Subgradient Projection Algorithm with Computational Errors

Alexander J. Zaslavski

February 17, 2021

We study the subgradient projection algorithm for minimization of convex and nonsmooth functions, under the presence of computational errors. We show that our algorithms generate a good approximate solution, if computational errors are bounded from above by a small positive constant. Moreover, for a known computational error, we find out what an approximate solution can be obtained and how many iterates one needs for this.

We study the subgradient projection algorithm for minimization of convex and nonsmooth functions and for computing the saddle points of convex-concave functions, under the presence of computational errors. The problem is described by an objective function and a set of feasible points. For this algorithm each iteration consists of two steps. The first step is a calculation of a subgradient of the objective function while in the second one we calculate a projection on the feasible set. In each of these two steps there is a computational error.

In general, these two computational errors are different. In our recent research (see . J. Zaslavski, Numerical optimization with computational errors, Springer, 2016 and A. J. Zaslavski, Convex optimization with computational errors, Springer, 2020) we showed that our algorithm generates a good approximate solution, if all the computational errors are bounded from above by a small positive constant. Moreover, if we know computational errors for the two steps of our algorithm, we find out what an approximate solution can be obtained and how many iterates one needs for this. In this talk we discuss a generalization of these results for an extension of the projected subgradient method, when instead of the projection on the feasible set it is used a quasi-nonexpansive retraction on this set. This generalization is studied in A. J. Zaslavski, The projected subgradient algorithm in convex optimization, SpringerBriefs in Optimization, 2020.

We study the subgradient algorithm for constrained minimization problems in Hilbert spaces equipped with an inner product denoted by $\langle \cdot, \cdot \rangle$ which induces a complete norm $\| \cdot \|$ and use the following notation.

For every $z \in R^1$ denote by $\lfloor z \rfloor$ the largest integer which does not exceed z:

$$\lfloor z \rfloor = \max\{i \in R^1 : i \text{ is an integer and } i \leq z\}.$$

For every nonempty set D, every function $f : D \to R^1$ and every nonempty set $C \subset D$ we set

$$\inf(f, C) = \inf\{f(x) : \ x \in C\}.$$

Let X be a Hilbert space equipped with an inner product denoted by $\langle \cdot, \cdot \rangle$ which induces a complete norm $\| \cdot \|$.

For each $x \in X$ and each $r > 0$ set

$$B_X(x, r) = \{y \in X : \|x - y\| \leq r\}$$

and for each $x \in X$ and each nonempty set $E \subset X$ set

$$d(x, E) = \inf\{\|x - y\| : y \in E\}.$$

For each nonempty open convex set $U \subset X$ and each convex function $f : U \to R^1$, for every $x \in U$ set

$$\partial f(x) = \{l \in X : f(y) - f(x)$$

$$\geq \langle l, y - x \rangle \text{ for all } y \in U\}$$

which is called the subdifferential of the function f at the point x.

Let C be a nonempty closed convex subset of X and let $f : X \to R^1$ be a convex function.

Suppose that there exist $L > 0$, $M_0 > 0$ such that

$$C \subset B_X(0, M_0),$$

$|f(x)-f(y)| \le L\|x-y\|$ for all $x, y \in B_X(0, M_0+2)$.

It is not difficult to see that for each $x \in B_X(0, M_0 + 1)$,

$$\emptyset \neq \partial f(x) \subset B_X(0, L).$$

It is well-know that for every nonempty closed convex set $D \subset X$ and every $x \in X$ there is a unique point $P_D(x) \in D$ satisfying

$$\|x - P_D(x)\| = \inf\{\|x - y\| : \ y \in D\}.$$

We consider the minimization problem

$$f(z) \to \min, \ z \in C.$$

Suppose that $\{\alpha_k\}_{k=0}^{\infty} \subset (0, \infty)$. Let us describe our algorithm.

Subgradient projection algorithm

Initialization: select an arbitrary

$$x_0 \in B_X(0, M_0 + 1).$$

Iterative step: given a current iteration vector $x_t \in U$ calculate $\xi_t \in \partial f(x_t)$ and the next iteration vector $x_{t+1} = P_C(x_t - \alpha_t \xi_t)$.

In [1] we study this algorithm under the presence of computational errors. Namely, in [1] we suppose that $\delta \in (0, 1]$ is a computational error produced by our computer system, and study the following algorithm.

Subgradient projection algorithm with computational errors

Initialization: select an arbitrary

$$x_0 \in B_X(0, M_0 + 1).$$

Iterative step: given a current iteration vector $x_t \in B_X(0, M_0 + 1)$ calculate $\xi_t \in \partial f(x_t) + B_X(0, \delta)$ and the next iteration vector $x_{t+1} \in U$ such that $\|x_{t+1} - P_C(x_t - a_t\xi_t)\| \leq \delta$.

In [2] we consider more complicated, but more realistic, version of this algorithm. Clearly, for the algorithm each iteration consists of two steps. The first step is a calculation of a subgradient of the objective function f while in the second one we calculate a projection on the set C. In each of these two steps there is a computational error produced by our computer system. In general, these two computational errors are different.

This fact is taken into account in the following projection algorithm studied in Chapter 2 of [2].

Suppose that $\{\alpha_k\}_{k=0}^{\infty} \subset (0, \infty)$ and $\delta_f, \delta_C \in (0, 1]$.

Initialization: select an arbitrary

$$x_0 \in B_X(0, M_0 + 1).$$

Iterative step: given a current iteration vector $x_t \in B_X(0, M_0 + 1)$ calculate $\xi_t \in \partial f(x_t) + B_X(0, \delta_f)$ and the next iteration vector $x_{t+1} \in U$ such that $\|x_{t+1} - P_C(x_t - \alpha_t \xi_t)\| \leq \delta_C$.

Note that in practice for some problems the set C is simple but the function f is complicated. In this case δ_C is essentially smaller than δ_f. On the other hand, there are cases when f is simple but the set C is complicated and therefore δ_f is much smaller than δ_C.

