_n, \quad n \ge 0,$$

where $\{\gamma_n\}_{n=0}^{\infty}$ and $\{\beta_n\}_{n=0}^{\infty}$ are sequence in $(0, 1)$ and $\{\delta_n\}_{n=0}^{\infty}$ is a sequence in $\mathbb{R}$ such that

    (i) $\sum_{n=0}^{\infty} \gamma_n = \infty$;

    (ii) *either* $\limsup_{n\to\infty} \delta_n \le 0$ *or* $\sum_{n=0}^{\infty} \gamma_n |\delta_n| < \infty$;

    (iii) $\sum_{n=0}^{\infty} \beta_n < \infty$.

Then $\lim_{n\to\infty} a_n = 0$.

Lemma 2.6 *Let H be a Hilbert space, then, for all $x, y \in H$, we have the inequality*

$$\|x + y\|^2 \le \|x\|^2 + 2\langle y, x + y \rangle.$$

Lemma 2.7 [17] *Let H be a Hilbert space, C a closed convex subset of H, and $T : C \to C$ a nonexpansive mapping with* $\text{Fix}(T) \ne \emptyset$. *If* $\{x_n\}_{n=1}^{\infty}$ *is a sequence in C weakly converging to x and if* $\{(I - T)x_n\}_{n=1}^{\infty}$ *converges strongly to y, then $(I - T)x = y$. In particular, if $y = 0$, then $x \in \text{Fix}(T)$.*

Lemma 2.8 [14] *Let H be a Hilbert space, $h : H \to H$ a contraction with coefficient $0 < \rho < 1$. Then*

$$\langle x - y, (I - h)x - (I - h)y \rangle \ge (1 - \rho)\|x - y\|^2, \quad x, y \in H.$$

That is, $I - h$ is strong monotone with coefficient $1 - \rho$.

We will use the notation $\rightharpoonup$ for weak convergence and $\to$ for strong convergence.

3 Strong convergence of proximal algorithms

Let H be a real Hilbert space and let Γ_0 be the space of convex functions in H that are proper, lower semicontinuous, and convex. Consider the following minimization problem:

$$\min_{x \in H} f(x) + g(x), \tag{3.1}$$

where $f, g \in \Gamma_0(H)$.

Proposition 3.1 [11] *Let $f, g \in \Gamma_0(H)$. Let $x^* \in H$ and $\lambda > 0$. Assume that f is finite-valued and differentiable on H. Then x^* is a solution to* (3.1) *if and only if x^* solves the fixed point equation*:

$$x^* = \left(\text{prox}_{\lambda g} \circ (I - \lambda \nabla f) \right) x^*.$$

Consider a mapping S_t on H defined by

$$S_t(x) = th(x) + (1 - t)\left(\text{prox}_{\lambda g} \circ (I - \lambda_t \nabla f) \right) x, \tag{3.2}$$

where h is a contraction with the coefficient $0 < \rho < 1$, $t \in (0, 1)$, $0 < \lambda_t \le \frac{2}{L}$. Assume that ∇f is L-Lipschitzian.

Proposition 3.2 *The mapping* $\text{prox}_{\lambda_t g} \circ (I - \lambda_t \nabla f)$ *is nonexpansive.*

Proof We show it in two cases.

Case 1: As $\lambda_t = \frac{2}{L}$, we have

$$
\left\| (I - \lambda_t \nabla f)x - (I - \lambda_t \nabla f)y \right\|^2
$$

$$
= \left\| \left(I - \frac{2}{L}\nabla f\right)x - \left(I - \frac{2}{L}\nabla f\right)y \right\|^2
$$

$$
= \left\langle x - y - \frac{2}{L}\big(\nabla f(x) - \nabla f(y)\big), x - y - \frac{2}{L}\big(\nabla f(x) - \nabla f(y)\big) \right\rangle
$$

$$
= \|x - y\|^2 + \frac{4}{L^2}\left\| \nabla f(x) - \nabla f(y) \right\|^2 - \frac{2}{L}\langle x - y, \nabla f(x) - \nabla f(y)\rangle
$$

$$
- \frac{2}{L}\langle \nabla f(x) - \nabla f(y), x - y \rangle
$$

$$
\leq \|x - y\|^2 + \frac{4}{L^2}\left\| \nabla f(x) - \nabla f(y) \right\|^2 - \frac{2}{L}\cdot\frac{1}{L}\left\| \nabla f(x) - \nabla f(y) \right\|^2
$$

$$
- \frac{2}{L}\cdot\frac{1}{L}\left\| \nabla f(x) - \nabla f(y) \right\|^2
$$

$$
\leq \|x - y\|^2.
$$

Hence,

$$
\left\| (I - \lambda_t)\nabla f(x) - (I - \lambda_t)\nabla f(y) \right\| \leq \|x - y\|.
$$

