数理解析研究所講究録 1362 巻 2004 年 88-94

Two Generalizations of the Projected Gradient Method for Convexly Constrained Inverse Problems —

Hybrid steepest descent method, Adaptive projected subgradient method

山田 功 (Isao Yamada) † 小倉信彦 (Nobuhiko Ogura) ‡

† Dept. of Communications & Integrated Systems, Tokyo Institute of Technology

‡ Precision and Intelligence Laboratory, Tokyo Institute of Technology

Abstract In this paper, we present a brief review on the central results of two generalizations of a classical convex optimization technique named the *projected gradient method* [1, 2]. The 1st generalization has been made by extending the convex projection operator, used in the projected gradient method, to the (quasi-)nonexpansive mapping in a real Hilbert space. By this generalization, we deduce the *hybrid steepest descent method* [3–10] (see also [11]) that can minimize the convex cost function over the fixed point set of nonexpansive mapping [3–9, 11] (these results can also be interpreted as generalizations of fixed point iterations found for example in [12–15]) or, more generally, over the fixed point set of quasi-nonexpansive mapping [10]. Since (i) the solution set of wide range of convexly constrained inverse problems, for example in *signal processing and image reconstruction*, can be characterized as the fixed point set of certain nonexpansive mapping [5, 6, 9, 16–18], and (ii) subgradient projection operator and its variations are typical examples of quasi-nonexpansive mapping [10, 19], the hybrid steepest descent method has rich applications in broad range of mathematical sciences and engineerings. The 2nd generalization has been made for the *Polyak's subgradient algorithm* [20] that was originally developed as a version, of the *projected gradient method*, for unsmooth convex optimization problem with a fixed target value. By extending the *Polyak's subgradient algorithm* to the case where the convex cost function itself keeps changing in the whole process, we deduce the *adaptive projected subgradient method* [21–23] that can minimize asymptotically the sequence of unsmooth nonnegative convex cost functions. The *adaptive projected subgradient method* can serve as a unified guiding principle of a wide range of *set theoretic adaptive filtering schemes* [24–30] for nonstationary random processes. The great flexibilities in the choice of (quasi-)nonexpansive mapping as well as unsmooth convex cost functions in the proposed methods yield naturally inherently parallel structures (in the sense of [31]).

1 Preliminaries

Let $\mathcal{H}$ be a real Hilbert space equipped with an inner product $\langle\cdot,\cdot\rangle$ and its induced norm $\|\cdot\|$. For a continuous convex function $\Phi:\mathcal{H}\to\mathbb{R}$, the *subdifferential* of Φ at $\forall y\in\mathcal{H}$, the set of all *subgradients* of Φ at y: $\partial\Phi(y) := \{g\in\mathcal{H} \mid \langle x-y,g\rangle+\Phi(y)\le\Phi(x), \forall x\in\mathcal{H}\}$ is nonempty. The convex function $\Phi:\mathcal{H}\to\mathbb{R}$ has a unique subgradient at $y\in\mathcal{H}$ if Φ is Gâteaux differentiable at y. This unique subgradient is nothing but the Gâteaux differential $\Phi'(y)$. A *fixed point* of a mapping $T:\mathcal{H}\to\mathcal{H}$ is a point $x\in\mathcal{H}$ such that $T(x)=x$. $Fix(T) := \{x\in\mathcal{H} \mid T(x)=x\}$ denotes the fixed point set of T. A mapping $T:\mathcal{H}\to\mathcal{H}$ is called (i) *strictly contractive* if $\|T(x)-T(y)\|\le\kappa\|x-y\|$ for some $\kappa\in(0,1)$ and all $x,y\in\mathcal{H}$ [The *Banach-Picard fixed point theorem* guarantees the unique existence of the fixed point, say $x_*\in Fix(T)$, of T and the strong convergence of $(T^n(x_0))_{n\ge 0}$ to

†This work was supported in part by JSPS grants-in-Aid (C15500129).