In our analysis of the behaviour of the algorithm in [1,2] properties of the projection operator P_C play an important role. In [3] we obtain generalizations of the results obtained in [1,2] for the subgradient methods in the case when the set C is not necessarily convex and the projection operator P_C is replaced by a mapping $P : X \to C$ which satisfies

$$Px = x \text{ for all } x \in C, \tag{1.1}$$

$$\|Px - z\| \leq \|x - z\| \text{ for all } z \in C \text{ and all } x \in X. \tag{1.2}$$

In other words, P is a quasi-nonexpansive retraction on C. Note that there are many mappings $P : X \to C$ satisfying (1.1) and (1.2).

Indeed, in S. Reich and A. J. Zaslavski, Genericity in nonlinear analysis, Developments in Mathematics, Springer, 2014 we consider a space of mappings $P : X \to X$ satisfying (1.1) and (1.2), which is equipped with a natural complete metric, and show that for a generic (typical) mapping from the space its powers converge to a mapping which also satisfies (1.1) and (1.2) and such that its image is C.

Note that the generalizations considered in this book have, besides their obvious mathematical interest, also a significant practical meaning. Usually, the projection operator $P_C : X \to C$ can be calculated when C is a simple set like a linear subspace, a half-space or a simplex. In practice, C is an intersection of simple sets C_i, $i = 1, \dots, q$, where q is a large natural number. The calculation of P_C is not possible in principle. Instead of it it is possible to calculate the product $P_{C_q} \cdots P_{C_1}$ and its powers $(P_{C_q} \cdots P_{C_1})^m$, $m = 1, 2, \dots$.

It is well-known (see, for example, A. J. Zaslavski, Approximate solutions of common fixed point problems, Springer Optimization and Its Applications, Springer, Cham, 2016) that under certain regularity conditions on C_i, $i = 1, \dots, q$ the powers $(P_{C_q} \cdots P_{C_1})^m$ converges as $m \to \infty$ to a mapping $P : X \to C$ which satisfies (1.1) and (1.2). Thus in practice we cannot calculate the projection operator P_C but only a mapping $P : X \to C$ satisfying (1.1) and (1.2) or, more exactly, its approximations. This shows that the results of [3] are indeed important from the point of view of practice.

Optimization problems on bounded sets

Let $(X, \langle \cdot, \cdot \rangle)$ be a Hilbert space with a inner product $\langle \cdot, \cdot \rangle$ which induces a complete norm $\| \cdot \|$.

Let C be a closed nonempty subset of the space X and U be an open convex subset of X such that $C \subset U$. Suppose that $L, M > 0$, $C \subset B_X(0, M)$ and that a convex function $f : U \to R^1$ satisfies

$$|f(u_1) - f(u_2)| \le L\|u_1 - u_2\| \text{ for all } u_1, u_2 \in U.$$

For each point $x \in U$ and each positive number ϵ let

$$\partial f(x) = \{l \in X : \ f(y) - f(x)$$

$$\ge \langle l, y - x\rangle \text{ for all } y \in U\}$$

and let

$$\partial_\epsilon f(x) = \{l \in X : \ f(y) - f(x)$$

$$\ge \langle l, y - x\rangle - \epsilon \text{ for all } y \in U\}.$$

Denote by $\mathcal{M}$ the set of all mappings $P : X \to C$ such that

$$Pz = z \text{ for all } z \in C,$$

$$\|Px - z\| \leq \|x - z\| \text{ for all } x \in X \text{ and all } z \in C.$$

Define

$$\inf(f, C) = \inf\{f(z) : \; z \in C\}.$$

It is clear that $\inf(f, C)$ is finite.

Set

$$C_{min} = \{x \in C : \; f(x) = \inf(f, C)\}.$$

For all $P \in \mathcal{M}$ set $P^0 x = x$, $x \in X$. We assume that

$$C_{min} \neq \emptyset.$$

Clearly, for each $x \in U$,

$$\partial f(x) \subset B_X(0, L).$$

Theorem 1 *Assume that* $\delta_f, \delta_C \in (0,1]$, $T \geq 1$ *is an integer,* $\{\alpha_t\}_{t=0}^{T-1} \subset (0,1]$,

$$\{P_i\}_{i=0}^{T-1} \subset \mathcal{M},$$

$$\{x_i\}_{i=0}^{T} \subset U,\ \{\xi_i\}_{i=0}^{T-1} \subset X,$$

$$\|x_0\| \leq M+1,$$

and that for $i = 0, \dots, T-1$,

$$B_X(\xi_i, \delta_f) \cap \partial f(x_i) \neq \emptyset,$$

$$\|x_{i+1} - P_i(x_i - \alpha_i \xi_i)\| \leq \delta_C.$$

Then

$$\min\{f(x_t):\ t=0,\dots,T-1\}-\inf(f,C),$$

$$f((\sum_{i=0}^{T-1}\alpha_i)^{-1}\sum_{t=0}^{T-1}\alpha_t x_t)-\inf(f,C)$$

$$\le (\sum_{j=0}^{T-1}\alpha_j)^{-1}\sum_{t=0}^{T-1}\alpha_t(f(x_t)-\inf(f,C))$$

$$\le 2^{-1}(2M+1)^2(\sum_{t=0}^{T-1}\alpha_t)^{-1}$$

$$+2^{-1}L^2(\sum_{t=0}^{T-1}\alpha_t^2)(\sum_{t=0}^{T-1}\alpha_t)^{-1}$$

$$+T\delta_C(\sum_{t=0}^{T-1}\alpha_t)^{-1}(2M+L+3)+\delta_f(2M+L+2). \tag{1}$$

Theorem 2 *Assume that* $r > 0$,

$$B_X(z, 2r) \subset U \text{ for all } z \in C,$$

$\Delta > 0$, $\delta_f, \delta_C \in (0, 1]$, $\delta_C \le r$, $T \ge 1$ *is an integer,* $\{\alpha_t\}_{t=0}^{T-1} \subset (0, 1]$,

