Gradient projection method for convex functions and strongly convex sets.

M. O. Golubev
maksimkane@mail.ru

Department of Higher Mathematics
Moscow Institute of Physics and Technology (SU)

SOPHIA ANTIPOLIS

June 29, 2015

Definitions 1

$\mathbb{H}$ - the Hilbert space over $\mathbb{R}$. $\langle p, x\rangle$ is scalar product for vectors $p, x \in \mathbb{H}$. Let $B_R(x) = \{y \in \mathbb{H} \mid \|y - x\| \leq R\}$.

Definition

The metric projection of the point $x \in \mathbb{H}$ on the set $A \subset \mathbb{H}$ is defined as follows $P_A(x) = \left\{ a \in A \mid \|x - a\| = \inf\limits_{y \in A} \|x - y\| \right\}$.

The set $P_A(x)$ is a singleton for any closed convex subset $A \subset \mathbb{H}$ and for any point $x \in \mathbb{H}$, i.e. $P_A(x) = \{a(x)\}$. Moreover, for any pair of points $x_0, x_1 \in \mathbb{H}$, $\{a_i\} = P_A(x_i)$, $i \in \{0, 1\}$ we have

$$\|a_0 - a_1\| \leq 1 \cdot \|x_0 - x_1\|.$$

For a subset $A \subset \mathbb{H}$ and a number $\varrho > 0$ we define the open ϱ-neighbourhood of the set A

$$U_A(\varrho) = \{x \in \mathbb{H} \mid \varrho_A(x) < \varrho\}.$$

Definitions 1

$\mathbb{H}$ - the Hilbert space over $\mathbb{R}$. $\langle p, x\rangle$ is scalar product for vectors $p, x \in \mathbb{H}$. Let $B_R(x) = \{y \in \mathbb{H} \mid \|y - x\| \leq R\}$.

Definition

The metric projection of the point $x \in \mathbb{H}$ on the set $A \subset \mathbb{H}$ is defined as follows $P_A(x) = \left\{ a \in A \mid \|x - a\| = \inf_{y \in A} \|x - y\| \right\}$.

The set $P_A(x)$ is a singleton for any closed convex subset $A \subset \mathbb{H}$ and for any point $x \in \mathbb{H}$, i.e. $P_A(x) = \{a(x)\}$. Moreover, for any pair of points $x_0, x_1 \in \mathbb{H}$, $\{a_i\} = P_A(x_i)$, $i \in \{0, 1\}$ we have

$$\|a_0 - a_1\| \leq 1 \cdot \|x_0 - x_1\|.$$

For a subset $A \subset \mathbb{H}$ and a number $\varrho > 0$ we define the open ϱ-neighbourhood of the set A

$$U_A(\varrho) = \{x \in \mathbb{H} \mid \varrho_A(x) < \varrho\}.$$

Definitions 1

$\mathbb{H}$ - the Hilbert space over $\mathbb{R}$. $\langle p, x\rangle$ is scalar product for vectors $p, x \in \mathbb{H}$. Let $B_R(x) = \{y \in \mathbb{H} \mid \|y - x\| \leq R\}$.

Definition

The metric projection of the point $x \in \mathbb{H}$ on the set $A \subset \mathbb{H}$ is defined as follows $P_A(x) = \left\{ a \in A \mid \|x - a\| = \inf\limits_{y \in A} \|x - y\| \right\}$.

The set $P_A(x)$ is a singleton for any closed convex subset $A \subset \mathbb{H}$ and for any point $x \in \mathbb{H}$, i.e. $P_A(x) = \{a(x)\}$. Moreover, for any pair of points $x_0, x_1 \in \mathbb{H}$, $\{a_i\} = P_A(x_i)$, $i \in \{0, 1\}$ we have

$$\|a_0 - a_1\| \leq 1 \cdot \|x_0 - x_1\|.$$

For a subset $A \subset \mathbb{H}$ and a number $\varrho > 0$ we define the open ϱ-neighbourhood of the set A

$$U_A(\varrho) = \{x \in \mathbb{H} \mid \varrho_A(x) < \varrho\}.$$

Definitions 1

$\mathbb{H}$ - the Hilbert space over $\mathbb{R}$. $\langle p, x\rangle$ is scalar product for vectors $p, x \in \mathbb{H}$. Let $B_R(x) = \{y \in \mathbb{H} \mid \|y - x\| \le R\}$.