x_* for any $x_0 \in \mathcal{H}$.]; (ii) *nonexpansive* if $\|T(x) - T(y)\| \le \|x - y\|$, $\forall x, y \in \mathcal{H}$; (iii) *firmly nonexpansive* if $\|T(x) - T(y)\|^2 \le \langle x - y, T(x) - T(y)\rangle$, $\forall x, y \in \mathcal{H}$ [32]; and (iv) *attracting nonexpansive* if T is nonexpansive with $Fix(T) \ne \emptyset$ and $\|T(x) - f\| < \|x - f\|$, $\forall f \in Fix(T)$ and $\forall x \notin Fix(T)$ [32]. Given a nonempty closed convex set $C \subset \mathcal{H}$, the mapping that assigns every point in $\mathcal{H}$ to its unique nearest point in C is called the *metric projection* or *convex projection* onto C and is denoted by P_C; i.e., $\|x - P_C(x)\| = d(x, C)$, where $d(x, C) := \inf_{y \in C} \|x - y\|$. P_C is firmly nonexpansive with $Fix(P_C) = C$ [32]. A mapping $T : \mathcal{H} \to \mathcal{H}$ is called *quasi-nonexpansive* if $\|T(x) - T(f)\| \le \|x - f\|$, $\forall (x, f) \in \mathcal{H} \times Fix(T)$. In this paper, for simplicity, a mapping $T : \mathcal{H} \to \mathcal{H}$ is called *attracting quasi-nonexpansive* if $Fix(T) \ne \emptyset$ and $\|T(x) - f\| < \|x - f\|$, $\forall (x, f) \in Fix(T)^C \times Fix(T)$. Moreover, a mapping $T : \mathcal{H} \to \mathcal{H}$ is called α*-averaged quasi-nonexpansive* if there exists $\alpha \in (0, 1)$ and a quasi-nonexpansive mapping $\mathcal{N} : \mathcal{H} \to \mathcal{H}$ such that $T = (1 - \alpha)I + \alpha\mathcal{N}$ (Note: $Fix(T) = Fix(\mathcal{N})$ holds automatically). In particular, 1/2-averaged quasi-nonexpansive mapping, which we specially call *firmly quasi-nonexpansive mapping.* Suppose that a continuous convex function $\Phi : \mathcal{H} \to \mathbb{R}$ satisfies $\mathrm{lev}_{\le 0}\Phi := \{x \in \mathcal{H} \mid \Phi(x) \le 0\} \ne \emptyset$. Then a mapping $T_{sp(\Phi)} : \mathcal{H} \to \mathcal{H}$ defined by

$$T_{sp(\Phi)} : x \mapsto \begin{cases} x - \frac{\Phi(x)}{\|g(x)\|^2} g(x) & \text{if } \Phi(x) > 0 \\ x & \text{if } \Phi(x) \le 0, \end{cases} \tag{1}$$

where g is a selection of the subdifferential $\partial\Phi$, is called a *subgradient projection relative to* Φ [32].The mapping $T_{sp(\Phi)} : \mathcal{H} \to \mathcal{H}$ is *firmly quasi-nonexpansive* and satisfies $Fix(T_{sp(\Phi)}) = \mathrm{lev}_{\le 0}\Phi$ (see for example [19]).

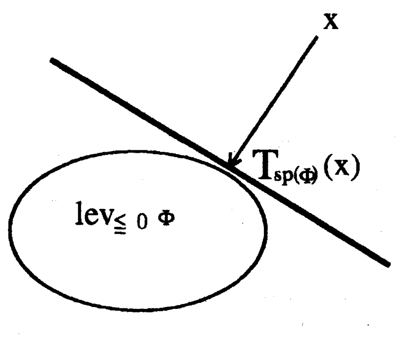

Figure 1: Subgradient projection relative to Φ

A mapping $\mathcal{F} : \mathcal{H} \to \mathcal{H}$ is called (i) *monotone* over $S \subset \mathcal{H}$ if $\langle \mathcal{F}(u) - \mathcal{F}(v), u - v\rangle \ge 0$, $\forall u, v \in S$. In particular, a mapping $\mathcal{F}$ which is monotone over $S \subset \mathcal{H}$ is called (ii) *paramonotone* over S if $\langle \mathcal{F}(u) - \mathcal{F}(v), u - v\rangle = 0 \Leftrightarrow \mathcal{F}(u) = \mathcal{F}(v)$, $\forall u, v \in S$ [34]; (iii) *uniformly monotone* over S if there exists a strictly monotone increasing continuous function $a : [0, \infty) \to [0, \infty)$, with $a(0) = 0$ and $a(t) \to \infty$ as $t \to \infty$, satisfying $\langle \mathcal{F}(u) - \mathcal{F}(v), u - v\rangle \ge a(\|u - v\|)\|u - v\|$ for all $u, v \in S$ [38]; (iv) η*-strongly monotone* over S if there exists $\eta > 0$ such that $\langle \mathcal{F}(u) - \mathcal{F}(v), u - v\rangle \ge \eta\|u - v\|^2$ for all $u, v \in S$ [38].