$$\{P_i\}_{i=0}^{T-1} \subset \mathcal{M},$$

$$\{x_i\}_{i=0}^{T} \subset U, \ \{\xi_i\}_{i=0}^{T-1} \subset X,$$

$$\|x_0\| \le M + 1,$$

$$B_X(x_0, r) \subset U,$$

and that for $i = 0, \dots, T - 1$,

$$B_X(\xi_i, \delta_f) \cap \partial_\Delta f(x_i) \ne \emptyset,$$

$$\|x_{i+1} - P_i(x_i - \alpha_i \xi_i)\| \le \delta_C.$$

Then

$$\min\{f(x_t):\ t = 0,\dots,T-1\} - \inf(f,C),$$

$$f((\sum_{i=0}^{T-1}\alpha_i)^{-1}\sum_{t=0}^{T-1}\alpha_t x_t) - \inf(f,C)$$

$$\leq (\sum_{j=0}^{T-1}\alpha_j)^{-1}\sum_{t=0}^{T-1}\alpha_t(f(x_t) - \inf(f,C))$$

$$\leq 2^{-1}(2M+1)^2(\sum_{t=0}^{T-1}\alpha_t)^{-1} + \Delta$$

$$+2^{-1}(L+\Delta r^{-1})^2(\sum_{t=0}^{T-1}\alpha_t^2)(\sum_{t=0}^{T-1}\alpha_t)^{-1}$$

$$+T\delta_C(\sum_{t=0}^{T-1}\alpha_t)^{-1}(2M+L+3+\Delta r^{-1})$$

$$+\delta_f(2M+L+2+\Delta r^{-1}). \qquad (2)$$

Note that (1) is a particular case of (2) with $\Delta = 0$.

Let $T \geq 1$ be an integer and $A > 0$ be given. We are interested in an optimal choice of α_t, $t = 0, \ldots, T-1$ satisfying $\sum_{t=0}^{T-1} \alpha_t = A$ which minimizes the right-hand side of (2). It was shown in [1] that $\alpha_t = \alpha = T^{-1}A$, $t = 0, \ldots, T-1$. In this case the right-hand side of (2) is

$$2^{-1}(2M+1)^2 T^{-1}\alpha^{-1} + \Delta$$

$$+2^{-1}(L + \Delta r^{-1})^2 \alpha$$

$$+\delta_C \alpha^{-1}(2M + L + 3 + \Delta r^{-1})$$

$$+\delta_f (2M + L + 2 + \Delta r^{-1}).$$

Now we can make the best choice of the step-size $\alpha > 0$. Since T can be arbitrary large we need to minimize the function

$$\delta_C \alpha^{-1}(2M + L + 3 + \Delta r^{-1})$$

$$+2^{-1}(L + \Delta r^{-1})^2 \alpha, \ \alpha > 0$$

which has a minimizer

$$\alpha = (L + \Delta r^{-1})^{-1}(2\delta_C(2M + L + 3$$

$$+\Delta r^{-1}))^{1/2}.$$

With this choice of α the right-hand side of (2) is

$$2^{-1}(2M+1)^2T^{-1}(L+\Delta r^{-1})(2\delta_C(2M$$

$$+L+3+\Delta r^{-1}))^{1/2}+\Delta$$

$$+(L+\Delta r^{-1})(2^{-1}\delta_C(2M+L+3+\Delta r^{-1}))^{1/2}$$

$$+\delta_f(2M+L+2+\Delta r^{-1})$$

$$+2^{-1}(L+\Delta r^{-1})(2\delta_C(2M$$

$$+L+3+\Delta r^{-1}))^{1/2}.$$

Now we should make the best choice of T. It is clear that T should be at the same order as δ_C^{-1}. In this case the right-hand side of (2) does not exceed $c_1\delta_C^{1/2}+\Delta+\delta_f(2M+L+2+\Delta r^{-1})$, where $c_1>0$ is a constant.

Optimization on unbounded sets

Let $(X, \langle \cdot, \cdot \rangle)$ be a Hilbert space with a inner product $\langle \cdot, \cdot \rangle$ which induces a complete norm $\| \cdot \|$.

Let C be a closed nonempty subset of the space X, U be an open convex subset of X such that

$$C \subset U$$

and $f : U \to R^1$ be a convex function which is Lipschitz on all bounded subsets of U.

For each point $x \in U$ and each positive number ϵ let

$$\partial f(x) = \{l \in X : \; f(y) - f(x)$$

$$\geq \langle l, y - x \rangle \text{ for all } y \in U\}$$

and let

$$\partial_\epsilon f(x) = \{l \in X : \; f(y) - f(x)$$

$$\geq \langle l, y - x \rangle - \epsilon \text{ for all } y \in U\}.$$

Assume that

$$\lim_{x \in U, \|x\| \to \infty} f(x) = \infty.$$

It means that for each $M_0 > 0$ there exists $M_1 > 0$ such that if a point $x \in U$ satisfies the inequality $\|x\| \geq M_1$, then $f(x) > M_0$.

Define

$$\inf(f, C) = \inf\{f(z) : \ z \in C\}.$$

Since the function f is Lipschitz on all bounded subsets of the space X, we have that $\inf(f, C)$ is finite.

Set

$$C_{min} = \{x \in C : \ f(x) = \inf(f, C)\}.$$

It is well-known that if the set C is convex, then the set C_{min} is nonempty. Clearly, the set $C_{min} \neq \emptyset$ if the space X is finite-dimensional.

We assume that

$$C_{min} \neq \emptyset.$$

It is clear that C_{min} is a closed subset of C. Fix

$$\theta_0 \in C.$$

Set

$$U_0 = \{x \in U : \ f(x) \leq f(\theta_0) + 4\}.$$

Clearly, there exists a number $\bar{K} > 1$ such that

$$U_0 \subset B_X(0, \bar{K}).$$

Since the function f is Lipschitz on all bounded subsets of U there exists a number $\bar{L} > 1$ such that

