Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 894248, 5 pages
http://dx.doi.org/10.1155/2014/894248

Research Article

Stochastic Methods Based on $\mathscr{V}\mathscr{U}$-Decomposition Methods for Stochastic Convex Minimax Problems

Yuan Lu,[1] Wei Wang,[2] Shuang Chen,[3] and Ming Huang[3]

[1] *Normal College, Shenyang University, Shenyang 110044, China*
[2] *School of Mathematics, Liaoning Normal University, Dalian 116029, China*
[3] *School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China*

Correspondence should be addressed to Wei Wang; wei_0713@sina.com

Received 6 August 2014; Revised 29 November 2014; Accepted 29 November 2014; Published 4 December 2014

Academic Editor: Hamid R. Karimi

This paper applies sample average approximation (SAA) method based on $\mathscr{V}\mathscr{U}$-space decomposition theory to solve stochastic convex minimax problems. Under some moderate conditions, the SAA solution converges to its true counterpart with probability approaching one and convergence is exponentially fast with the increase of sample size. Based on the $\mathscr{V}\mathscr{U}$-theory, a superlinear convergent $\mathscr{V}\mathscr{U}$-algorithm frame is designed to solve the SAA problem.

1. Introduction

In this paper, the following stochastic convex minimax problem (SCMP) is considered:

$$\min_{x\in R^n} f(x), \tag{1}$$

where

$$f(x) = \max\{E[f_i(x,\xi)] : i = 0,\dots,m\}, \tag{2}$$

and the functions $f_i(x,\xi) : R^n \to R$, $i = 0,\dots,m$, are convex and C^2, $\xi : \Omega \to \Xi \subset R^n$ is a random vector defined on probability space (Ω, Υ, P); E denotes the mathematical expectation with respect to the distribution of ξ.

SCMP is a natural extension of deterministic convex minimax problems (CMP for short). The CMP has a number of important applications in operations research, engineering problems, and economic problems. While many practical problems only involve deterministic data, there are some important instances where problems data contains some uncertainties and consequently SCMP models are proposed to reflect the uncertainties.

A blanket assumption is made that, for every $x \in R^n$, $E[f_i(x,\xi)]$, $i = 0,\dots,m$, are well defined. Let $\xi^1,\dots,\xi^N$ be a sampling of ξ. A well-known approach based on the sampling is the so-called SAA method, that is, using sample average value of $f_i(x,\xi)$ to approximate its expected value because the classical law of large number for random functions ensures that the sample average value of $f_i(x,\xi)$ converges with probability 1 to $E[f_i(x,\xi)]$ when the sampling is independent and identically distributed (idd for short). Specifically, we can write down the SAA of our SCMP (1) as follows:

$$\min_{x\in R^n} \hat{f}^N(x), \tag{3}$$

where

$$\begin{gathered} \hat{f}^N(x) = \max\left\{\hat{f}_i^N(x) : i = 0,\dots,m\right\}, \\ \hat{f}_i^N(x) := \frac{1}{N}\sum_{j=1}^{N} f_i\left(x,\xi^j\right). \end{gathered} \tag{4}$$

The problem (3) is called the SAA problem and (1) the true problem.

The SAA method has been a hot topic of research in stochastic optimization. Pagnoncelli et al. [1] present the SAA method for chance constrained programming. Shapiro et al. [2] consider the stochastic generalized equation by using the SAA method. Xu [3] raises the SAA method for a class of stochastic variational inequality problems. Liu et al. [4]

give the penalized SAA methods for stochastic mathematical programs with complementarity constraints. Chen et al. [5] discuss the SAA methods based on Newton method to the stochastic variational inequality problem with constraint conditions. Since the objective functions of the SAA problems in the references talking above are smooth, then they can be solved by using Newton method.

More recently, new conceptual schemes have been developed, which are based on the $\mathcal{V}\mathcal{U}$-theory introduced in [6]; see else [7–11]. The idea is to decompose R^n into two orthogonal subspaces $\mathcal{V}$ and $\mathcal{U}$ at a point $\overline{x}$, where the nonsmoothness of f is concentrated essentially on $\mathcal{V}$ and the smoothness of f appears on the $\mathcal{U}$ subspace. More precisely, for a given $\overline{g} \in \partial f(\overline{x})$, where $\partial f(\overline{x})$ denotes the subdifferential of f at $\overline{x}$ in the sense of convex analysis, then R^n can be decomposed into direct sum of two orthogonal subspaces, that is, $R^n = \mathcal{U} \oplus \mathcal{V}$, where $\mathcal{V} = \text{lin}(\partial f(\overline{x}) - \overline{g})$, and $\mathcal{U} = \mathcal{V}^{\perp}$. As a result an algorithm frame can be designed for the SAA problem that makes a step in the $\mathcal{V}$ space, followed by a $\mathcal{U}$-Newton step in order to obtain superlinear convergence. A $\mathcal{V}\mathcal{U}$-space decomposition method for solving a constrained nonsmooth convex program is presented in [12]. A decomposition algorithm based on proximal bundle-type method with inexact data is presented for minimizing an unconstrained nonsmooth convex function in [13].