The *variational inequality problem* $VIP(\mathcal{F}, C)$ is defined as follows: given $\mathcal{F} : \mathcal{H} \to \mathcal{H}$ which is monotone over a nonempty closed convex set $C \subset \mathcal{H}$, find $u^* \in C$ such that $\langle v - u^*, \mathcal{F}(u^*)\rangle \ge 0$, $\forall v \in C$. If a function $\Theta : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ is convex over a closed convex set C and Gâteaux differentiable with derivative Θ' over an open set $U \supset C$, then Θ' is paramonotone over C. For such a Θ, the set $\Gamma := \{u \in C \mid \Theta(u) = \inf \Theta(C)\}$ is nothing but the solution set of $VIP(\Theta', C)$ [33]. Given $\mathcal{F} : \mathcal{H} \to \mathcal{H}$ which is monotone over a nonempty closed convex set C, $u^* \in C$ is a solution of VIP$(\mathcal{F}, C)$ if and only if $u^* \in Fix\,(P_C\,(I - \mu\mathcal{F}))$ for an arbitrarily fixed $\mu > 0$ (For related mathematical discussion in this section, the readers should consult, e.g., [6, 9, 19, 31–38]).

2 Hybrid Steepest Descent Method

Theorem 1 (Strong convergence for nonexpansive mapping [6, 9]**)** *Let $T : \mathcal{H} \to \mathcal{H}$ be a nonexpansive mapping with $Fix\,(T) \neq \emptyset$. Suppose that a mapping $\mathcal{F} : \mathcal{H} \to \mathcal{H}$ is κ-Lipschitzian and η-strongly monotone over $T(\mathcal{H})$. Then, by using any sequence $(\lambda_n)_{n\geq 1} \subset [0, \infty)$ satisfying (W1) $\lim_{n\to+\infty} \lambda_n = 0$, (W2) $\sum_{n\geq 1} \lambda_n = +\infty$, (W3) $\sum_{n\geq 1} |\lambda_n - \lambda_{n+1}| < +\infty$ [or $(\lambda_n)_{n\geq 1} \subset (0, \infty)$ satisfying (L1) $\lim_{n\to+\infty} \lambda_n = 0$, (L2) $\sum_{n\geq 1} \lambda_n = +\infty$, (L3) $\lim_{n\to\infty}(\lambda_n - \lambda_{n+1})\lambda_{n+1}^{-2} = 0$], the sequence $(u_n)_{n\geq 0}$ generated, with arbitrary $u_0 \in \mathcal{H}$, by*

$$u_{n+1} := T(u_n) - \lambda_{n+1}\mathcal{F}\,(T(u_n)) \tag{2}$$

converges strongly to the uniquely existing solution of the VIP: find $u^* \in Fix(T)$ such that $\langle v - u^*, \mathcal{F}(u^*)\rangle \geq 0$, $\forall v \in Fix(T)$. (Note: The condition *(L3)* was relaxed recently to $\lim_{n\to\infty} \frac{\lambda_n}{\lambda_{n+1}} = 1$ [11].) □

Theorem 1 is a generalization of a fixed point iteration [12–15] so called the *anchor method*:

$$u_{n+1} := \lambda_{n+1}a + (1 - \lambda_{n+1})T(u_n),$$

which converges strongly to $P_{Fix(T)}(a)$.

The *hybrid steepest descent method* (2) can be applied to more genaral monotone operators [7, 8] if $\dim(\mathcal{H}) < \infty$. Moreover, by the use of slowly changing sequence of nonexpansive mappings having same fixed point sets, a variation of the hybrid steepest descent method is gifted with notable robustness to the numerical errors possibly unavoidable in the iterative computations [9].