As the mapping $\mathrm{prox}_{\lambda_t}$ is nonexpansive, we get

$$
\left\| \big(\mathrm{prox}_{\lambda_t g} \circ (I - \lambda_t \nabla f)\big)x - \big(\mathrm{prox}_{\lambda_t g} \circ (I - \lambda_t \nabla f)\big)y \right\|
$$

$$
\leq \left\| (I - \lambda_t)\nabla f(x) - (I - \lambda_t)\nabla f(y) \right\|
$$

$$
\leq \|x - y\|.
$$

Case 2: $0 < \lambda_t < \frac{2}{L}$. We follow the proof of [16]. Since ∇f is L-Lipschitzian, ∇f is $(1/L)$-ism [18], which then implies that $\lambda_t \nabla f$ is $(1/\lambda_t L)$-ism. So $I - \lambda_t \nabla f$ is $(\lambda_t L/2)$-averaged. Since the proximal mapping $\mathrm{prox}_{\lambda_t g}$ is $(1/2)$-averaged, the composite $\mathrm{prox}_{\lambda_t g} \circ (I - \lambda_t)\nabla f$ is $((2 + \lambda_t L)/4)$-averaged for $0 < \lambda_t < 2/L$, so the mapping $\mathrm{prox}_{\lambda_t g} \circ (I - \lambda_t)\nabla f$ is nonexpansive. □

By Proposition 3.2 it is not hard to see that S_t is a contraction on H. For $x, y \in H$, we have

$$
\left\| S_t(x) - S_t(y) \right\|
$$

$$
= \left\| t\big[h(x) - h(y)\big] + (1 - t)\big[\big(\mathrm{prox}_{\lambda_t g} \circ (I - \lambda_t \nabla f)\big)x - \big(\mathrm{prox}_{\lambda_t g} \circ (I - \lambda_t \nabla f)\big)y\big] \right\|
$$

$$
\leq t\rho\|x - y\| + (1 - t)\|x - y\|
$$

$$
= \big(1 - t(1 - \rho)\big)\|x - y\|.
$$

Hence, S_t has a unique point, we denote it by x_t. Thus x_t is the unique solution of the fixed point equation

$$
x_t = th(x_t) + (1 - t)\big(\mathrm{prox}_{\lambda_t g} \circ (I - \lambda_t \nabla f)\big)x_t. \tag{3.3}
$$

We will use the following notation in Proposition 3.3 and Theorem 3.4:

$$V_t = V_{\lambda_t} = \text{prox}_{\lambda_t g} \circ (I - \lambda_t \nabla f);$$

$$V = V_\lambda = \text{prox}_{\lambda g} \circ (I - \lambda \nabla f).$$

The properties of V_t and V are helpful for the following proof [19].

Proposition 3.3 *Assume that* (3.1) *is consistent, and let S denote its solution set. Assume that* λ_t *is continuous with respect to t. Since* $0 < \lambda_t \le 2/L$, *we assume that* $\lambda_{t_j} \to \lambda$ ($t_j \to 0$) *and that* $\lambda > 0$. *Let* x_t *be defined by* (3.3), *we have*
   (i) $\{x_t\}$ *is bounded for* $t \in (0, 1)$;
   (ii) $\lim_{t\to 0} \|x_t - (\text{prox}_{\lambda g} \circ (I - \lambda \nabla f))x_t\| = 0$;
   (iii) x_t *defines a continuous curve from* $(0, 1)$ *into H.*

Proof (i) Take a $p \in S$, then we have $p \in \text{Fix}(V)$, and

$$\|x_t - p\|$$
$$= \|th(x_t) + (1 - t)V_t x_t - p\|$$
$$\le (1 - t)\|V_t x_t - p\| + t\|h(x_t) - p\|$$
$$= (1 - t)\|V_t x_t - V_t p\| + t\|h(x_t) - p\|$$
$$\le (1 - t)\|x_t - p\| + t\|h(x_t) - p\|.$$

It follows that

$$\|x_t - p\|$$
$$\le \|h(x_t) - p\|$$
$$\le \|h(x_t) - h(p)\| + \|h(p) - p\|$$
$$\le \rho\|x_t - p\| + \|h(p) - p\|.$$

Hence, $\|x_t - p\| \le \frac{1}{1-\rho}\|h(p) - p\|$, and $\{x_t\}$ is bounded, so are $\{\text{prox}_{\lambda g} \circ (I - \lambda \nabla f)x_t\}$ and $\{h(x_t)\}$.
   (ii) By the definition of $\{x_t\}$, we have, for any $\{t_j\} \to 0$,

$$\|x_{t_j} - (\text{prox}_{\lambda g} \circ (I - \lambda \nabla f))x_{t_j}\|$$
$$= \|t_j h(x_{t_j}) + (1 - t_j)V_{t_j} x_{t_j} - V x_{t_j}\|$$
$$\le t_j\|h(x_{t_j}) - V_{t_j} x_{t_j}\| + \|V_{t_j} x_{t_j} - V x_{t_j}\|$$

and

$$\|V_{t_j} x_{t_j} - V x_{t_j}\|$$
$$= \|(\text{prox}_{\lambda_{t_j} g} \circ (I - \lambda_{t_j} \nabla f))x_{t_j} - (\text{prox}_{\lambda g} \circ (I - \lambda \nabla f))x_{t_j}\|$$