In this paper, the objective function in (1) is nonsmooth, but it has the structure which has the connection with $\mathcal{V}\mathcal{U}$-space decomposition. Based on the $\mathcal{V}\mathcal{U}$-theory, a superlinear convergent $\mathcal{V}\mathcal{U}$-algorithm frame is designed to solve the SAA problem. The rest of the paper is organized as follows. In the next section, the SCMP is transformed to the nonsmooth problem and the proof of the approximation solution set converges to the true solution set in the sense that Hausdorff distance is obtained. In Section 3, the $\mathcal{V}\mathcal{U}$-theory of the SAA problem is given. In the final section, the $\mathcal{V}\mathcal{U}$-decomposition algorithm frame of the SAA problem is designed.

2. Convergence Analysis of SAA Problem

In this section, we discuss the convergence of (3) to (1) as N increases. Specifically, we investigate the fact that the solution of the SAA problem (3) converges to its true counterpart as $N \to \infty$. Firstly, we make the basic assumptions for SAA method. In the following, we give the basic assumptions for SAA method.

Assumption 1. (a) Letting X be a set, for $i = 1, \ldots, n$, the limits

$$M_{F_i}(t) := \lim_{N \to \infty} M_{F_i}^{N}(t) \tag{5}$$

exist for every $x \in X$.

(b) For every $s \in X$, the moment-generating function $M_s(t)$ is finite-valued for all t in a neighborhood of zero.

(c) There exists a measurable function $\kappa : \Omega \to R_+$ such that

$$\left| g(s', \xi) - g(s, \xi) \right| \le \kappa(\xi) \left\| s' - s \right\| \tag{6}$$

for all $\xi \in \Omega$ and all $s', s \in X$.

(d) The moment-generating function $M_\kappa(t) = E[e^{t\kappa(\xi)}]$ of $\kappa(\xi)$ is finite-valued for all t in a neighborhood of zero, where $M_s(t) = E[e^{t(g(s,\xi)-G(s))}]$ is the moment-generating function of the random variable $g(s,\xi) - G(s)$.

Theorem 2. *Let S^* and S^N denote the solution sets of (1) and (3). Assuming that both S^* and S^N are nonempty, then, for any $\varepsilon > 0$, one has $D(S^N, S^*) < \varepsilon$, where $D(S^N, S^*) = \sup_{x \in S^N} d(x, S^*)$.*

Proof. For any points $\tilde{x} \in S^N$ and $x \in R^n$, we have

$$\begin{aligned} \hat{f}^N(\tilde{x}) &= \max\left\{\hat{f}_i^N(\tilde{x}),\ i = 0, \ldots, m\right\} \\ &= \max\left\{\frac{1}{N}\sum_{j=1}^{N} f_i\left(\tilde{x}, \xi^j\right),\ i = 0, \ldots, m\right\} \\ &\le \max\left\{\frac{1}{N}\sum_{j=1}^{N} f_i\left(x, \xi^j\right),\ i = 0, \ldots, m\right\}. \end{aligned} \tag{7}$$

From Assumption 1, we know that, for any $\varepsilon > 0$, there exist M; if $N > M$, $i = 0, \ldots, m$, then

$$\left|\frac{1}{N}\sum_{j=1}^{N} f_i\left(x, \xi^j\right) - E\left[f_i\left(x, \xi^j\right)\right]\right| < \varepsilon. \tag{8}$$

By letting $N > M$, we obtain

$$\begin{aligned} \hat{f}^N(\tilde{x}) &\le \max\left\{\frac{1}{N}\sum_{j=1}^{N} f_i\left(x, \xi^j\right),\ i = 0, \ldots, m\right\} \\ &\le \max\left\{E\left[f_i\left(x, \xi^j\right)\right] + \varepsilon,\ i = 0, \ldots, m\right\} \\ &= f(x) + \varepsilon. \end{aligned} \tag{9}$$

This shows that $d(\tilde{x}^N, S^*) < \varepsilon$, which implies $D(\tilde{x}^N, s^*) < \varepsilon$. □

We now move on to discuss the exponential rate of convergence of SAA problem (3) to the true problem (1) as sample increases.