$$|f(z_1) - f(z_2)| \leq \bar{L}\|z_1 - z_2\|$$

$$\text{for all } z_1, z_2 \in U \cap B_X(0, \bar{K} + 4).$$

Denote by $\mathcal{M}_C$ the set of all mappings $P : X \to C$ such that ,

$$Pz = z \text{ for all } z \in C,$$

$$\|Pz - x\| \leq \|z - x\| \text{ for all } x \in C \text{ and all } z \in X.$$

Theorem 3 *Assume that*

$$K_1 \geq \bar{K} + 4,\ L_1 \geq \bar{L},$$

$\delta_f, \delta_C \in (0, 1]$,

$$|f(z_1) - f(z_2)| \leq L_1\|z_1 - z_2\|$$

$$\textit{for all } z_1, z_2 \in B_X(0, 3K_1 + 2) \cap U,$$

$$\alpha \in (0, (1 + \bar{L})^{-2})$$

and that

$$\delta_f(\bar{K} + 3K_1 + 2 + L_1) \leq \alpha,$$

$$\delta_C(\bar{K} + 3K_1 + L_1 + 3) \leq \alpha.$$

Let $T \geq 2$ be an integer

$$\{P_t\}_{t=0}^{T-1} \subset \mathcal{M}_C,$$

$$\{x_t\}_{t=0}^{T} \subset U,\ \{\xi_t\}_{t=0}^{T-1} \subset X,$$

$$\|x_0\| \leq K_1,$$

$$B_X(x_0, \delta_C) \cap C \neq \emptyset$$

and that for $t = 0, \dots, T-1$,

$$B_X(\xi_t, \delta_f) \cap \partial f(x_t) \neq \emptyset,$$

$$\|x_{t+1} - P_t(x_t - \alpha\xi_t)\| \leq \delta_C.$$

Then

$$\|x_t\| \leq 2\bar{K} + K_1,\ t = 0, \dots, T$$

and

$$\min\{f(x_t):\ t=0,\dots,T-1\}-\inf(f,C),$$

$$f(T^{-1}\sum_{i=0}^{T-1}x_i)-\inf(f,C)$$

$$\leq T^{-1}\sum_{i=0}^{T-1}f(x_i)-\inf(f,C)$$

$$\leq (2T\alpha)^{-1}(K_1+\bar{K})^2+L_1^2\alpha$$

$$+\alpha^{-1}\delta_C(\bar{K}+3K_1+L_1+3)+\delta_f(3K_1+\bar{K}+L_1+2).$$

Theorem 4 *Assume that*

$$K_1 \geq \bar{K} + 4,\ L_1 \geq \bar{L},\ r_0 \in (0, 1],$$

$$B_X(z, r_0) \subset U,\ z \in C,$$

$$|f(z_1) - f(z_2)| \leq L_1 \|z_1 - z_2\|$$

$$\textit{for all } z_1, z_2 \in B_X(0, 3K_1 + 1) \cap U,$$

$$\Delta \in (0, r_0],\ \delta_f, \delta_C \in (0, 2^{-1} r_0],$$

$$\alpha \in (0, (\bar{L} + 3)^{-2}],$$

and that

$$\delta_f(3\bar{K} + K_1 + 4 + L_1) \leq \alpha,$$

$$\delta_C(3\bar{K} + K_1 + L_1 + 2) \leq \alpha.$$

Let $T \geq 2$ *be an integer*

$$\{P_t\}_{t=0}^{T-1} \subset \mathcal{M}_C,$$

$$\{x_t\}_{t=0}^{T} \subset U,\ \{\xi_t\}_{t=0}^{T-1} \subset X,$$

$$\|x_0\| \leq K_1,$$

$$B_X(x_0, \delta_C) \cap C \neq \emptyset$$

and that for $i = 1, \dots, T-1$,

$$B_X(\xi_t, \delta_f) \cap \partial_\Delta f(x_t) \neq \emptyset,$$

$$\|x_{t+1} - P_t(x_t - \alpha \xi_t)\| \leq \delta_C.$$

Then

$$\|x_t\| \leq 2\bar{K} + K_1,\ t = 0, \dots, T$$

and

$$\min\{f(x_t):\ t=0,\dots,T-1\}-\inf(f,C),$$

$$f(T^{-1}\sum_{i=0}^{T-1} x_i)-\inf(f,C)$$

$$\leq T^{-1}\sum_{i=0}^{T-1} f(x_i)-\inf(f,C)$$

$$\leq (2T\alpha)^{-1}(K_1+\bar{K})^2+(L_1+2)^2\alpha$$

$$+\alpha^{-1}\delta_C(3\bar{K}+K_1+L_1+2)$$

$$+\Delta+\delta_f(K_1+3\bar{K}+L_1+4).$$

Zero-sum games with two players

Let $(X, \langle\cdot,\cdot\rangle)$, $(Y, \langle\cdot,\cdot\rangle)$ be Hilbert spaces equipped with the complete norms $\|\cdot\|$ which are induced by their inner products. Let C be a nonempty closed convex subset of X, D be a nonempty closed convex subset of Y, U be an open convex subset of X and V be an open convex subset of Y such that

$$C \subset U, \ D \subset V$$

and let a function $f : U \times V \to R^1$ possess the following properties:

(i) for each $v \in V$, the function $f(\cdot, v) : U \to R^1$ is convex;

(ii) for each $u \in U$, the function $f(u, \cdot) : V \to R^1$ is concave.

Assume that a function $\phi : R^1 \to [0, \infty)$ is bounded on all bounded sets and that positive numbers M_1, M_2, L_1, L_2 satisfy

$$C \subset B_X(0, M_1),$$

$$D \subset B_Y(0, M_2),$$

$$|f(u_1, v) - f(u_2, v)| \leq L_1 \|u_1 - u_2\|$$

$$\text{for all } v \in V \text{ and all } u_1, u_2 \in U,$$

$$|f(u, v_1) - f(u, v_2)| \leq L_2 \|v_1 - v_2\|$$

$$\text{for all } u \in U \text{ and all } v_1, v_2 \in V.$$

Let

$$x_* \in C \text{ and } y_* \in D$$

satisfy

$$f(x_*, y) \leq f(x_*, y_*) \leq f(x, y_*)$$

for each $x \in C$ and each $y \in D$.

The following result was obtained in [2].