$$
\begin{aligned}
&= \left\| \left(\mathrm{prox}_{\lambda g} \left(\frac{\lambda}{\lambda_{t_j}} (I - \lambda_{t_j} \nabla f) \right) x_{t_j} + \left(1 - \frac{\lambda}{\lambda_{t_j}} \right) (\mathrm{prox}_{\lambda_{t_j} g} \circ (I - \lambda_{t_j} \nabla f)) x_{t_j} \right) \right. \\
&\quad \left. - (\mathrm{prox}_{\lambda g} \circ (I - \lambda \nabla f)) x_{t_j} \right\| \\
&\le \left\| \frac{\lambda}{\lambda_{t_j}} (I - \lambda_{t_j} \nabla f) x_{t_j} + \left(1 - \frac{\lambda}{\lambda_{t_j}} \right) (\mathrm{prox}_{\lambda_{t_j} g} \circ (I - \lambda_{t_j} \nabla f)) x_{t_j} - (I - \lambda \nabla f) x_{t_j} \right\| \\
&= \left\| \left(\frac{\lambda}{\lambda_{t_j}} - 1 \right) x_{t_j} + \left(1 - \frac{\lambda}{\lambda_{t_j}} \right) (\mathrm{prox}_{\lambda_{t_j} g} \circ (I - \lambda_{t_j} \nabla f)) x_{t_j} \right\| \\
&= \left| 1 - \frac{\lambda}{\lambda_{t_j}} \right| \cdot \| x_{t_j} - V_{t_j} x_{t_j} \|.
\end{aligned}
$$

Then we obtain

$$
\| x_{t_j} - \mathrm{prox}_{\lambda g}(I - \lambda \nabla f) x_{t_j} \| \le t_j \| h(x_{t_j}) - V_{t_j} x_{t_j} \| + \left| 1 - \frac{\lambda}{\lambda_{t_j}} \right| \cdot \| x_{t_j} - V_{t_j} x_{t_j} \|. \tag{3.4}
$$

Since $\{x_t\}$, $\{h(x_t)\}$, and $\{V_t x_t\}$ are bounded, and $\lambda_{t_j} \to \lambda$ $(t_j \to 0)$, we can obtain by (3.4) that

$$
\lim_{t_j \to 0} \| x_{t_j} - \mathrm{prox}_{\lambda g} \circ (I - \lambda \nabla f) x_{t_j} \| = 0.
$$

By the arbitrariness of t_j, we get

$$
\lim_{t \to 0} \| x_t - \mathrm{prox}_{\lambda g} \circ (I - \lambda \nabla f) x_t \| = 0.
$$

(iii) For any given $t, t_0 \in (0,1)$,

$$
\begin{aligned}
&\| x_t - x_{t_0} \| \\
&\quad = \| t h(x_t) + (1 - t) V_t x_t - t_0 h(x_{t_0}) - (1 - t_0) V_{t_0} x_{t_0} \| \\
&\quad \le \| (t - t_0) h(x_t) + t_0 (h(x_t) - h(x_{t_0})) \| \\
&\qquad + \| (1 - t_0) V_{t_0} x_t - (1 - t_0) V_{t_0} x_{t_0} + (1 - t) V_t x_t - (1 - t_0) V_{t_0} x_t \| \\
&\quad \le |t - t_0| \cdot \| h(x_t) \| + t_0 \rho \| x_t - x_{t_0} \| + (1 - t_0) \| x_t - x_{t_0} \| \\
&\qquad + \| V_t x_t - V_{t_0} x_t \| + \| t V_t x_t - t_0 V_{t_0} x_t \| \\
&\quad \le |t - t_0| \cdot \| h(x_t) \| + t_0 \rho \| x_t - x_{t_0} \| + (1 - t_0) \| x_t - x_{t_0} \| + \left| 1 - \frac{\lambda_{t_0}}{\lambda_t} \right| \cdot \| x_t - V_t x_t \| \\
&\qquad + \| t_0 V_t x_t - t_0 V_{t_0} x_t \| + \| t V_t x_t - t_0 V_t x_t \| \\
&\quad \le \left(1 - t_0 (1 - \rho) \right) \| x_t - x_{t_0} \| + (t_0 + 1) \left| 1 - \frac{\lambda_{t_0}}{\lambda_t} \right| \cdot \| x_t - V_t x_t \| \\
&\qquad + |t - t_0| (\| h(x_t) \| + \| V_t x_t \|).
\end{aligned}
$$

Hence,

$$
\| x_t - x_{t_0} \| \le \frac{(\| h(x_t) \| + \| V_t x_t \|)}{t_0 (1 - \rho)} |t - t_0| + \frac{(t_0 + 1)}{t_0 (1 - \rho)} \left| 1 - \frac{\lambda_{t_0}}{\lambda_t} \right| \cdot \| x_t - V_t x_t \|. \tag{3.5}
$$

Since λ_t is continuous with respect to t, we can obtain by (3.5) that $x_t \to x_{t_0}$ as $t \to t_0$, that is, $\{x_t\}$ defines a continuous curve from $(0,1)$ into H. □

Theorem 3.4 *Let $\{x_t\}$ be defined by* (3.3). *Assume that the minimization problem* (3.1) *is consistent, and let S denote its solution set. Assume that $\lim_{t \to 0} \lambda_t = \lambda > 0$ and that λ_t is continuous with respect to t. Assume that ∇f is L-Lipschitzian. Then x_t converges as $t \to 0$ to a solution x^* of* (3.1), *which also solves the variational inequality*

$$\langle (I - h)x^*, \tilde{x} - x^* \rangle \geq 0, \quad \tilde{x} \in S. \tag{3.6}$$

Proof By Lemma 2.8 we know that $I - h$ is strongly monotone, so the variational inequality (3.6) has only one solution. Below we use $x^* \in S$ to denote the unique solution of (3.6).