Theorem 3. *Let x^N be a solution to the SAA problem (3) and S^* is the solution set of the true problem (1). Suppose Assumption 1 holds. Then, for every $\varepsilon > 0$, there exist positive constants $c(\varepsilon)$ and $d(\varepsilon)$, such that*

$$\text{Prob}\left\{d\left(x^N, S^*\right) \ge \varepsilon\right\} \le c(\varepsilon)\exp^{-Nd(\varepsilon)} \tag{10}$$

for N sufficiently large.

Proof. Let $\varepsilon > 0$ be any small positive number. By Theorem 2 and

$$\left|\frac{1}{N}\sum_{j=1}^{N} f_i\left(x, \xi^j\right) - E\left[f_i\left(x, \xi^j\right)\right]\right| < \varepsilon, \tag{11}$$

we have $d(x^N, S) < \varepsilon$. Therefore, by Assumption 1, we have

$$\begin{aligned}&\text{Prob}\left\{d\left(x^N, S^*\right) \geq \varepsilon\right\}\\ &\quad\leq \text{Prob}\left\{\left|\frac{1}{N}\sum_{j=1}^{N} f_i\left(x, \xi^j\right) - E\left[f_i\left(x, \xi^j\right)\right]\right| \geq \delta\right\} \\ &\quad\leq c(\varepsilon)\exp^{-Nd(\varepsilon)}.\end{aligned} \tag{12}$$

The proof is complete. □

3. The $\mathcal{V}\mathcal{U}$-Theory of the SAA Problem

In the following sections, we give the $\mathcal{V}\mathcal{U}$-theory, $\mathcal{V}\mathcal{U}$-decomposition algorithm frame, and convergence analysis of the SAA problem.

The subdifferential of $\hat{f}^N$ at a point $x \in R^n$ can be computed in terms of the gradients of the function that are active at x. More precisely,

$$\begin{aligned}&\partial \hat{f}^N(x)\\ &\quad= \text{Conv}\left\{g \in R^n \mid g = \sum_{i\in I(x)} \frac{\alpha_i}{N}\sum_{j=1}^{N}\nabla f_i\left(x, \xi^j\right) : \right. \\ &\qquad\qquad \left. \alpha = (\alpha_i)_{i\in I(x)} \in \Delta_{|I(x)|}\right\},\end{aligned} \tag{13}$$

where

$$I(x) = \left\{i \in I \mid \hat{f}^N(x) = \hat{f}_i^N(x)\right\} \tag{14}$$

is the set of active indices at x, and

$$\Delta_s = \left\{\alpha \in R^s \mid \alpha_i \geq 0, \sum_{i=1}^{s}\alpha_i = 1\right\}. \tag{15}$$

Let $\bar{x} \in R^n$ be a solution of (3). By continuity of the structure functions, there exists a ball $B_\varepsilon(\bar{x}) \subseteq R^n$ such that

$$\forall x \in B_\varepsilon(\bar{x}), \quad I(x) \subseteq I(\bar{x}). \tag{16}$$

For convenience, we assume that the cardinality of $I(\bar{x})$ is $m_1 + 1$ $(m_1 \leq m)$ and reorder the structure functions, so that $I(\bar{x}) = \{0, \dots, m_1\}$. From now on, we consider that

$$\forall x \in B_\varepsilon(\bar{x}), \quad \hat{f}^N(x) = \hat{f}_i^N(x), \quad i \in I(\bar{x}). \tag{17}$$

The following assumption will be used in the rest of this paper.

Assumption 4. The set

$$\left\{\nabla \hat{f}_i^N(x) - \nabla \hat{f}_0^N(x)\right\}_{0\neq i\in I(\bar{x})} \tag{18}$$

is linearly independent.

Theorem 5. *Suppose Assumption 4 holds. Then R^n can be decomposition at $\bar{x}$: $R^n = \mathcal{U} \oplus \mathcal{V}$, where*

$$\begin{aligned}\mathcal{V} &= \text{lin}\left\{\nabla \hat{f}_i^N(x) - \nabla \hat{f}_0^N(x)\right\}_{0\neq i\in I(\bar{x})},\\ \mathcal{U} &= \left\{d \in R^n \mid \left\langle d, \left\{\nabla \hat{f}_i^N(x) - \nabla \hat{f}_0^N(x)\right\}_{0\neq i\in I(\bar{x})}\right\rangle\right\} = 0.\end{aligned} \tag{19}$$

Proof. The proof can be directly obtained by using Assumption 4 and the definition of the spaces $\mathcal{V}$ and $\mathcal{U}$.