Definition

The metric projection of the point $x \in \mathbb{H}$ on the set $A \subset \mathbb{H}$ is defined as follows $P_A(x) = \left\{ a \in A \mid \|x - a\| = \inf_{y \in A} \|x - y\| \right\}$.

The set $P_A(x)$ is a singleton for any closed convex subset $A \subset \mathbb{H}$ and for any point $x \in \mathbb{H}$, i.e. $P_A(x) = \{a(x)\}$. Moreover, for any pair of points $x_0, x_1 \in \mathbb{H}$, $\{a_i\} = P_A(x_i)$, $i \in \{0, 1\}$ we have

$$\|a_0 - a_1\| \le 1 \cdot \|x_0 - x_1\|.$$

For a subset $A \subset \mathbb{H}$ and a number $\varrho > 0$ we define the open ϱ-neighbourhood of the set A

$$U_A(\varrho) = \{x \in \mathbb{H} \mid \varrho_A(x) < \varrho\}.$$

$\mathbb{H}$ - the Hilbert space over $\mathbb{R}$. $\langle p, x \rangle$ is scalar product for vectors $p, x \in \mathbb{H}$. Let $B_R(x) = \{y \in \mathbb{H} \mid \|y - x\| \leq R\}$.

Definition

The metric projection of the point $x \in \mathbb{H}$ on the set $A \subset \mathbb{H}$ is defined as follows $P_A(x) = \left\{ a \in A \mid \|x - a\| = \inf_{y \in A} \|x - y\| \right\}$.

The set $P_A(x)$ is a singleton for any closed convex subset $A \subset \mathbb{H}$ and for any point $x \in \mathbb{H}$, i.e. $P_A(x) = \{a(x)\}$. Moreover, for any pair of points $x_0, x_1 \in \mathbb{H}$, $\{a_i\} = P_A(x_i)$, $i \in \{0, 1\}$ we have

$$\|a_0 - a_1\| \leq 1 \cdot \|x_0 - x_1\|.$$

For a subset $A \subset \mathbb{H}$ and a number $\varrho > 0$ we define the open ϱ-neighbourhood of the set A

$$U_A(\varrho) = \{x \in \mathbb{H} \mid \varrho_A(x) < \varrho\}.$$

$\mathbb{H}$ - the Hilbert space over $\mathbb{R}$. $\langle p, x \rangle$ is scalar product for vectors $p, x \in \mathbb{H}$. Let $B_R(x) = \{y \in \mathbb{H} \mid \|y - x\| \leq R\}$.

Definition

The metric projection of the point $x \in \mathbb{H}$ on the set $A \subset \mathbb{H}$ is defined as follows $P_A(x) = \left\{ a \in A \mid \|x - a\| = \inf_{y \in A} \|x - y\| \right\}$.

The set $P_A(x)$ is a singleton for any closed convex subset $A \subset \mathbb{H}$ and for any point $x \in \mathbb{H}$, i.e. $P_A(x) = \{a(x)\}$. Moreover, for any pair of points $x_0, x_1 \in \mathbb{H}$, $\{a_i\} = P_A(x_i)$, $i \in \{0, 1\}$ we have

$$\|a_0 - a_1\| \leq 1 \cdot \|x_0 - x_1\|.$$

For a subset $A \subset \mathbb{H}$ and a number $\varrho > 0$ we define the open ϱ-neighbourhood of the set A

$$U_A(\varrho) = \{x \in \mathbb{H} \mid \varrho_A(x) < \varrho\}.$$

Definitions 2

Definition

A nonempty subset $A \subset \mathbb{H}$ is called a strongly convex set of radius $R > 0$ if it can be represented as the intersection of closed balls of radius $R > 0$, that is there exists a subset $X \subset \mathbb{H}$ such that $A = \bigcap\limits_{x \in X} B_R(x)$.