To prove that $x_t \to x^*$ ($t \to 0$), we write, for a given $\tilde{x} \in S$,

$$x_t - \tilde{x} = th(x_t) + (1 - t)V_t x_t - \tilde{x} = t\big(h(x_t) - \tilde{x}\big) + (1 - t)(V_t x_t - \tilde{x});$$

$$\|x_t - \tilde{x}\|^2$$

$$= \langle x_t - \tilde{x}, x_t - \tilde{x} \rangle$$

$$= (1 - t)\langle V_t x_t - \tilde{x}, x_t - \tilde{x} \rangle + t\langle h(x_t) - \tilde{x}, x_t - \tilde{x} \rangle$$

$$\leq (1 - t)\|x_t - \tilde{x}\|^2 + t\langle h(x_t) - h(\tilde{x}), x_t - \tilde{x} \rangle + t\langle h(\tilde{x}) - \tilde{x}, x_t - \tilde{x} \rangle$$

$$\leq \big(1 - t(1 - \rho)\big)\|x_t - \tilde{x}\|^2 + t\langle h(\tilde{x}) - \tilde{x}, x_t - \tilde{x} \rangle.$$

Hence,

$$\|x_t - \tilde{x}\|^2 \leq \frac{\langle h(\tilde{x}) - \tilde{x}, x_t - \tilde{x} \rangle}{1 - \rho}. \tag{3.7}$$

Since $\{x_t\}$ is bounded as $t \to 0$, it is obvious that if $\{t_j\}$ is a sequence in $(0,1)$ such that $t_j \to 0$ ($j \to \infty$) and $x_{t_j} \rightharpoonup \bar{x}$, then by (3.7) we get $x_{t_j} \to \bar{x}$. Since $0 < \lambda_t \leq \frac{2}{L}$, we may assume that $\lambda_{t_j} \to \lambda \in (0, \frac{2}{L}]$ and that $\lambda > 0$, then by the proof of Proposition 3.2 we see that $\text{prox}_{\lambda g}(I - \lambda \nabla f)$ is also nonexpansive. Applying Lemma 2.7 and Proposition 3.3(ii), we get $\bar{x} \in S$.

Next, we show that $\bar{x} \in S$ solves the variational inequality (3.6). Indeed, we notice that x_t solves the fixed point equation

$$x_t = th(x_t) + (1 - t)\big(\text{prox}_{\lambda_t g} \circ (I - \lambda_t \nabla f)\big)x_t,$$

$$h(x_t) = \frac{1}{t}\big[x_t - (1 - t)V_t x_t\big],$$

$$(I - h)x_t = -\frac{1 - t}{t}(I - V_t)x_t.$$

Since V_t is nonexpansive, $I - V_t$ is monotone, so for any $\tilde{x} \in S$,

$$\langle (I - h)x_t, x_t - \tilde{x} \rangle$$

$$= -\frac{1 - t}{t}\langle (I - V_t)x_t, x_t - \tilde{x} \rangle$$

$$= -\frac{1 - t}{t}\langle (I - V_t)x_t - (I - V_t)\tilde{x}, x_t - \tilde{x} \rangle - \frac{1 - t}{t}\langle (I - V_t)\tilde{x}, x_t - \tilde{x} \rangle$$

$$\leq -\frac{1-t}{t}\langle (I - V_t)\tilde{x}, x_t - \tilde{x}\rangle$$

$$= 0.$$

Taking the limit through $t = t_n \to 0$, we obtain

$$\langle (I - h)\bar{x}, \bar{x} - \tilde{x}\rangle \leq 0.$$

Therefore $\bar{x} = x^*$ by uniqueness. □

Initialize $x_0 \in H$ and iterate

$$x_{n+1} = \alpha_n h(x_n) + (1 - \alpha_n)\big(\mathrm{prox}_{\lambda_n g} \circ (I - \lambda_n \nabla f)\big)x_n, \tag{3.8}$$

where $\{\alpha_n\}$ is a sequence in $(0, 1)$, $0 < \lambda_n \leq \frac{2}{L}$, $\liminf_{n \to \infty} \lambda_n > 0$, and $h : H \to H$ is a contraction with the coefficient $0 < \rho < 1$.