Given a subgradient $\bar{g} \in \partial \hat{f}^N$ with $\mathcal{V}$-component $\bar{g}_v = \bar{V}^T\bar{g}$, the $\mathcal{U}$-Lagrangian of $\hat{f}^N$, depending on $\bar{g}_v$, is defined by

$$\begin{aligned}R^{\dim \mathcal{U}} \ni u \longmapsto &L_u(u; \bar{g}_v)\\ &:= \min_{v\in R^{\dim \mathcal{V}}}\left\{\hat{f}^N\left(\bar{x} + \bar{U}u + \bar{V}v\right) - \langle \bar{g}_v, v\rangle_v\right\}.\end{aligned} \tag{20}$$

The associated set of $\mathcal{V}$-space minimizers is defined by

$$\begin{aligned}&W(u; \bar{g}_v)\\ &\quad:= \left\{v : L_u(u; \bar{g}_v) = \hat{f}^N\left(\bar{x} + \bar{U}u + \bar{V}v\right) - \langle \bar{g}_v, v\rangle_v\right\}.\end{aligned} \tag{21}$$

□

Theorem 6. *Suppose Assumption 4 holds. Let $\chi(u) = \bar{x} + u \oplus v(u)$ be a trajectory leading to $\bar{x}$ and let $\bar{H} := \nabla^2 L_u(0, 0)$. Then for all u sufficiently small the following hold:*

(i) *the nonlinear system, with variable v and the parameter u,*

$$\hat{f}_i^N\left(\bar{x} + \bar{U}u + \bar{V}v\right) - \hat{f}_0^N\left(\bar{x} + \bar{U}u + \bar{V}v\right) = 0, \quad 0 \neq i \in I(\bar{x}) \tag{22}$$

has a unique solution $v = v(u)$ and $v : R^{\dim \mathcal{U}} \to R^{\dim \mathcal{V}}$ is a C^2 function;

(ii) *$\chi(u)$ is a C^2-function with $J\chi(u) = \bar{U} + \bar{V}Jv(u)$;*

(iii) *$L_u(u; 0) = \hat{f}_i^N(\bar{x} + u \oplus v(u)) = \hat{f}^N(\bar{x} + u \oplus v(u)) = \hat{f}_i^N(\bar{x}) + (1/2)u^T\bar{H}u + o(|u|^2)$;*

(iv) *$\nabla L_u(u; 0) = \bar{H}u + o(|u|)$;*

(v) *$\hat{f}^N(\chi(u)) = \hat{f}_i^N(\chi(u))$, $i \in I(\bar{x})$.*

Proof. Item (i) follows from the assumption that f_i are C^2 and applying a Second-Order Implicit Function Theorem (see [14], Theorem 2.1). Since $v(u)$ is C^2, $\chi(u)$ is C^2 and the Jacobians $Jv(u)$ exist and are continuous. Differentiating the primal track with respect to u, we obtain the expression of $J\chi(u)$ and item (ii) follows.

(iii) By the definition of $L_u(u; \bar{g}_v)$ and $W(u; \bar{g}_v)$, we have

$$L_u(u; 0) = \hat{f}_i^N(\bar{x} + u \oplus v(u)) = \hat{f}^N(\bar{x} + u \oplus v(u)). \tag{23}$$

According to the second-order expansion of L_u, we obtain

$$L_u(u;0) = L_u(0;0) + \langle \nabla L_u(0;0), u\rangle + \frac{1}{2}u^T\nabla^2 L_u(0;0)u + o(|u|). \tag{24}$$

Since $L_u(0;0) = \hat{f}_i^N(\bar{x})$, $i \in I(\bar{x})$, $\nabla L_u(0;0) = 0$, and $\overline{H} = \nabla^2 L_u(0;0)$,

$$L_u(u;0) = \hat{f}_i^N(\bar{x}) + \frac{1}{2}u^T\overline{H}u + o(|u|^2). \tag{25}$$

Similar to (iii), we get (iv):

$$\nabla L_u(u;0) = \nabla L_u(0;0) + \langle u^T\nabla^2 L_u(0;0), u\rangle + o(|u|^2) = \overline{H} + o(|u|). \tag{26}$$

The conclusion of (v) can be obtained in terms of (i) and the definition of $\chi(u)$. □

4. Algorithm and Convergence Analysis

Supposing $0 \in \partial\hat{f}^N(\bar{x})$, we give an algorithm frame which can solve (3). This algorithm makes a step in the $\mathcal{V}$-subspace, followed by a $\mathcal{U}$-Newton step in order to obtain superlinear convergence rate.

