Stochastic Optimization with Inequality Constraints Using Simultaneous Perturbations and Penalty Functions

I-JENG WANG*† and JAMES C. SPALL†

†The Johns Hopkins University Applied Physics Laboratory
11100 Johns Hopkins Road, Laurel, MD 20723-6099, U.S.A.

We present a stochastic approximation algorithm based on penalty function method and a simultaneous perturbation gradient estimate for solving stochastic optimization problems with general inequality constraints. We present a general convergence result that applies to a class of penalty functions including the quadratic penalty function, the augmented Lagrangian, and the absolute penalty function. We also establish an asymptotic normality result for the algorithm with smooth penalty functions under minor assumptions. Numerical results are given to compare the performance of the proposed algorithm with different penalty functions.

1 Introduction

In this paper, we consider a constrained stochastic optimization problem for which only noisy measurements of the cost function are available. More specifically, we are aimed to solve the following optimization problem:

$$\min_{\theta \in G} L(\theta), \tag{1}$$

where $L: \mathcal{R}^d \to \mathcal{R}$ is a real-valued cost function, $\theta \in \mathcal{R}^d$ is the parameter vector, and $G \subset \mathcal{R}^d$ is the constraint set. We also assume that the gradient of $L(\cdot)$ exists and is denoted by $g(\cdot)$. We assume that there exists a unique solution θ^* for the constrained optimization problem defined by (1). We consider the situation where no explicit closed-form expression of the function L is available (or is very complicated even if available), and the only information are noisy measurements of L at specified values of the parameter vector θ. This scenario arises naturally for simulation-based optimization where the cost function L is defined as the expected value of a random cost associated with the stochastic simulation of a complex system. We also assume that significant costs (in term of time and/or computational costs) are involved in obtaining each measurement (or sample) of $L(\theta)$. These constraint prevent us from estimating the gradient (or Hessian) of $L(\cdot)$ accurately, hence prohibit the application of effective nonlinear programming techniques for inequality constraint, for example, the sequential quadratic programming methods (see; for example, section 4.3 of (1)). Throughout the paper we use θ_n to denote the nth estimate of the solution θ^*.

Several results have been presented for constrained optimization in the stochastic domain. In the area of stochastic approximation (SA), most of the available results are based on the simple idea of projecting the estimate θ_n back to its nearest point in G whenever θ_n lies outside the constraint set G. These projection-based SA algorithms are typically of the following form:

$$\theta_{n+1} = \pi_G[\theta_n - a_n \hat{g}_n(\theta_n)], \tag{2}$$

---

*Corresponding author. Email: i-jeng.wang@jhuapl.edu

where $\pi_G\colon \mathcal{R}^d \to G$ is the set projection operator, and $\hat{g}_n(\theta_n)$ is an estimate of the gradient $g(\theta_n)$; see, for example (2, 3, 5, 6). The main difficulty for this projection approach lies in the implementation (calculation) of the projection operator π_G. Except for simple constraints like interval or linear constraints, calculation of $\pi_G(\theta)$ for an arbitrary vector θ is a formidable task.

Other techniques for dealing with constraints have also been considered: Hiriart-Urruty (7) and Pflug (8) present and analyze a SA algorithm based on the penalty function method for stochastic optimization of a convex function with convex inequality constraints; Kushner and Clark (3) present several SA algorithms based on the Lagrange multiplier method, the penalty function method, and a combination of both. Most of these techniques rely on the Kiefer-Wofolwitz (KW) (4) type of gradient estimate when the gradient of the cost function is not readily available. Furthermore, the convergence of these SA algorithms based on "non-projection" techniques generally requires complicated assumptions on the cost function L and the constraint set G. In this paper, we present and study the convergence of a class of algorithms based on the penalty function methods and the simultaneous perturbation (SP) gradient estimate (9). The advantage of the SP gradient estimate over the KW-type estimate for unconstrained optimization has been demonstrated with the simultaneous perturbation stochastic approximation (SPSA) algorithms. And whenever possible, we present sufficient conditions (as remarks) that can be more easily verified than the much weaker conditions used in our convergence proofs.