Theorem 3.5 *Let $f, g \in \Gamma_0(H)$. Assume that the minimization problem (3.1) is consistent and let S denote its solution set. Assume in addition that*

(C1) *∇f is L-Lipschitzian on H;*

(C2) *$\alpha_n \to 0$;*

(C3) *$\sum_{n=0}^{\infty} \alpha_n = \infty$;*

(C4) *$\sum_{n=0}^{\infty} |\alpha_{n+1} - \alpha_n| < \infty$;*

(C5) *$\sum_{n=0}^{\infty} |\lambda_{n+1} - \lambda_n| < \infty$.*

Then the sequence $\{x_n\}_{n=0}^{\infty}$ generated by (3.8) converges to x^* as defined in Theorem 3.4.

Proof Putting

$$V_n = V_{\lambda_n} = \mathrm{prox}_{\lambda_n g} \circ (I - \lambda_n \nabla f);$$

$$V = V_\lambda = \mathrm{prox}_{\lambda g} \circ (I - \lambda \nabla f).$$

We then get $x_{n+1} = \alpha_n h(x_n) + (1 - \alpha_n)V_n x_n$.

First we show that the sequence $\{x_n\}_{n=0}^{\infty}$ is bounded. Indeed, we have, for $\bar{x} \in S$,

$$\|x_{n+1} - \bar{x}\|$$

$$= \big\|\alpha_n h(x_n) + (1 - \alpha_n)V_n x_n - \bar{x}\big\|$$

$$= \big\|\alpha_n\big(h(x_n) - h(\bar{x})\big) + \alpha_n\big(h(\bar{x}) - \bar{x}\big) + (1 - \alpha_n)(V_n x_n - \bar{x})\big\|$$

$$\leq \alpha_n \rho \|x_n - \bar{x}\| + \alpha_n\big\|h(\bar{x}) - \bar{x}\big\| + (1 - \alpha_n)\|x_n - \bar{x}\|$$

$$= \big(1 - \alpha_n(1 - \rho)\big)\|x_n - \bar{x}\| + \alpha_n\big\|h(\bar{x}) - \bar{x}\big\|$$

$$\leq \max\big\{\|x_n - \bar{x}\|, \big\|h(\bar{x}) - \bar{x}\big\|/(1 - \rho)\big\}.$$

So, an induction argument shows that

$$\|x_n - \bar{x}\| \leq \max\big\{\|x_0 - \bar{x}\|, \big\|h(\bar{x}) - \bar{x}\big\|/(1 - \rho)\big\}, \quad n \geq 0.$$

We next prove that $\|x_{n+1} - x_n\| \to 0$ as $n \to \infty$. For the sake of simplicity, we assume that $0 < a \leq \lambda_n \leq b \leq \frac{2}{L}$.

We compute

$$\|x_{n+1} - x_n\|$$
$$= \left\| \alpha_n h(x_n) + (1 - \alpha_n) V_n x_n - \alpha_{n-1} h(x_{n-1}) - (1 - \alpha_{n-1}) V_{n-1} x_{n-1} \right\|$$
$$\leq \left\| \alpha_n \left[h(x_n) - h(x_{n-1}) \right] + \left[\alpha_n h(x_{n-1}) - \alpha_{n-1} h(x_{n-1}) \right] \right\|$$
$$\quad + \left\| (1 - \alpha_n)(V_n x_n - V_n x_{n-1}) + (1 - \alpha_n) V_n x_{n-1} - (1 - \alpha_{n-1}) V_{n-1} x_{n-1} \right\|$$
$$\leq \alpha_n \rho \|x_n - x_{n-1}\| + |\alpha_n - \alpha_{n-1}| \left\| h(x_{n-1}) \right\| + (1 - \alpha_n) \|x_n - x_{n-1}\|$$
$$\quad + (1 - \alpha_n) \|V_n x_{n-1} - V_{n-1} x_{n-1}\| + \|\alpha_n V_{n-1} x_{n-1} - \alpha_{n-1} V_{n-1} x_{n-1}\|$$
$$\leq \left(1 - \alpha_n(1 - \rho) \right) \|x_n - x_{n-1}\| + |\alpha_n - \alpha_{n-1}| \left(\left\| h(x_{n-1}) \right\| + \|V_{n-1} x_{n-1}\| \right)$$
$$\quad + \|V_n x_{n-1} - V_{n-1} x_{n-1}\|$$

and

$$\|V_n x_{n-1} - V_{n-1} x_{n-1}\| \leq \left| 1 - \frac{\lambda_n}{\lambda_{n-1}} \right| \cdot \|x_{n-1} - V_n x_{n-1}\|$$
$$\leq \frac{|\lambda_n - \lambda_{n-1}|}{a} \|x_{n-1} - V_n x_{n-1}\|,$$

so we obtain

$$\|x_{n+1} - x_n\| \leq \left(1 - \alpha_n(1 - \rho) \right) \|x_n - x_{n-1}\| + M \left(|\alpha_n - \alpha_{n-1}| + |\lambda_n - \lambda_{n-1}| \right), \qquad (3.9)$$

where $M \leq \max\{\|h(x_{n-1})\| + \|V_{n-1} x_{n-1}\|, \|x_{n-1} - V_n x_{n-1}\|/a\}$. By assumptions (C3)-(C5) in the theorem, we have $\sum_{n=0}^{\infty} \alpha_n = \infty$, and $\sum_{n=1}^{\infty}(|\alpha_n - \alpha_{n-1}| + |\lambda_n - \lambda_{n-1}|) < \infty$. Hence, Lemma 2.5 is applicable to (3.9) and we conclude that $\|x_{n+1} - x_n\| \to 0$.