Prop 1 *Let* T *be a natural number,* $\delta_C,\ \delta_D \in (0,1]$, $\{a_t\}_{t=0}^{T} \subset (0,\infty)$ *and let* $\{b_{t,1}\}_{t=0}^{T}$, $\{b_{t,2}\}_{t=0}^{T} \subset (0,\infty)$. *Assume that* $\{x_t\}_{t=0}^{T+1} \subset U$, $\{y_t\}_{t=0}^{T+1} \subset V$, *for each* $t \in \{0,\dots,T+1\}$,

$$B_X(x_t,\delta_C)\cap C \neq \emptyset,$$

$$B_Y(y_t,\delta_D)\cap D \neq \emptyset,$$

for each $z \in C$ *and each* $t \in \{0,\dots,T\}$,

$$a_t(f(x_t,y_t) - f(z,y_t))$$

$$\leq \phi(\|z - x_t\|) - \phi(\|z - x_{t+1}\|) + b_{t,1}$$

and that for each $v \in D$ *and each* $t \in \{0,\dots,T\}$,

$$a_t(f(x_t,v) - f(x_t,y_t))$$

$$\leq \phi(\|v - y_t\|) - \phi(\|v - y_{t+1}\|) + b_{t,2}.$$

Let

$$\widehat{x}_T = (\sum_{i=0}^{T} a_i)^{-1} \sum_{t=0}^{T} a_t x_t,$$

$$\widehat{y}_T = (\sum_{i=0}^{T} a_i)^{-1} \sum_{t=0}^{T} a_t y_t.$$

Then $B_X(\widehat{x}_T, \delta_C) \cap C \neq \emptyset$,

$$B_Y(\widehat{y}_T, \delta_D) \cap D \neq \emptyset,$$

$$|(\sum_{t=0}^{T} a_t)^{-1} \sum_{t=0}^{T} a_t f(x_t, y_t) - f(x_*, y_*)|$$

$$\leq (\sum_{t=0}^{T} a_t)^{-1} \max\{\sum_{t=0}^{T} b_{t,1},\ \sum_{t=0}^{T} b_{t,2}\}$$

$$+ \max\{L_1\delta_C,\ L_2\delta_D\}$$

$$+(\sum_{t=0}^{T} a_t)^{-1} \sup\{\phi(s) :$$

$$s \in [0, \max\{2M_1,\ 2M_2\} + 1]\},$$

$$|f(\widehat{x}_T,\widehat{y}_T)-(\sum_{t=0}^T a_t)^{-1}\sum_{t=0}^T a_t f(x_t,y_t)|$$

$$\leq (\sum_{t=0}^T a_t)^{-1}\sup\{\phi(s):\ s\in[0,\max\{2M_1,\ 2M_2\}+1]\}$$

$$+(\sum_{t=0}^T a_t)^{-1}\max\{\sum_{t=0}^T b_{t,1},\ \sum_{t=0}^T b_{t,2}\}$$

$$+\max\{L_1\delta_C,\ L_2\delta_D\}$$

and for each $z\in C$ and each $v\in D$,

$$f(z,\widehat{y}_T)\geq f(\widehat{x}_T,\widehat{y}_T)$$

$$-2(\sum_{t=0}^T a_t)^{-1}\sup\{\phi(s):\ s\in[0,\max\{2M_1,\ 2M_2\}+1]\}$$

$$-2(\sum_{t=0}^T a_t)^{-1}\max\{\sum_{t=0}^T b_{t,1},\ \sum_{t=0}^T b_{t,2}\}$$

$$-\max\{L_1\delta_C,\ L_2\delta_D\},$$

$$f(\widehat{x}_T, v) \le f(\widehat{x}_T, \widehat{y}_T)$$

$$+2(\sum_{t=0}^{T} a_t)^{-1} \sup\{\phi(s) :\ s \in [0, \max\{2M_1,\ 2M_2\}+1]\}$$

$$+2(\sum_{t=0}^{T} a_t)^{-1} \max\{\sum_{t=0}^{T} b_{t,1},\ \sum_{t=0}^{T} b_{t,2}\}$$

$$+\max\{L_1\delta_C,\ L_2\delta_D\}.$$

Zero-sum games on bounded sets

Let $(X, \langle\cdot,\cdot\rangle)$, $(Y, \langle\cdot,\cdot\rangle)$ be Hilbert spaces equipped with the complete norms $\|\cdot\|$ which are induced by their inner products. Let C be a nonempty closed convex subset of X, D be a nonempty closed convex subset of Y, U be an open convex subset of X and V be an open convex subset of Y such that

$$C \subset U,\ D \subset V.$$

For each concave function $g : V \to R^1$, each $x \in V$ and each $\epsilon > 0$, set

$$\partial g(x) = \{l \in Y :\ \langle l, y - x\rangle$$

$$\geq g(y) - g(x) \text{ for all } y \in V\},$$

$$\partial_\epsilon g(x) = \{l \in Y :\ \langle l, y - x\rangle + \epsilon$$

$$\geq g(y) - g(x) \text{ for all } y \in V\}.$$

Suppose that there exist $L_1, L_2, M_1, M_2 > 0$ such that $C \subset B_X(0, M_1)$, $D \subset B_Y(0, M_2)$, a function $f : U \times V \to R^1$ possesses the following properties:

(i) for each $v \in V$, the function $f(\cdot, v) : U \to R^1$ is convex;

(ii) for each $u \in U$, the function $f(u, \cdot) : V \to R^1$ is concave,

for each $v \in V$,

$$|f(u_1, v) - f(u_2, v)| \leq L_1 \|u_1 - u_2\|$$

for all $u_1, u_2 \in U$

and that for each $u \in U$,

$$|f(u, v_1) - f(u, v_2)| \leq L_2 \|v_1 - v_2\|$$

for all $v_1, v_2 \in V$.