Algorithm 7 (algorithm frame).

Step 0. Initialization: given $\varepsilon > 0$, choose a starting point $x^{(0)}$ close to $\bar{x}$ enough and a subgradient $\tilde{g}^{(0)} \in \partial\hat{f}^N(x^{(0)})$ and set $k = 0$.

Step 1. Stop if

$$\|\tilde{g}^{(k)}\| \le \varepsilon. \tag{27}$$

Step 2. Find the active index set $I(\bar{x})$.

Step 3. Construct $\mathcal{V}\mathcal{U}$-decomposition at $\bar{x}$; that is, $R^n = \mathcal{V} \oplus \mathcal{U}$. Compute

$$\nabla^2 L_u(0;0) = \overline{U}^T M(0)\overline{U}, \tag{28}$$

where

$$M(0) = \sum_{i\in I(\bar{x})} \bar{\alpha}_i \nabla^2 \hat{f}_i^N(\bar{x}). \tag{29}$$

Step 4. Perform $\mathcal{V}$-step. Compute $\delta_{\mathcal{V}}^{(k)}$ which denotes $v(u)$ in (22) and set $x^{(k)} = x^{(k)} + 0 \oplus \delta_{\mathcal{V}}^{(k)}$.

Step 5. Perform $\mathcal{U}$-step. Compute $\delta_{\mathcal{U}}^{(k)}$ from the system

$$\overline{U}^T M(0)\overline{U}\delta_{\mathcal{U}} + \overline{U}^T\tilde{g}^{(k)} = 0, \tag{30}$$

where

$$\sum_{i\in I(\bar{x})} \alpha_i(u)\nabla\hat{f}_i^{N_i}(x^{(k)}) = \tilde{g}^{(k)} \in \partial\hat{f}^N(x^{(k)}) \tag{31}$$

is such that $\overline{V}^T\tilde{g}^{(k)} = 0$. Compute $x^{(k+1)} = \tilde{x}^{(k)} + \delta_{\mathcal{U}}^{(k)} \oplus 0 = x^{(k)} + \delta_{\mathcal{U}}^{(k)} \oplus \delta_{\mathcal{V}}^{(k)}$.

Step 6. Update: set $k = k + 1$ and return to Step 1.

Theorem 8. *Suppose the starting point $x^{(0)}$ is close to $\bar{x}$ enough and $0 \in \mathrm{ri}\partial\hat{f}(\bar{x})$, $\nabla^2 L_u(0;0) \succ 0$. Then the iteration points $\{x^{(k)}\}_{k=1}^{\infty}$ generated by Algorithm 7 converge and satisfy*

$$\|x^{(k+1)} - \bar{x}\| = o(\|x^{(k)} - \bar{x}\|). \tag{32}$$

Proof. Let $u^{(k)} = (x^{(k)} - \bar{x})_{\mathcal{U}}$ and $v^{(k)} = (x^{(k)} - \bar{x})_{\mathcal{V}} + \delta_{\mathcal{V}}^{(k)}$. It follows from Theorem 6(i) that

$$\|(x^{(k+1)} - \bar{x})_{\mathcal{V}}\| = \|(x^{(k)} - \bar{x})_{\mathcal{V}}\| = o\|(x^{(k)} - \bar{x})_{\mathcal{U}}\| = o\|(x^{(k)} - \bar{x})\|. \tag{33}$$

Since $\nabla^2 L_u(0;0)$ exists and $\nabla L_u(0;0) = 0$, we have from the definition of $\mathcal{U}$-Hessian matrix that

$$\nabla L_u(u^{(k)};0) = \overline{U}^T\tilde{g}^{(k)} = 0 + \nabla^2 L_u(0;0)u^{(k)} + o(\|u^{(k)}\|_{\mathcal{U}}). \tag{34}$$