Definition

Normal cone to the set $A \subset \mathbb{H}$ at the point $a \in \overline{A}$ is the following set

$$N(A; a) = \left\{ p \in \mathbb{H} \,\middle|\, \sup_{x \in A} \langle p, x \rangle \leq \langle p, a \rangle \right\}.$$

Definitions 2

Definition

A nonempty subset $A \subset \mathbb{H}$ is called a strongly convex set of radius $R > 0$ if it can be represented as the intersection of closed balls of radius $R > 0$, that is there exists a subset $X \subset \mathbb{H}$ such that $A = \bigcap_{x \in X} B_R(x)$.

Definition

Normal cone to the set $A \subset \mathbb{H}$ at the point $a \in \overline{A}$ is the following set

$$N(A; a) = \left\{ p \in \mathbb{H} \mid \sup_{x \in A} \langle p, x \rangle \leq \langle p, a \rangle \right\}.$$

Theorem A

Theorem (M. V. Balashov, M. O. Golubev)

Let $A \subset \mathbb{H}$ be a closed convex subset. Then the following properties are equivalent

1) A is a strongly convex set of radius $R > 0$,

2) $\forall \varrho > 0, \forall x_0, x_1 \in \mathbb{H} \setminus U_A(\varrho), \{a_i\} = P_A(x_i), i \in \{0, 1\},$

$$\|a_0 - a_1\| \leq \frac{R}{(R + \varrho)} \cdot \|x_0 - x_1\|.$$

Theorem A

Theorem (M. V. Balashov, M. O. Golubev)

Let $A \subset \mathbb{H}$ be a closed convex subset. Then the following properties are equivalent
1) A is a strongly convex set of radius $R > 0$,
2) $\forall \varrho > 0, \forall x_0, x_1 \in \mathbb{H} \setminus U_A(\varrho), \{a_i\} = P_A(x_i), i \in \{0, 1\}$,

$$\|a_0 - a_1\| \leq \frac{R}{(R + \varrho)} \cdot \|x_0 - x_1\|.$$

Theorem A

Theorem (M. V. Balashov, M. O. Golubev)

Let $A \subset \mathbb{H}$ be a closed convex subset. Then the following properties are equivalent
1) A is a strongly convex set of radius $R > 0$,
2) $\forall \varrho > 0, \forall x_0, x_1 \in \mathbb{H} \setminus U_A(\varrho), \{a_i\} = P_A(x_i), i \in \{0, 1\}$,

$$\|a_0 - a_1\| \le \frac{R}{(R + \varrho)} \cdot \|x_0 - x_1\|.$$

Consider the minimization problem

$$f(x) \to \min, \quad x \in A \subset \mathbb{H}. \tag{1}$$

Consider the standard gradient projection algorithm:

$$x_{k+1} = P_A(x_k - \alpha_k f'(x_k)), \quad x_1 \in \partial A, \quad \alpha_k > 0. \tag{2}$$

Consider the minimization problem

$$f(x) \to \min, \quad x \in A \subset \mathbb{H}. \tag{1}$$

Consider the standard gradient projection algorithm:

$$x_{k+1} = P_A(x_k - \alpha_k f'(x_k)), \quad x_1 \in \partial A, \quad \alpha_k > 0. \tag{2}$$

Conditions

Suppose that:

(i) $f : \mathbb{H} \to \mathbb{R}$ is convex, differentiable and the gradient $f'(x)$ satisfies the Lipschitz condition with constant $M > 0$, i.e. for all $x_1, x_2 \in \mathbb{H}$

$$\|f'(x_1) - f'(x_2)\| \leq M\|x_1 - x_2\|,$$

(ii) $A \subset \mathbb{H}$ is strongly convex with radius R,

(iii) for any $k \in \mathbb{N}$ there exists a unit vector $n(x_k) \in N(A; x_k)$ such that $\langle n(x_k), f'(x_k)\rangle \leq 0$, (i.e. $x_k - \alpha_k f'(x_k) \notin A$ for any $\alpha_k > 0$),

(iv) the problem (1) has a unique solution $x_* \in \partial A$.