Since $\{x_n\}$ is bounded, there exists a subsequence $\{x_{n_j}\}$ such that $x_{n_j} \rightharpoonup z$; below we will prove that $z \in S$. Since $0 < \lambda_n \leq \frac{2}{L}$, we may assume that $\lambda_{n_j} \to \lambda$. We have

$$\left\| x_{n_j} - \text{prox}_{\lambda g} \circ (I - \lambda \nabla f) x_{n_j} \right\| \leq \|x_{n_j} - x_{n_{j+1}}\| + \left\| x_{n_{j+1}} - \text{prox}_{\lambda g} \circ (I - \lambda \nabla f) x_{n_j} \right\|. \qquad (3.10)$$

We compute

$$\left\| x_{n_{j+1}} - \text{prox}_{\lambda g} \circ (I - \lambda \nabla f) x_{n_j} \right\|$$
$$= \left\| \alpha_{n_j} h(x_{n_j}) + (1 - \alpha_{n_j}) \left(\text{prox}_{\lambda_{n_j} g} \circ (I - \lambda_{n_j} \nabla f) \right) x_{n_j} - \text{prox}_{\lambda g} \circ (I - \lambda \nabla f) x_{n_j} \right\|$$
$$\leq \alpha_{n_j} \left\| h(x_{n_j}) - V x_{n_j} \right\| + (1 - \alpha_{n_j}) \|V_{n_j} x_{n_j} - V x_{n_j}\|$$

and

$$\|V_{n_j} x_{n_j} - V x_{n_j}\|$$
$$= \left\| \text{prox}_{\lambda_{n_j} g} \circ (I - \lambda_{n_j} \nabla f) x_{n_j} - \text{prox}_{\lambda g} \circ (I - \lambda \nabla f) x_{n_j} \right\|$$

$$\leq \left\| \frac{\lambda}{\lambda_{n_j}} (I - \lambda_{n_j} \nabla f) x_{n_j} + \left(1 - \frac{\lambda}{\lambda_{n_j}}\right) \left(\text{prox}_{\lambda_{n_j} g} \circ (I - \lambda_{n_j} \nabla f) \right) x_{n_j} - (I - \lambda \nabla f) x_{n_j} \right\|$$

$$= \left\| \left(\frac{\lambda}{\lambda_{n_j}} - 1 \right) x_{n_j} + \left(1 - \frac{\lambda}{\lambda_{n_j}}\right) \text{prox}_{\lambda_{n_j} g} (I - \lambda_{n_j} \nabla f) x_{n_j} \right\|$$

$$= \left| 1 - \frac{\lambda}{\lambda_{n_j}} \right| \cdot \| x_{n_j} - V_{n_j} x_{n_j} \|$$

$$\leq |\lambda_{n_j} - \lambda| \frac{\| x_{n_j} - V_{n_j} x_{n_j} \|}{a}.$$

So we obtain

$$\left\| x_{n_{j+1}} - \text{prox}_{\lambda g} \circ (I - \lambda \nabla f) x_{n_j} \right\|$$

$$\leq \alpha_{n_j} \| h(x_{n_j}) - V x_{n_j} \| + (1 - \alpha_{n_j}) |\lambda_{n_j} - \lambda| \frac{\| x_{n_j} - V_{n_j} x_{n_j} \|}{a}. \tag{3.11}$$

Combining (3.10) and (3.11) we get

$$\left\| x_{n_j} - \left(\text{prox}_{\lambda g} \circ (I - \lambda \nabla f) \right) x_{n_j} \right\|$$

$$\leq \| x_{n_j} - x_{n_{j+1}} \| + \alpha_{n_j} \| h(x_{n_j}) - V x_{n_j} \| + (1 - \alpha_{n_j}) |\lambda_{n_j} - \lambda| \frac{\| x_{n_j} - V_{n_j} x_{n_j} \|}{a}. \tag{3.12}$$

Since $\lambda_{n_j} \to \lambda$, and $\alpha_n \to 0$, by (3.12) we get

$$\left\| x_{n_j} - \left(\text{prox}_{\lambda g} \circ (I - \lambda \nabla f) \right) x_{n_j} \right\| \to 0. \tag{3.13}$$

By the proof of Theorem 3.4 we know that $\text{prox}_{\lambda g} \circ (I - \lambda \nabla f)$ is nonexpansive. It follows from Lemma 2.7 and (3.13) that $z \in S$.