The next theorem shows that the hybrid steepest descent method can also be applied to the variational inequality problem over the fixed point set of *quasi-nonexpansive mappings*.

Definition 2 (Quasi-shrinking mapping[10]**)** *Suppose that $T : \mathcal{H} \to \mathcal{H}$ is quasi- nonexpansive with $Fix(T) \cap C \neq \emptyset$ for some closed convex set $C \subset \mathcal{H}$. Then $T : \mathcal{H} \to \mathcal{H}$ is called quasi-shrinking on $C(\subset \mathcal{H})$ if*

$$D : r \in [0, \infty) \mapsto \begin{cases} \displaystyle\inf_{u \in \rhd(Fix(T), r)\cap C} d(u, Fix(T)) - d(T(u), Fix(T)) & \\ \qquad\qquad \text{if } \rhd(Fix(T), r) \cap C \neq \emptyset \\ \infty \qquad\qquad \text{otherwise} \end{cases}$$

satisfies $D(r) = 0 \Leftrightarrow r = 0$, where $\rhd(Fix(T), r) := \{x \in \mathcal{H} \mid d(x, Fix(T)) \geq r\}$. □

Proposition 3 [10] *Suppose that a continuous convex function $\Phi : \mathcal{H} \to \mathbb{R}$ has $\mathrm{lev}_{\leq 0} \neq \emptyset$ and bounded subdiferential $\partial\Phi : \mathcal{H} \to 2^{\mathcal{H}}$, i.e., $\partial\Phi$ maps bounded sets to bounded sets. Define $T_\alpha := (1 - \alpha)I + \alpha T_{sp(\Phi)}$ for $\alpha \in (0, 2)$ [hence $Fix(T_\alpha) = \mathrm{lev}_{\leq 0}\Phi$: see (1) for the definition of $T_{sp(\Phi)}$]. Then, we have the followings:*

(a) *If a selection of subgradient of Φ, say $\Phi' : \mathcal{H} \to \mathcal{H}$, is uniformly monotone over $\mathcal{H}$, then T_α is quasi-shrinking on any nonempty bounded closed convex set C satisfying $C \cap \mathrm{lev}_{\leq 0}\Phi \neq \emptyset$.*

(b) *Assume* $\dim(\mathcal{H}) < \infty$. *Then* T_α *is quasi-shrinking on any nonempty bounded closed convex set* $C(\subset \mathcal{H})$ *satisfying* $C \cap lev_{\leq 0}\Phi \neq \emptyset$. □

Theorem 4 *(***Strong convergence for quasi-shrinking mapping** [10]*) Suppose that* $T : \mathcal{H} \to \mathcal{H}$ *is a quasi-nonexpansive mapping with* $Fix(T) \neq \emptyset$. *Let* $\mathcal{F} : \mathcal{H} \to \mathcal{H}$ *be* κ*-Lipschitzian and* η*-strongly monotone over* $T(\mathcal{H})$ *[Hence* $VIP(\mathcal{F}, Fix(T))$ *has its unique solution* $u^* \in Fix(T)$*]. Suppose also that there exists some* $(f, u_0) \in Fix(T) \times \mathcal{H}$ *for which* T *is quasi-shrinking on*

$$C_f(u_0) := \left\{ x \in \mathcal{H} \mid \|x - f\| \leq R_f := \max\left(\|u_0 - f\|, \frac{\|\mu \mathcal{F}(f)\|}{1 - \sqrt{1 - \mu(2\eta - \mu\kappa^2)}} \right) \right\}.$$

Then for any $\mu \in (0, \frac{2\eta}{\kappa^2})$ *and any* $(\lambda_n)_{n \geq 1} \subset [0, 1]$ *satisfying (H1)* $\lim_{n\to\infty} \lambda_n = 0$, *and (H2)* $\sum_{n \geq 1} \lambda_n = \infty$, *the sequence* $(u_n)_{n \geq 0}$, *generated by*

$$u_{n+1} := T(u_n) - \lambda_{n+1}\mu\mathcal{F}(T(u_n)),$$

converges strongly to u^*. □

If $\dim(\mathcal{H}) < \infty$, in a way similar to the discussions in [7–9], we can generalize Theorem 4 for application to more genaral monotone operators [10].