Conditions

Suppose that:

(i) $f : \mathbb{H} \to \mathbb{R}$ is convex, differentiable and the gradient $f'(x)$ satisfies the Lipschitz condition with constant $M > 0$, i.e. for all $x_1, x_2 \in \mathbb{H}$

$$\|f'(x_1) - f'(x_2)\| \leq M\|x_1 - x_2\|,$$

(ii) $A \subset \mathbb{H}$ is strongly convex with radius R,

(iii) for any $k \in \mathbb{N}$ there exists a unit vector $n(x_k) \in N(A; x_k)$ such that $\langle n(x_k), f'(x_k)\rangle \leq 0$, (i.e. $x_k - \alpha_k f'(x_k) \notin A$ for any $\alpha_k > 0$),

(iv) the problem (1) has a unique solution $x_* \in \partial A$.

Conditions

Suppose that:

(i) $f : \mathbb{H} \to \mathbb{R}$ is convex, differentiable and the gradient $f'(x)$ satisfies the Lipschitz condition with constant $M > 0$, i.e. for all $x_1, x_2 \in \mathbb{H}$

$$\|f'(x_1) - f'(x_2)\| \leq M\|x_1 - x_2\|,$$

(ii) $A \subset \mathbb{H}$ is strongly convex with radius R,

(iii) for any $k \in \mathbb{N}$ there exists a unit vector $n(x_k) \in N(A; x_k)$ such that $\langle n(x_k), f'(x_k)\rangle \leq 0$, (i.e. $x_k - \alpha_k f'(x_k) \notin A$ for any $\alpha_k > 0$),

(iv) the problem (1) has a unique solution $x_* \in \partial A$.

Conditions

Suppose that:

(i) $f : \mathbb{H} \to \mathbb{R}$ is convex, differentiable and the gradient $f'(x)$ satisfies the Lipschitz condition with constant $M > 0$, i.e. for all $x_1, x_2 \in \mathbb{H}$

$$\|f'(x_1) - f'(x_2)\| \leq M\|x_1 - x_2\|,$$

(ii) $A \subset \mathbb{H}$ is strongly convex with radius R,

(iii) for any $k \in \mathbb{N}$ there exists a unit vector $n(x_k) \in N(A; x_k)$ such that $\langle n(x_k), f'(x_k)\rangle \leq 0$, (i.e. $x_k - \alpha_k f'(x_k) \notin A$ for any $\alpha_k > 0$),

(iv) the problem (1) has a unique solution $x_* \in \partial A$.

Theorem 1

Theorem

a. *Suppose that conditions (i)-(iv) hold. Let* $\alpha_k = \alpha \in \left(0, \frac{2}{M}\right]$. *Let* $t = \min\limits_{x \in \partial A} \|f'(x)\| > 0$.

Then the sequence x_k, *generated by the rule (2), converges to the solution of (1) at a rate of geometric progression:*
$\|x_{k+1} - x_*\| \leq q\|x_k - x_*\|$, *where* $q = \frac{R}{\sqrt[4]{(R^2+\alpha^2 t^2)(R+\alpha t)^2}}$;

b. *Suppose that conditions (i)-(iv) hold. Let* $\alpha_k = \alpha \in \left(0, \frac{2}{M}\right]$.
Then the sequence x_k, *generated by the rule (2), converges to the solution of (1) and the estimate holds:* $\|x_{k+1} - x_*\| \leq q_k\|x_k - x_*\|$,
where $q_k = \sqrt[4]{\frac{R^2}{R^2+\alpha^2\|f'(x_k)\|^2}}$.

Theorem

a. *Suppose that conditions (i)-(iv) hold. Let* $\alpha_k = \alpha \in \left(0, \frac{2}{M}\right]$. *Let* $t = \min\limits_{x \in \partial A} \|f'(x)\| > 0$.
Then the sequence x_k, *generated by the rule (2), converges to the solution of (1) at a rate of geometric progression:*
$\|x_{k+1} - x_*\| \le q\|x_k - x_*\|$, *where* $q = \frac{R}{\sqrt[4]{(R^2+\alpha^2 t^2)(R+\alpha t)^2}}$;
b. *Suppose that conditions (i)-(iv) hold. Let* $\alpha_k = \alpha \in \left(0, \frac{2}{M}\right]$.
Then the sequence x_k, *generated by the rule (2), converges to the solution of (1) and the estimate holds:* $\|x_{k+1} - x_*\| \le q_k\|x_k - x_*\|$, *where* $q_k = \sqrt[4]{\frac{R^2}{R^2+\alpha^2\|f'(x_k)\|^2}}$.