For each $(\xi, \eta) \in U \times V$ and each $\epsilon > 0$, set

$$\partial_x f(\xi, \eta) = \{l \in X :$$

$$f(y, \eta) - f(\xi, \eta) \geq \langle l, y - \xi \rangle$$

$$\text{for all } y \in U\},$$

$$\partial_y f(\xi, \eta) = \{l \in Y :$$

$$\langle l, y - \eta \rangle \geq f(\xi, y) - f(\xi, \eta)$$

$$\text{for all } y \in V\},$$

$$\partial_{x,\epsilon} f(\xi, \eta) = \{l \in X :$$

$$f(y, \eta) - f(\xi, \eta) + \epsilon \geq \langle l, y - \xi \rangle$$

$$\text{for all } y \in U\},$$

$$\partial_{y,\epsilon} f(\xi, \eta) = \{l \in Y :$$

$$\langle l, y - \eta \rangle + \epsilon \geq f(\xi, y) - f(\xi, \eta)$$

$$\text{for all } y \in V\}.$$

In view of properties (i) and (ii), for each $\xi \in U$ and each $\eta \in V$,

$$\emptyset \neq \partial_x f(\xi, \eta) \subset B_X(0, L_1),$$

$$\emptyset \neq \partial_y f(\xi, \eta) \subset B_Y(0, L_2).$$

Let

$$x_* \in C \text{ and } y_* \in D$$

satisfy

$$f(x_*, y) \le f(x_*, y_*) \le f(x, y_*)$$

for each $x \in C$ and each $y \in D$.

Denote by $\mathcal{M}_U$ the set of all mappings $P : X \to X$ such that

$$Px = x,\ x \in C,$$

$$\|Px - z\| \le \|x - z\| \text{ for all } x \in X \text{ and all } z \in C$$

and by $\mathcal{M}_V$ the set of all mappings $P : Y \to Y$ such that

$$Py = y, \ y \in D,$$

$$\|Py - z\| \le \|y - z\| \text{ for all } y \in Y \text{ and all } z \in C.$$

Let $\delta_{f,1}, \delta_{f,2}, \delta_C, \delta_D \in (0, 1]$ and $\{\alpha_k\}_{k=0}^{\infty} \subset (0, \infty)$.

Let us describe our algorithm.

Subgradient projection algorithm for zero-sum games

Initialization: select arbitrary $x_0 \in U$ and $y_0 \in V$.

Iterative step: given current iteration vectors $x_t \in U$ and $y_t \in V$ calculate

$$\xi_t \in \partial_x f(x_t, y_t) + B_X(0, \delta_{f,1}),$$

$$\eta_t \in \partial_y f(x_t, y_t) + B_Y(0, \delta_{f,2})$$

and the next pair of iteration vectors $x_{t+1} \in U$, $y_{t+1} \in V$ such that

$$\|x_{t+1} - P_t(x_t - \alpha_t \xi_t)\| \le \delta_C,$$

$$\|y_{t+1} - Q_t(y_t + \alpha_t \eta_t)\| \le \delta_D,$$

where $P_t \in \mathcal{M}_U$, $Q_t \in \mathcal{M}_V$.

Theorem 5 *Let* $\delta_{f,1}, \delta_{f,2}, \delta_C, \delta_D \in (0,1]$, $\{\alpha_k\}_{k=0}^{\infty} \subset (0,\infty)$,

$$\{P_t\}_{t=0}^{\infty} \subset \mathcal{M}_U,\ P_t(X) = C,\ t = 0,1,\dots,$$

$$\{Q_t\}_{t=0}^{\infty} \subset \mathcal{M}_V,\ Q_t(Y) = D,\ t = 0,1,\dots.$$

Assume that $\{x_t\}_{t=0}^{\infty} \subset U$, $\{y_t\}_{t=0}^{\infty} \subset V$, $\{\xi_t\}_{t=0}^{\infty} \subset X$, $\{\eta_t\}_{t=0}^{\infty} \subset Y$,

$$B_X(x_0, \delta_C) \cap C \neq \emptyset,\ B_Y(y_0, \delta_D) \cap D \neq \emptyset$$

and that for each integer $t \geq 0$,

$$\xi_t \in \partial_x f(x_t, y_t) + B_X(0, \delta_{f,1}),$$

$$\eta_t \in \partial_y f(x_t, y_t) + B_Y(0, \delta_{f,2}),$$

$$\|x_{t+1} - P_t(x_t - \alpha_t \xi_t)\| \leq \delta_C$$

and

$$\|y_{t+1} - Q_t(y_t + \alpha_t \eta_t)\| \leq \delta_D.$$

For each integer $t \geq 0$ set

$$b_{t,1} = \alpha_t^2 L_1^2 + \delta_C(2M_1 + L_1 + 3) + \alpha_t \delta_{f,1}(2M_1 + L_1 + 2),$$

$$b_{t,2} = \alpha_t^2 L_2^2 + \delta_D(2M_2 + L_2 + 3) + \alpha_t \delta_{f,2}(2M_2 + L_2 + 2).$$

Let for each natural number T,

$$\widehat{x}_T = (\sum_{i=0}^{T} \alpha_t)^{-1} \sum_{t=0}^{T} \alpha_t x_t,$$

$$\widehat{y}_T = (\sum_{i=0}^{T} \alpha_t)^{-1} \sum_{t=0}^{T} \alpha_t y_t.$$

Then for each natural number T,

$$B_X(\widehat{x}_T, \delta_C) \cap C \neq \emptyset,\ B_Y(\widehat{y}_T, \delta_D) \cap D \neq \emptyset,$$

$$|(\sum_{t=0}^{T} \alpha_t)^{-1} \sum_{t=0}^{T} \alpha_t f(x_t, y_t) - f(x_*, y_*)|$$

$$\leq (\sum_{t=0}^{T} \alpha_t)^{-1} \max\{\sum_{t=0}^{T} b_{t,1},$$

$$\sum_{t=0}^{T} b_{t,2}\} + \max\{L_1\delta_C,\ L_2\delta_D\}$$

$$+(2\sum_{t=0}^{T} \alpha_t)^{-1} \max\{(2M_1,\ 2M_2\} + 1)^2,$$

$$|f(\widehat{x}_T,\widehat{y}_T) - (\sum_{t=0}^{T} \alpha_t)^{-1} \sum_{t=0}^{T} \alpha_t f(x_t,y_t)|$$

$$\leq (2\sum_{t=0}^{T} \alpha_t)^{-1}(\max\{2M_1, 2M_2\} + 1)^2$$

$$+ \max\{L_1\delta_C,\ L_2\delta_D\}$$

$$+(\sum_{t=0}^{T} \alpha_t)^{-1} \max\{\sum_{t=0}^{T} b_{t,1}, \sum_{t=0}^{T} b_{t,2}\}$$