3 Adaptive Projected Subgradient Method

Theorem 5 (Adaptive Projected Subgradient Method [21, 22]**)** *Let* $\Theta_n : \mathcal{H} \to [0, \infty)$ $(\forall n \in \mathbb{N})$ *be a sequence of continuous convex functions and* $K \subset \mathcal{H}$ *a nonempty closed convex set. For an arbitrarily given* $u_0 \in K$, *the adaptive projected subgradient method produces a sequence* $(u_n)_{n\in\mathbb{N}} \subset K$ *by*

$$u_{n+1} := \begin{cases} P_K\left(u_n - \lambda_n \frac{\Theta_n(u_n)}{\|\Theta_n'(u_n)\|^2}\Theta_n'(u_n)\right) & \text{if } \Theta_n'(u_n) \neq 0, \\ u_n \text{ otherwise,} & \end{cases}$$

where $\Theta_n'(u_n) \in \partial\Theta_n(u_n)$ *and* $0 \leq \lambda_n \leq 2$. *Then the sequence* $(u_n)_{n\in\mathbb{N}}$ *satisfies the followings.*

(a) *(Monotone approximation) Suppose that*

$$u_n \notin \Omega_n := \{u \in K \mid \Theta_n(u) = \Theta_n^*\} \neq \emptyset,$$

where $\Theta_n^* := \inf_{u\in K} \Theta_n(u)$. [1] *Then, by using* $\forall \lambda_n \in \left(0, 2\left(1 - \frac{\Theta_n^*}{\Theta_n(u_n)}\right)\right)$, *we have*

$$\forall u^{*(n)} \in \Omega_n, \|u_{n+1} - u^{*(n)}\| < \|u_n - u^{*(n)}\|.$$

(b) *(Boundedness, Asymptotic optimality) Suppose*

$$\exists N_0 \in \mathbb{N} \text{ s.t.} \begin{cases} \Theta_n^* = 0, \forall n \geq N_0 \text{ and} \\ \Omega := \bigcap_{n \geq N_0} \Omega_n \neq \emptyset. \end{cases} \tag{3}$$

Then $(u_n)_{n\in\mathbb{N}}$ *is bounded. Moreover if we specially use* $\forall \lambda_n \in [\varepsilon_1, 2 - \varepsilon_2] \subset (0, 2)$, *we have* $\lim_{n\to\infty} \Theta_n(u_n) = 0$ *provided that* $(\Theta_n'(u_n))_{n\in\mathbb{N}}$ *is bounded.*

[1] In this case, $\Theta_n(u_n) > \Theta_n^* \geq 0$.

(c) *(Strong convergence) Assume (3) and* Ω *has some relative interior w.r.t. a hyperplane* $\Pi(\subset \mathcal{H})$, *i.e., there exist* $\widetilde{u} \in \Pi \cap \Omega$ *and* $\exists \varepsilon > 0$ *satisfying* $\{v \in \Pi \mid \|v - \widetilde{u}\| \leq \varepsilon\} \subset \Omega$. *Then, by using* $\forall \lambda_n \in [\varepsilon_1, 2 - \varepsilon_2] \subset (0, 2)$, $(u_n)_{n \in \mathbb{N}}$ *converges strongly to some* $\widehat{u} \in K$, *i.e.,* $\lim_{n\to\infty} \|u_n - \widehat{u}\| = 0$. *Moreover* $\lim_{n\to\infty} \Theta_n(\widehat{u}) = 0$ *if (i)* $(\Theta'_n(u_n))_{n\in\mathbb{N}}$ *is bounded and (ii) there exists bounded* $(\Theta'_n(\widehat{u}))_{n\in\mathbb{N}}$ *where* $\Theta'_n(\widehat{u}) \in \partial\Theta_n(\widehat{u}), \forall n \in \mathbb{N}$.