Theorem 2

Theorem

Suppose that conditions (i)-(ii) hold. Let $\frac{RM}{t} < 1$, *where* $t = \min_{x \in \partial A} \|f'(x)\| > 0$.

The sequence x_k *is generated by the rule (2) with* $\alpha_k = \alpha > 0$ *for any* k.

a. *if* $\alpha \in \left(\frac{2R}{t}, \frac{2}{M}\right]$ *then the sequence* x_k *converges to the solution of (1) at a rate of geometric progression:* $\|x_{k+1} - x_*\| \leq q\|x_k - x_*\|$, *where* $q = \frac{R}{\alpha t - R}$;

b. *if* $\alpha > \frac{2}{M}$ *then the sequence* x_k *converges to the solution of (1) at a rate of geometric progression:* $\|x_{k+1} - x_*\| \leq q\|x_k - x_*\|$, *where* $q = \frac{R(\alpha M - 1)}{\alpha t - R}$.

Theorem 2

Theorem

Suppose that conditions (i)-(ii) hold. Let $\frac{RM}{t} < 1$, *where* $t = \min_{x \in \partial A} \|f'(x)\| > 0$.
The sequence x_k *is generated by the rule (2) with* $\alpha_k = \alpha > 0$ *for any* k.
a. *if* $\alpha \in \left(\frac{2R}{t}, \frac{2}{M}\right]$ *then the sequence* x_k *converges to the solution of (1) at a rate of geometric progression:*
$\|x_{k+1} - x_*\| \leq q\|x_k - x_*\|$, *where* $q = \frac{R}{\alpha t - R}$;
b. *if* $\alpha > \frac{2}{M}$ *then the sequence* x_k *converges to the solution of (1) at a rate of geometric progression:* $\|x_{k+1} - x_*\| \leq q\|x_k - x_*\|$, *where* $q = \frac{R(\alpha M - 1)}{\alpha t - R}$.

Theorem 2

Theorem

Suppose that conditions (i)-(ii) hold. Let $\frac{RM}{t} < 1$, *where*
$t = \min_{x \in \partial A} \|f'(x)\| > 0$.
The sequence x_k *is generated by the rule (2) with* $\alpha_k = \alpha > 0$ *for any* k.
a. *if* $\alpha \in \left(\frac{2R}{t}, \frac{2}{M}\right]$ *then the sequence* x_k *converges to the solution of (1) at a rate of geometric progression:*
$\|x_{k+1} - x_*\| \leq q\|x_k - x_*\|$, *where* $q = \frac{R}{\alpha t - R}$;
b. *if* $\alpha > \frac{2}{M}$ *then the sequence* x_k *converges to the solution of (1) at a rate of geometric progression:* $\|x_{k+1} - x_*\| \leq q\|x_k - x_*\|$, *where* $q = \frac{R(\alpha M - 1)}{\alpha t - R}$.

Conditions

Definition

Suppose that subset $U \subset \mathbb{H}$ is convex. Function $f : U \to \mathbb{R}$ is called strongly convex with constant $\gamma > 0$ on the subset U if function $f(x) - \frac{\gamma}{2}\|x\|^2$ is convex on the subset U.

Suppose that:

(v) $f : \mathbb{H} \to \mathbb{R}$ is strongly convex with $\gamma > 0$, differentiable and the gradient $f'(x)$ satisfies the Lipschitz condition with constant $M > 0$, i.e. for all $x_1, x_2 \in \mathbb{H}$

$$\|f'(x_1) - f'(x_2)\| \leq M\|x_1 - x_2\|.$$

Conditions

Definition

Suppose that subset $U \subset \mathbb{H}$ is convex. Function $f : U \to \mathbb{R}$ is called strongly convex with constant $\gamma > 0$ on the subset U if function $f(x) - \frac{\gamma}{2}\|x\|^2$ is convex on the subset U.