We next show that

$$\limsup_{n \to \infty} \langle h(x^*) - x^*, x_n - x^* \rangle \leq 0, \tag{3.14}$$

where x^* is obtained in Theorem 3.4. Indeed, replacing n with n_j in (3.14), and letting $j \to \infty$, we have

$$\limsup_{n \to \infty} \langle h(x^*) - x^*, x_n - x^* \rangle = \lim_{j \to \infty} \langle h(x^*) - x^*, x_{n_j} - x^* \rangle.$$

Hence, by (3.6), we obtain

$$\limsup_{n \to \infty} \langle h(x^*) - x^*, x_n - x^* \rangle = \langle h(x^*) - x^*, z - x^* \rangle \leq 0.$$

We finally show that $x_n \to x^*$. We have

$$\| x_{n+1} - x^* \|^2$$

$$= \| \alpha_n h(x_n) + (1 - \alpha_n) V_n x_n - x^* \|^2$$

$$= \| \alpha_n (h(x_n) - h(x^*)) + (1 - \alpha_n)(V_n x_n - x^*) + \alpha_n (h(x^*) - x^*) \|^2$$

$$\leq \left\| \alpha_n \big(h(x_n) - h(x^*) \big) + (1 - \alpha_n)\big(V_n x_n - x^* \big) \right\|^2 + 2\big\langle \alpha_n \big(h(x^*) - x^* \big), x_{n+1} - x^* \big\rangle$$

$$\leq \alpha_n \left\| h(x_n) - h(x^*) \right\|^2 + (1 - \alpha_n) \left\| V_n x_n - x^* \right\|^2 + 2\big\langle \alpha_n \big(h(x^*) - x^* \big), x_{n+1} - x^* \big\rangle$$

$$\leq \alpha_n \rho^2 \left\| x_n - x^* \right\|^2 + (1 - \alpha_n) \left\| x_n - x^* \right\|^2 + 2\big\langle \alpha_n \big(h(x^*) - x^* \big), x_{n+1} - x^* \big\rangle$$

$$\leq \big(1 - \alpha_n(1 - \rho^2) \big) \left\| x_n - x^* \right\|^2 + 2\alpha_n \big\langle h(x^*) - x^*, x_{n+1} - x^* \big\rangle.$$

It then follows that

$$\left\| \alpha_n h(x_n) + (1 - \alpha_n) V_n x_n - x^* \right\|^2 \leq \big(1 - \alpha_n(1 - \rho^2) \big) \left\| x_n - x^* \right\|^2 + \alpha_n \delta_n, \tag{3.15}$$

where $\delta_n = 2\langle h(x^*) - x^*, x_{n+1} - x^* \rangle$. Applying Lemma 2.5 to the inequality (3.15), together with (3.14), we get $x_n \to x^*$ as $n \to \infty$. $\square$

4 An application of proximal algorithm to the lasso and Q-lasso

Take $f(x) = \frac{1}{2}\|Ax - b\|_2^2$ and $g(x) = \gamma\|x\|_1$, then lasso (1.2) can be solved by the proximal algorithms (3.8). We have $\nabla f(x) = A^t(Ax - b)$, and we show that ∇f is Lipschitz continuous with constant $L = \|A\|_2^2$ as follows:

$$\left\| A^t(Ax - b) - A^t(Ay - b) \right\|_2 = \left\| A^t A(x - y) \right\|_2 \leq \|A\|_2^2 \|x - y\|_2.$$

Then the proximal algorithm (3.8) is equivalent to

$$x_{n+1} = \alpha_n h(x_n) + (1 - \alpha_n)\Big[\mathrm{prox}_{\lambda_n \gamma \|\cdot\|_1}\big(I - \lambda_n A^t(Ax - b) \big)\Big]x_n. \tag{4.1}$$

Here we have, for $\alpha > 0$ and $x = (x_i)^t \in \mathbb{R}^n$,

$$\mathrm{prox}_{\alpha\|\cdot\|_1}(x) = \big(\mathrm{prox}_{\alpha|\cdot|}(x_1), \ldots, \mathrm{prox}_{\alpha|\cdot|}(x_n) \big)^t,$$

and $\mathrm{prox}_{\alpha|\cdot|}(\beta) = \mathrm{sgn}(\beta)\max\{|\beta| - \alpha, 0\}$ for $\beta \in \mathbb{R}$.

Theorem 4.1 *Assume that*

(C1) $0 < \lambda_n \leq 2/\|A\|_2^2$;

(C2) $\alpha_n \to 0$;

(C3) $\sum_{n=0}^{\infty} \alpha_n = \infty$;

(C4) $\sum_{n=0}^{\infty} |\alpha_{n+1} - \alpha_n| < \infty$;

(C5) $\sum_{n=0}^{\infty} |\lambda_{n+1} - \lambda_n| < \infty$.

Then the sequence $\{x_n\}_{n=0}^{\infty}$ generated by (4.1) converges to a solution x^* of lasso (1.2), which also solves the variational inequality (3.6).