We focus on general explicit inequality constraints where G is defined by

$$G \triangleq \{\theta \in \mathcal{R}^d\colon q_j(\theta) \leq 0, j = 1, \ldots, s\}, \tag{3}$$

where $q_j\colon \mathcal{R}^d \to \mathcal{R}$ are continuously differentiable real-valued functions. We assume that the analytical expression of the function q_j is available. We extend the result presented in (10) to incorporate a larger classes of penalty functions based on the augmented Lagragian method. We also establish the asymptotic normality for the proposed algorithm. Simulation results are presented to illustrated the performance of the technique for stochastic optimization.

2   Constrained SPSA Algorithms

2.1   *Penalty Functions*

The basic idea of the penalty-function approach is to convert the originally constrained optimization problem (1) into an unconstrained one defined by

$$\min_\theta L_r(\theta) \triangleq L(\theta) + rP(\theta), \tag{4}$$

where $P\colon \mathcal{R}^d \to \mathcal{R}$ is the penalty function and r is a positive real number normally referred to as the *penalty parameter*. The penalty functions are defined such that P is an increasing function of the constraint functions q_j; $P > 0$ if and only if $q_j > 0$ for any j; $P \to \infty$ as $q_j \to \infty$; and $P \to -l$ ($l \geq 0$) as $q_j \to -\infty$. In this paper, we consider a penalty function method based on the augmented Lagrangian function defined by

$$L_r(\theta, \lambda) = L(\theta) + \frac{1}{2r} \sum_{j=1}^{s} \left\{ [\max\{0, \lambda_j + rq_j(\theta)\}]^2 - \lambda_j^2 \right\}, \tag{5}$$

where $\lambda \in \mathcal{R}^s$ can be viewed as an estimate of the Lagrange multiplier vector. The associated penalty function is

$$P(\theta) = \frac{1}{2r^2} \sum_{j=1}^{s} \left\{ [\max\{0, \lambda_j + rq_j(\theta)\}]^2 - \lambda_j^2 \right\}. \tag{6}$$

Let $\{r_n\}$ be a positive and strictly increasing sequence with $r_n \to \infty$ and $\{\lambda_n\}$ be a bounded nonnegative sequence in $\mathcal{R}^s$. It can be shown (see; for example, section 4.2 of (1)) that the minimum of the sequence functions $\{L_n\}$, defined by

$$L_n(\theta) \triangleq L_{r_n}(\theta, \lambda_n),$$

converges to the solution of the original constrained problem (1). Since the penalized cost function (or the augmented Lagrangian) (5) is a differentiable function of θ, we can apply the standard stochastic approximation technique with the SP gradient estimate for L to minimize $\{L_n(\cdot)\}$. In other words, the original problem can be solved with an algorithm of the following form:

$$\theta_{n+1} = \theta_n - a_n \hat{\nabla} L_n(\theta_n)$$
$$= \theta_n - a_n \hat{g}_n - a_n r_n \nabla P(\theta_n),$$

where $\hat{g}_n$ is the SP estimate of the gradient $g(\cdot)$ at θ_n that we shall specify later. Note that since we assume the constraints are explicitly given, the gradient of the penalty function $P(\cdot)$ is directly used in the algorithm.