Suppose that:

(v) $f : \mathbb{H} \to \mathbb{R}$ is strongly convex with $\gamma > 0$, differentiable and the gradient $f'(x)$ satisfies the Lipschitz condition with constant $M > 0$, i.e. for all $x_1, x_2 \in \mathbb{H}$

$$\|f'(x_1) - f'(x_2)\| \leq M\|x_1 - x_2\|.$$

Theorem

a. *Suppose that conditions (ii)-(v) hold. Let* $\alpha_k = \alpha \in \left(0, \min\left\{\frac{2\gamma}{M^2}, \frac{2}{\gamma+M}\right\}\right)$, $M > \gamma$. *Let* $t = \min_{x\in\partial A} \|f'(x)\| > 0$. *Define the number*

$$L = L(\alpha, \gamma, M) = \\ = \min\left\{\sqrt{1 - 2\alpha\gamma + \alpha^2 M^2}, \sqrt{1 - \frac{2\alpha\gamma M}{\gamma + M}}\right\}$$

Then the sequence x_k, *generated by the rule (2), converges to the solution of (1) at a rate of geometric progression:* $\|x_{k+1} - x_*\| \le q\|x_k - x_*\|$, *where* $q = \frac{R}{\sqrt[4]{(R^2+\alpha^2 t^2)(R+\alpha t)^2}} L$;

Theorem 3.a

Theorem

a. *Suppose that conditions (ii)-(v) hold. Let* $\alpha_k = \alpha \in \left(0, \min\left\{\frac{2\gamma}{M^2}, \frac{2}{\gamma+M}\right\}\right)$, $M > \gamma$. *Let* $t = \min_{x\in\partial A} \|f'(x)\| > 0$. *Define the number*

$$L = L(\alpha, \gamma, M) = \\ = \min\left\{\sqrt{1 - 2\alpha\gamma + \alpha^2 M^2}, \sqrt{1 - \frac{2\alpha\gamma M}{\gamma + M}}\right\}$$

Then the sequence x_k, *generated by the rule (2), converges to the solution of (1) at a rate of geometric progression:* $\|x_{k+1} - x_*\| \leq q\|x_k - x_*\|$, *where* $q = \frac{R}{\sqrt[4]{(R^2+\alpha^2 t^2)(R+\alpha t)^2}} L$;

Theorem 3.b

Theorem

b. *Suppose that conditions (ii)-(v) hold. Let* $\alpha_k = \alpha \in \left(0, \min\left\{\frac{2\gamma}{M^2}, \frac{2}{\gamma+M}\right\}\right)$. *Define the number*

$$L = L(\alpha, \gamma, M) = \\ = \min\left\{\sqrt{1 - 2\alpha\gamma + \alpha^2 M^2}, \sqrt{1 - \frac{2\alpha\gamma M}{\gamma + M}}\right\}$$

Then the sequence x_k, *generated by the rule (2), converges to the solution of (1) and the estimate holds:* $\|x_{k+1} - x_*\| \leq q_k \|x_k - x_*\|$, *where* $q_k = \sqrt[4]{\frac{R^2}{R^2 + \alpha^2 \|f'(x_k)\|^2}} L$.

Theorem 3.b

Theorem

b. *Suppose that conditions (ii)-(v) hold. Let* $\alpha_k = \alpha \in \left(0, \min\left\{\frac{2\gamma}{M^2}, \frac{2}{\gamma+M}\right\}\right)$. *Define the number*

$$L = L(\alpha, \gamma, M) = \\ = \min\left\{\sqrt{1 - 2\alpha\gamma + \alpha^2 M^2}, \sqrt{1 - \frac{2\alpha\gamma M}{\gamma + M}}\right\}$$

Then the sequence x_k*, generated by the rule (2), converges to the solution of (1) and the estimate holds:* $\|x_{k+1} - x_*\| \leq q_k \|x_k - x_*\|$, *where* $q_k = \sqrt[4]{\frac{R^2}{R^2 + \alpha^2 \|f'(x_k)\|^2}} L$.

Thank you for your attention!