For Q-lasso (1.6), we take $f(x) = (1/2)\|(I - P_Q)Ax\|_2^2$ and $g(x) = \gamma\|x\|_1$. Since P_Q is $\frac{1}{2}$-averaged, $I - P_Q$ is nonexpansive, so we can show that $\nabla f(x) = A^t(I - P_Q)Ax$ is Lipschitz continuous with constant $L = \|A\|_2^2$ as follows:

$$\left\| A^t(I - P_Q)Ax - A^t(I - P_Q)Ay \right\|_2$$

$$\leq \|A\|_2 \left\| (I - P_Q)Ax - (I - P_Q)Ay \right\|_2$$

$$\leq \|A\|_2 \|Ax - Ay\|_2$$

$$\leq \|A\|_2^2 \|x - y\|_2.$$

Then the proximal algorithm (3.8) is reduced to the following algorithm for Q-lasso (1.6):

$$x_{n+1} = \alpha_n f(x_n) + (1 - \alpha_n)\big[\mathrm{prox}_{\lambda_n \gamma \|\cdot\|_1}\big(I - \lambda_n A^t(I - P_Q)A\big)\big]x_n. \tag{4.2}$$

The convergence of Theorem 3.5 reads as follows for Q-lasso (1.6).

Theorem 4.2 *Assume that*

(C1) $0 < \lambda_n \leq 2/\|A\|_2^2$;

(C2) $\alpha_n \to 0$;

(C3) $\sum_{n=0}^{\infty} \alpha_n = \infty$;

(C4) $\sum_{n=0}^{\infty} |\alpha_{n+1} - \alpha_n| < \infty$;

(C5) $\sum_{n=0}^{\infty} |\lambda_{n+1} - \lambda_n| < \infty$.

Then the sequence $\{x_n\}_{n=0}^{\infty}$ generated by (4.2) converges to a solution x^* of Q-lasso (1.6), which is also a solution of the variational inequality (3.6).

5 Conclusion

1. We modify the proximal-gradient algorithm based on the viscosity proximation method; thus, we obtain strong convergence of results in [11]. Then we apply our results to the lasso and the Q-lasso.

2. Theorem 3.4 proves the continuous version of Theorem 3.5, which is not presented in [11].

3. In our main result, we extend the scope of L, that is, the condition in our main results is weaker than the condition in [11] and [12].

References
1. Tibshirani, R: Regression shrinkage and selection via the lasso. J. R. Stat. Soc., Ser. B **58**, 267-288 (1996)
2. Chen, SS, Donoho, DL, Saunders, MA: Atomic decomposition by basis pursuit. SIAM J. Sci. Comput. **20**(1), 33-61 (1998)
3. Donoho, D: Compressed sensing. IEEE Trans. Inf. Theory **52**(4), 1289-1306 (2000)
4. Candes, EJ, Tao, T: Near-optimal signal recovery from random projections: universal encoding strategies? IEEE Trans. Inf. Theory **52**(12), 5406-5425 (2006)
5. Candes, EJ, Romberg, J, Tao, T: Stable signal recovery from incomplete and inaccurate measurement. Commun. Pure Appl. Math. **59**(2), 1207-1223 (2006)
6. Candes, EJ, Romberg, J, Tao, T: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theory **52**(2), 489-509 (2006)
7. Cipra, BA: l_1-magic. SIAM News. 39, 9 (2006)
8. Censor, Y, Elfving, T: A multiprojection algorithm using Bregman projections in a product space. Numer. Algorithms **8**(2-4), 221-239 (1994)
9. Censor, Y, Elfving, T, Kopt, N, Bortfeld, T: The multiple-sets split-feasibility problem and its applications for inverse problem. Inverse Probl. **21**, 2071-2084 (2005)

10. Censor, Y, Elfving, T, Martin, B, Trfimov, A: A unified approach for inversion problems in intensity-modulated radiation therapy. Phys. Med. Biol. **51**, 2353-2365 (2006)
11. Xu, HK: Properties and iterative methods for the lasso and its variants. Chin. Ann. Math., Ser. B **35**(3), 1-18 (2014)
12. Alghamdi, MA, Alghamdi, MA, Shahzad, N, Xu, HK: Properties and iterative methods for the Q-lasso. Abstr. Appl. Anal. **2013**, Article ID 250943 (2013)
13. Moudafi, A: Viscosity approximation methods for fixed-points problems. J. Math. Anal. Appl. **241**, 46-55 (2000)
14. Xu, HK: Viscosity approximation methods for nonexpansive mapping. J. Math. Anal. Appl. **298**, 279-291 (2004)
15. Combettes, PL, Wajs, R: Signal recovery by proximal forward-backward splitting. Multiscale Model. Simul. **4**(4), 1168-1200 (2005)
16. Xu, HK: Averaged mappings and the gradient-projection algorithm. J. Optim. Theory Appl. **150**, 360-378 (2011)
17. Geobel, K, Kirk, WA: Topics in Metric Fixed Point Theory. Cambridge Studies in Advanced Mathematics, vol. 28. Cambridge University Press, Cambridge (1990)
18. Baillon, JB, Haddad, G: Quelques propriétés des opérateurs angle-bornés et n-cycliquement monotones. Isr. J. Math. **26**, 137-150 (1977)
19. Marino, G, Muglia, L: On the auxiliary mappings generated by a family of mappings and solutions of variational inequalities problems. Optim. Lett. **9**, 263-282 (2015)