(d) *(A characterization of* $\widehat{u}$*) Assume the existence of some interior* $\widetilde{u}$ *of* Ω, *i.e., there exists* $\varrho > 0$ *satisfying* $\{v \in \mathcal{H} \mid \|v - \widetilde{u}\| \leq \varrho\} \subset \Omega$. *In addition to the conditions (i) and (ii) in (c), assume that there exists* $\delta > 0$ *satisfying*

$$\forall n \geq N_0, \forall u \in \Gamma \setminus (\mathrm{lev}_{\leq 0}\Theta_n), \exists \Theta'_n(u) \in \partial\Theta_n(u), \|\Theta'_n(u)\| \geq \delta,$$

where $\Gamma := \{(1-s)\widetilde{u} + s\hat{u} \in \mathcal{H} \mid s \in (0,1)\}$. *Then, by using* $\forall \lambda_n \in [\varepsilon_1, 2 - \varepsilon_2] \subset (0, 2)$, $\lim_{n\to\infty} u_n =: \widehat{u} \in \overline{\liminf_{n\to\infty} \Omega_n}$, *where* $\overline{\liminf_{n\to\infty} \Omega_n}$ *stands for the closure of* $\liminf_{n\to\infty} \Omega_n := \bigcup_{n\geq 0} \bigcap_{k\geq n} \Omega_k$. □

4 Concluding Remarks

In this paper, we briefly present central results on the *hybrid steepest descent method* and the *adaptive projected subgradient method* recently developed by our research group. For detailed mathematical discussions of the methods and their applications to inverse problems and signal processing problems, see [3–10, 16, 17, 21–23, 30] and references therein.

References

[1] A.A. Goldstein, Convex programming in Hilbert space, *Bull. Amer. Math. Soc.* **70** (1964) 709-710.

[2] E.S. Levitin and B.T. Polyak, Constrained Minimization Method, *USSR Computational Mathematics and Mathematical Physics* **6** (1966) 1-50.

[3] I. Yamada, N. Ogura, Y. Yamashita and K. Sakaniwa, Quadratic optimization of fixed points of nonexpansive mappings in Hilbert space, *Numer. Funct. Anal. Optim.* **19** (1998) 165-190 (see also *Technical Report of IEICE,* DSP96-106 (1996) 63-70).

[4] F. Deutsch and I. Yamada, Minimizing certain convex functions over the intersection of the fixed point sets of nonexpansive mappings, *Numer. Funct. Anal. Optim.* **19** (1998) 33-56.

[5] I. Yamada, Convex projection algorithm from POCS to Hybrid steepest descent method (in Japanese), *Journal of the IEICE,* **83** (2000) 616-623.

[6] I. Yamada, The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings, in *Inherently Parallel Algorithm for Feasibility and Optimization and Their Applications,* (D. Butnariu, Y. Censor, and S. Reich, Eds.) Elsevier, (2001) 473-504.

[7] N. Ogura and I. Yamada, Non-strictly convex minimization over the fixed point set of the asymptotically shrinking nonexpansive mapping, *Numer. Funct. Anal. Optim.* **23** (2002) 113-137.

[8] N. Ogura and I. Yamada, Non-strictly convex minimization over the bounded fixed point set of nonexpansive mapping, *Numer. Funct. Anal. Optim.* **24** (2003) 129-135.

[9] I. Yamada, N. Ogura and N. Shirakawa, A numerically robust hybrid steepest descent method for the convexly constrained generalized inverse problems, in *Inverse Problems, Image Analysis, and Medical Imaging,* (Z. Nashed and O. Scherzer, Eds.) *Contemporary Mathematics,* **313** Amer. Math. Soc., (2002) 269-305.

[10] I. Yamada and N. Ogura, Hybrid steepest descent method for the variational inequality problem over the fixed point sets of certain quasi-nonexpansive mappings, *Victoria International Conference 2004,* Wellington, (to be presented) Feb., (2004).

[11] H.K. Xu and T.H. Kim, Convergence of hybrid steepest descent methods for variational inequalities, *Journal of Optimization Theory and Applications,* vol.119, no.1, (2003) 185-201.

[12] B. Halpern, Fixed points of nonexpanding maps, *Bull. Amer. Math. Soc.* **73** (1967) 957-961.

[13] P.L. Lions, Approximation de points fixes de contractions, *C. R. Acad. Sci. Paris Sèrie A-B* **284** (1977) 1357-1359.

[14] R. Wittmann, Approximation of fixed points of nonexpansive mappings, *Arch. Math.* **58** (1992) 486-491.