and for each $z \in C$ and each $v \in D$,

$$f(z, \widehat{y}_T) \geq f(\widehat{x}_T, \widehat{y}_T)$$

$$-(\sum_{t=0}^{T} \alpha_t)^{-1}(\max\{2M_1, 2M_2\} + 1)^2$$

$$-\max\{L_1\delta_C,\ L_2\delta_D\}$$

$$-2(\sum_{t=0}^{T} \alpha_t)^{-1} \max\{\sum_{t=0}^{T} b_{t,1}, \sum_{t=0}^{T} b_{t,2}\}$$

$$f(\widehat{x}_T, v) \leq f(\widehat{x}_T, \widehat{y}_T)$$

$$+(\sum_{t=0}^{T} \alpha_t)^{-1}(\max\{2M_1, 2M_2\} + 1)^2$$

$$+\max\{L_1\delta_C,\ L_2\delta_D\}$$

$$+2(\sum_{t=0}^{T} \alpha_t)^{-1} \max\{\sum_{t=0}^{T} b_{t,1}, \sum_{t=0}^{T} b_{t,2}\}.$$

Theorem 6 *Let* $r_1, r_2 > 0$,

$$B_X(z, 2r_1) \subset U \text{ for all } z \in C,$$

$$B_Y(u, 2r_2) \subset V \text{ for all } u \in D,$$

$\Delta_1, \Delta_2 > 0$, $\delta_{f,1}, \delta_{f,2}, \delta_C, \delta_D \in (0, 1]$,

$$\delta_C \le r_1, \ \delta_D \le r_2,$$

and $\{\alpha_t\}_{k=0}^{\infty} \subset (0, 1]$,

$$\{P_t\}_{t=0}^{\infty} \subset \mathcal{M}_U, \ \{Q_t\}_{t=0}^{\infty} \subset \mathcal{M}_V,$$

$P_t(X) = C$, $t = 0, 1, \dots$, $Q_t(Y) = D$, $t = 0, 1, \dots$.

Assume that $\{x_t\}_{t=0}^{\infty} \subset U$, $\{y_t\}_{t=0}^{\infty} \subset V$, $\{\xi_t\}_{t=0}^{\infty} \subset X$, $\{\eta_t\}_{t=0}^{\infty} \subset Y$,

$$B_X(x_0, \delta_C) \cap C \ne \emptyset, \ B_Y(y_0, \delta_D) \cap D \ne \emptyset$$

and that for each integer $t \geq 0$,

$$B_X(\xi_t, \delta_{f,1}) \cap \partial_{x,\Delta_1} f(x_t, y_t) \neq \emptyset,$$

$$B_Y(\eta_t, \delta_{f,2}) \cap \partial_{y,\Delta_2} f(x_t, y_t) \neq \emptyset,$$

$$\|x_{t+1} - P_t(x_t - a_t \xi_t)\| \leq \delta_C$$

and

$$\|y_{t+1} - Q_t(y_t + a_t \eta_t)\| \leq \delta_D.$$

For each integer $t \geq 0$ set

$$b_{t,1} = \alpha_t \Delta_1 + 2^{-1}\alpha_t^2 (L_1 + \Delta_1 r_1^{-1})$$

$$+\delta_C (2M_1 + L_1 + 3 + \Delta_1 r_1^{-1})$$

$$+\alpha_t \delta_{f,1} (2M_1 + L_1 + 2 + \Delta_1 r_1^{-1}),$$

$$b_{t,2} = \alpha_t \Delta_2 + 2^{-1}\alpha_t^2 (L_2 + \Delta_2 r_2^{-1})$$

$$+\delta_D (2M_2 + L_2 + 3 + \Delta_2 r_2^{-1})$$

$$+\alpha_t \delta_{f,2} (2M_2 + L_2 + 2 + \Delta_2 r_2^{-1}).$$

Let for each natural number T,

$$\widehat{x}_T = (\sum_{i=0}^{T} \alpha_t)^{-1} \sum_{t=0}^{T} \alpha_t x_t,$$

$$\widehat{y}_T = (\sum_{i=0}^{T} \alpha_t)^{-1} \sum_{t=0}^{T} \alpha_t y_t.$$

Then for each natural number T,

$$B_X(\widehat{x}_T, \delta_C) \cap C \neq \emptyset,\ B_Y(\widehat{y}_T, \delta_D) \cap D \neq \emptyset,$$

$$|(\sum_{t=0}^{T} \alpha_t)^{-1} \sum_{t=0}^{T} \alpha_t f(x_t, y_t) - f(x_*, y_*)|$$

$$\leq (\sum_{t=0}^{T} \alpha_t)^{-1} \max\{\sum_{t=0}^{T} b_{t,1},$$

$$\sum_{t=0}^{T} b_{t,2}\} + \max\{L_1\delta_C,\ L_2\delta_D\}$$

$$+(2\sum_{t=0}^{T} \alpha_t)^{-1} \max\{(2M_1,\ 2M_2\} + 1)^2,$$

$$|f(\widehat{x}_T, \widehat{y}_T) - (\sum_{t=0}^{T} \alpha_t)^{-1} \sum_{t=0}^{T} \alpha_t f(x_t, y_t)|$$

$$\leq (2 \sum_{t=0}^{T} \alpha_t)^{-1} (\max\{2M_1, 2M_2\} + 1)^2$$

$$+ \max\{L_1 \delta_C,\ L_2 \delta_D\}$$

$$+ (\sum_{t=0}^{T} \alpha_t)^{-1} \max\{\sum_{t=0}^{T} b_{t,1}, \sum_{t=0}^{T} b_{t,2}\}$$

and for each $z \in C$ and each $v \in D$,

$$f(z, \widehat{y}_T) \geq f(\widehat{x}_T, \widehat{y}_T)$$

$$-(\sum_{t=0}^{T} \alpha_t)^{-1} (\max\{2M_1, 2M_2\} + 1)^2$$

$$- \max\{L_1 \delta_C,\ L_2 \delta_D\}$$

$$-2(\sum_{t=0}^{T} \alpha_t)^{-1} \max\{\sum_{t=0}^{T} b_{t,1}, \sum_{t=0}^{T} b_{t,2}\},$$

$$f(\widehat{x}_T, v) \le f(\widehat{x}_T, \widehat{y}_T)$$

$$+(\sum_{t=0}^{T} \alpha_t)^{-1}(\max\{2M_1, 2M_2\} + 1)^2$$

$$+\max\{L_1\delta_C,\ L_2\delta_D\}$$

$$+2(\sum_{t=0}^{T} \alpha_t)^{-1} \max\{\sum_{t=0}^{T} b_{t,1}, \sum_{t=0}^{T} b_{t,2}\}.$$