By virtue of (30), we have $\nabla^2 L_u(0;0)(u^{(k)} + \delta_{\mathcal{U}}^{(k)}) = o(\|u^{(k)}\|_{\mathcal{U}})$. It follows from the hypothesis $\nabla^2 L_u(0;0) \succ 0$ that $\nabla^2 L_u(0;0)$ is invertible and hence $\|u^{(k)} + \delta_{\mathcal{U}}^{(k)}\| = o(\|u^{(k)}\|_{\mathcal{U}})$. In consequence, one has

$$\begin{aligned}(x^{(k+1)} - \bar{x})_{\mathcal{U}} &= (x^{(k+1)} - x^{(k)})_{\mathcal{U}} + (x^{(k)} - x^{(k)})_{\mathcal{U}} + (x^{(k)} - \bar{x})_{\mathcal{U}} \\ &= u^{(k)} + \delta_{\mathcal{U}}^{(k)} = o(\|u^{(k)}\|_{\mathcal{U}}) \\ &= o(\|x^{(k)} - \bar{x}\|).\end{aligned} \tag{35}$$

The proof is completed by combining (33) and (35). □

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The research is supported by the National Natural Science Foundation of China under Project nos. 11301347, 11171138, and 11171049 and General Project of the Education Department of Liaoning Province no. L201242.

References

[1] B. K. Pagnoncelli, S. Ahmed, and A. Shapiro, "Sample average approximation method for chance constrained programming: theory and applications," *Journal of Optimization Theory and Applications*, vol. 142, no. 2, pp. 399–416, 2009.

[2] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lecture on Stochastic Programming: Modelling and Theory*, SIAM, Philadelphia, Pa, USA, 2009.

[3] H. Xu, "Sample average approximation methods for a class of stochastic variational inequality problems," *Asia-Pacific Journal of Operational Research*, vol. 27, no. 1, pp. 103–119, 2010.

[4] Y. Liu, H. Xu, and J. J. Ye, "Penalized sample average approximation methods for stochastic mathematical programs with complementarity constraints," *Mathematics of Operations Research*, vol. 36, no. 4, pp. 670–694, 2011.

[5] S. Chen, L.-P. Pang, F.-F. Guo, and Z.-Q. Xia, "Stochastic methods based on Newton method to the stochastic variational inequality problem with constraint conditions," *Mathematical and Computer Modelling*, vol. 55, no. 3-4, pp. 779–784, 2012.

[6] C. Lemarechal, F. Oustry, and C. Sagastizabal, "The U-Lagrangian of a convex function," *Transactions of the American Mathematical Society*, vol. 352, no. 2, pp. 711–729, 2000.

[7] R. Mifflin and C. Sagastizábal, "VU-decomposition derivatives for convex max-functions," in *Ill-Posed Variational Problems and Regularization Techniques*, M. Théra and R. Tichatschke, Eds., vol. 477 of *Lecture Notes in Economics and Mathematical Systems*, pp. 167–186, Springer, Berlin, Germany, 1999.

[8] C. Lemaréchal and C. Sagastizábel, "More than first-order developments of convex functions: primal-dual relations," *Journal of Convex Analysis*, vol. 3, no. 2, pp. 255–268, 1996.

[9] R. Mifflin and C. Sagastizabal, "On VU-theory for functions with primal-dual gradient strcture," *SIAM Journal on Optimization*, vol. 11, no. 2, pp. 547–571, 2000.

[10] R. Mifflin and C. Sagastizábal, "Functions with primal-dual gradient structure and U-Hessians," in *Nonlinear Optimization and Related Topics*, G. Pillo and F. Giannessi, Eds., vol. 36 of *Applied Optimization*, pp. 219–233, Kluwer Academic, 2000.

[11] R. Mifflin and C. Sagastizábal, "Primal-dual gradient structured functions: second-order results; links to EPI-derivatives and partly smooth functions," *SIAM Journal on Optimization*, vol. 13, no. 4, pp. 1174–1194, 2003.

[12] Y. Lu, L. P. Pang, F. F. Guo, and Z. Q. Xia, "A superlinear space decomposition algorithm for constrained nonsmooth convex program," *Journal of Computational and Applied Mathematics*, vol. 234, no. 1, pp. 224–232, 2010.

[13] Y. Lu, L.-P. Pang, J. Shen, and X.-J. Liang, "A decomposition algorithm for convex nondifferentiable minimization with errors," *Journal of Applied Mathematics*, vol. 2012, Article ID 215160, 15 pages, 2012.

[14] S. Lang, *Real and Functional Analysis*, vol. 142 of *Graduate Texts in Mathematics*, Springer, New York, NY, USA, 3rd edition, 1993.