[15] H.H. Bauschke, The approximation of fixed points of compositions of nonexpansive mappings in Hilbert space, *J. Math. Anal. Appl.* **202** (1996) 150-159.

[16] I. Yamada, Approximation of convexly constrained pseudoinverse by Hybrid Steepest Descent Method, *Proc. of IEEE ISCAS'99*, Florida, May (1999).

[17] K. Slavakis, I. Yamada and K. Sakaniwa, Computation of symmetric positive definite Toeplitz matrices by the Hybrid Steepest Descent Method, *Signal Processing* **83** (2003) 1135-1140.

[18] C. Byrn, A unified treatment of some iterative algorithms in signal processing and image reconstruction, *Inverse Problems,* **20** (2004) 103-120.

[19] H.H. Bauschke and P.L. Combettes, A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert space, *Math. Oper. Res.,* **26** (2001) 248-264.

[20] B. T. Polyak, "Minimization of unsmooth functionals," *USSR Comput. Math. Phys.* vol.9, (1969) 14-29.

[21] I. Yamada, Adaptive projected subgradient method — A unified view of projection based adaptive algorithms (in Japanese), *Journal of the IEICE,* **86** (2003) 654-658.

[22] I. Yamada and N. Ogura, Adaptive projected subgradient method and its applications to set theoretic adaptive filtering, *Proc. of the 37th Asilomar Conference on Signals, Systems and Computers,* California, Nov., (2003).

[23] I. Yamada, N. Ogura and M. Yukawa, Adaptive projected subgradient method and its acceleration techniques, *Proc. of IFAC Workshop on Adaptation and Learning in Control and Signal Processing (ALCOSP 04)* Yokohama, (to be presented) August, (2004).

[24] J. Nagumo and J. Noda, "A learning method for system identification," *IEEE Trans. Autom. Control*, vol.12, no.3, (1967) 282-287.

[25] A. E. Albert and L. S. Gardner, Jr, *Stochastic approximation and nonlinear regression*, MIT Press, 1967.

[26] T. Hinamoto and S. Maekawa, Extended theory of learning identification (in Japanese), *Trans. IEE Japan*, vol.95-C, no.10, (1975) 227-234.

[27] K. Ozeki and T. Umeda, An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties (in Japanese), *IEICE Trans.*, vol.67-A, no.5, (1984) 126-132.

[28] S. Gollamudi, S. Nagaraj, S. Kapoor and Y. H. Huang, Set-membership filtering and a set-membership normalized LMS algorithm with an adaptive step size, *IEEE Signal Processing Lett.*, vol.5, no.5, (1998) 111-114.

[29] L. Guo, A. Ekpenyong and Y. H. Huang, Frequency-domain adaptive filtering — A set-membership approach, *Proc. of the 37th Asilomar Conference on Signals, Systems and Computers,* California, Nov., (2003).

[30] I. Yamada, K. Slavakis and K. Yamada, An efficient robust adaptive filtering algorithm based on parallel subgradient projection techniques, *IEEE Trans. on Signal Processing,* vol.50, no.5, (2002) 1091-1101.

[31] Y. Censor and S.A Zenios, *Parallel Optimization: Theory, Algorithm, and Optimization* (Oxford University Press, 1997).

[32] H.H. Bauschke and J.M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Review* **38** (1996) 367-426.

[33] I. Ekeland and R. Themam, *Convex Analysis and Variational Problems , Classics in Applied Mathematics 28* (SIAM, 1999).

[34] Y. Censor, A.N. Iusem and S.A. Zenios, An interior point method with Bregman functions for the variational inequality problem with paramonotone operators, *Math. Programming* **81** (1998) 373-400.

[35] D. Butnariu and A.N. Iusem, *Totally Convex Functions for fixed point computation and infinite dimensional optimization* (Kluwer Academic Publishers, 2000).

[36] K. Goebel and S. Reich, *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings* (Dekker, New York and Basel, 1984).

[37] W. Takahashi, *Nonlinear Functional Analysis— Fixed Point Theory and its Applications* (Yokohama Publishers, 2000).

[38] E. Zeidler, *Nonlinear Functional Analysis and its Applications, II/B - Nonlinear Monotone Operators* (Springer-Verlag, 1990).