We are interested in the optimal choice of α_t, $t = 0, 1, \dots, T$. Let T be a natural number and $A_T = \sum_{t=0}^{T} \alpha_t$ be given. In order to make the best choice of α_t, $t = 0, \dots, T$, we need to minimize the function $\sum_{t=0}^{T} \alpha_t^2$ on the set

$$\{\alpha = (\alpha_0, \dots, \alpha_T) \in R^{T+1} :\ \alpha_i \ge 0,$$

$$i = 0, \dots, T,\ \sum_{i=0}^{T} \alpha_i = A_T\}.$$

This function has a unique minimizer

$$\alpha_i = (T+1)^{-1} A_T \, , i = 0, \dots, T.$$

Let T be a natural number and $\alpha_t = \alpha$ for all $t = 0, \dots, T$. Now we will find the best $\alpha > 0$. In order to meet this goal we need to choose a which is a minimizer of the function

$$((T+1)\alpha)^{-1}(\max\{2M_1, 2M_2\} + 1)^2$$

$$+2\alpha^{-1}(T+1)^{-1} \max\{\sum_{t=0}^{T} b_{t,1}, \sum_{t=0}^{T} b_{t,2}\}$$

$$= ((T+1)\alpha)^{-1}(\max\{2M_1, 2M_2\} + 1)^2$$

$$+2\alpha^{-1}(T+1)^{-1}\max\{(T+1)(\alpha\Delta_1$$

$$+\delta_C(2M_1 + 3 + L_1 + \Delta_1 r_1^{-1}))$$

$$+2^{-1}\alpha^2(L_1 + \Delta_1 r_1^{-1})$$

$$+\alpha\delta_{f,1}(2M_1 + 2 + L_1 + \Delta_1 r_1^{-1}),$$

$$(T+1)(\alpha\Delta_2 + \delta_D(2M_2 + 3 + L_2 + \Delta_2 r_2^{-1}))$$

$$+2^{-1}\alpha^2(L_2+\Delta_2 r_2^{-1})+\alpha\delta_{f,2}(2M_2+2+L_2+\Delta_2 r_2^{-1})\}$$

$$= ((T+1)\alpha)^{-1}(\max\{2M_1, 2M_2\} + 1)^2$$

$$+2\max\{\Delta_1 + \alpha^{-1}\delta_C(2M_1 + 3 + L_1 + \Delta_1 r_1^{-1})$$

$$+2^{-1}\alpha(L_1+\Delta_1 r_1^{-1})+\delta_{f,1}(2M_1+2+L_1+\Delta_1 r_1^{-1}),$$

$$\Delta_2 + \alpha^{-1}\delta_D(2M_2 + 3 + L_2 + \Delta_2 r_2^{-1})$$

$$+2^{-1}\alpha(L_2+\Delta_2 r_2^{-1})+\delta_{f,2}(2M_2+2+L_2+\Delta_2 r_2^{-1})\}$$

$$\leq ((T+1)\alpha)^{-1}(\max\{2M_1, 2M_2\}+1)^2$$

$$+2\max\{\Delta_1, \Delta_2\}$$

$$+2\max\{\delta_{f,1}(2M_1+2+L_1+\Delta_1 r_1^{-1}),$$

$$\delta_{f,2}(2M_2+2+L_2+\Delta_2 r_2^{-1})\}$$

$$+2\alpha^{-1}\max\{\delta_C(2M_1+3+L_1+\Delta_1 r_1^{-1}),$$

$$\delta_D(2M_2+3+L_2+\Delta_2 r_2^{-1})\}$$

$$+\alpha\max\{L_1+\Delta_1 r_1^{-1},\ L_2+\Delta_2 r_2^{-1}\}.$$

Since T can be arbitrary large, we need to find a minimizer of the function

$$\phi(\alpha) := 2\alpha^{-1} \max\{\delta_C(2M_1 + 3 + L_1 + \Delta_1 r_1^{-1}),$$

$$\delta_D(2M_2 + 3 + L_2 + \Delta_2 r_2^{-1})\}$$

$$+\alpha \max\{L_1 + \Delta_1 r_1^{-1},\ L_2 + \Delta_2 r_2^{-1}\},\ \alpha > 0.$$

This function has a minimizer

$$\alpha_* = 2^{1/2} \max\{\delta_C(2M_1 + 3 + L_1 + \Delta_1 r_1^{-1}),$$

$$\delta_D(2M_2 + 3 + L_2 + \Delta_2 r_2^{-1})\}^{1/2}$$

$$\times \max\{L_1 + \Delta_1 r_1^{-1},\ L_2 + \Delta_2 r_2^{-1}\}^{-1/2}$$

and

$$\phi(\alpha_*) = 2^{3/2} \max\{\delta_C(2M_1 + 3 + L_1 + \Delta_1 r_1^{-1}),$$
$$\delta_D(2M_2 + 3 + L_2 + \Delta_2 r_2^{-1})\}^{1/2}$$
$$\times \max\{L_1 + \Delta_1 r_1^{-1},\ L_2 + \Delta_2 r_2^{-1}\}^{1/2}.$$

The appropriate choice of T, it should be at the same order as $\max\{\delta_C,\ \delta_D\}^{-1}$.

1. A. J. Zaslavski, Numerical optimization with computational errors, Springer, 2016.

2. A. J. Zaslavski, Convex optimization with computational errors, Springer, 2020.

3. A. J. Zaslavski, *The projected subgradient algorithm in convex optimization*, SpringerBriefs in Optimization, 2020.