Note that when $\lambda_n = 0$, the penalty function defined by (6) reduces to the standard quadratic penalty function discussed in (10)

$$L_r(\theta, 0) = L(\theta) + r \sum_{j=1}^{s} [\max\{0, q_j(\theta)\}]^2 .$$

Even though the convergence of the proposed algorithm only requires $\{\lambda_n\}$ be bounded (hence we can set $\lambda_n = 0$), we can significantly improve the performance of the algorithm with appropriate choice of the sequence based on concepts from Lagrange multiplier theory. Moreover, it has been shown (1) that, with the standard quadratic penalty function, the penalized cost function $L_n = L + r_n P$ can become ill-conditioned as r_n increases (that is, the condition number of the Hessian matrix of L_n at θ_n^* diverges to ∞ with r_n). The use of the general penalty function defined in (6) can prevent this difficulty if $\{\lambda_n\}$ is chosen so that it is close to the true Lagrange multipliers. In Section 4, we will present an iterative method based on the method of multipliers (see; for example, (11)) to update λ_n and compare its performance with the standard quadratic penalty function.

2.2 A SPSA Algorithms for Inequality Constraints

In this section, we present the specific form of the algorithm for solving the constrained stochastic optimization problem. The algorithm we consider is defined by

$$\theta_{n+1} = \theta_n - a_n \hat{g}_n(\theta_n) - a_n r_n \nabla P(\theta_n), \tag{7}$$

where $\hat{g}_n(\theta_n)$ is an estimate of the gradient of L, $g(\cdot)$, at θ_n, $\{r_n\}$ is an increasing sequence of positive scalar with $\lim_{n \to \infty} r_n = \infty$, $\nabla P(\theta)$ is the gradient of $P(\theta)$ at θ, and $\{a_n\}$ is a positive scalar sequence satisfying $a_n \to 0$ and $\sum_{n=1}^{\infty} a_n = \infty$. The gradient estimate $\hat{g}_n$ is obtained from two noisy measurements of the cost function L by

$$\frac{(L(\theta_n + c_n \Delta_n) + \epsilon_n^+) - (L(\theta_n - c_n \Delta_n) + \epsilon_n^-)}{2c_n} \frac{1}{\Delta_n}, \tag{8}$$

where $\Delta_n \in \mathcal{R}^d$ is a random perturbation vector, $c_n \to 0$ is a positive sequence, ϵ_n^+ and ϵ_n^- are noise in the measurements. and $1/\Delta_n$ denotes the vector $[1/\Delta_n^1, \ldots, 1/\Delta_n^d]$. For analysis, we rewrite the algorithm (7)

into

$$\theta_{n+1} = \theta_n - a_n g(\theta_n) - a_n r_n \nabla P(\theta_n) + a_n d_n - a_n \frac{\epsilon_n}{2c_n \Delta_n}, \tag{9}$$

where d_n and ϵ_n are defined by

$$d_n \triangleq g(\theta_n) - \frac{L(\theta_n + c_n \Delta_n) - L(\theta_n - c_n \Delta_n)}{2c_n \Delta_n},$$

$$\epsilon_n \triangleq \epsilon_n^+ - \epsilon_n^-,$$

respectively.

We establish the convergence of the algorithm (7) and the associated asymptotic normality under appropriate assumptions in the next section.

3   Convergence and Asymptotic Normality

3.1   Convergence Theorem

To establish convergence of the algorithm (7), we need to study the asymptotic behavior of an SA algorithm with a "time-varying" regression function. In other words, we need to consider the convergence of an SA algorithm of the following form:

$$\theta_{n+1} = \theta_n - a_n f_n(\theta_n) + a_n d_n + a_n e_n, \tag{10}$$

where $\{f_n(\cdot)\}$ is a sequence of functions. We state here without proof a version of the convergence theorem given by Spall and Cristion in (13) for an algorithm in the generic form (10).

THEOREM 3.1 *Assume the following conditions hold:*

(A.1) *For each n large enough ($\geq N$ for some $N \in \mathcal{N}$), there exists a unique θ_n^* such that $f_n(\theta_n^*) = 0$. Furthermore, $\lim_{n \to \infty} \theta_n^* = \theta^*$.*

(A.2) *$d_n \to 0$, and $\sum_{k=1}^n a_k e_k$ converges.*

(A.3) *For some $N < \infty$, any $\rho > 0$ and for each $n \geq N$, if $\|\theta - \theta^*\| \geq \rho$, then there exists a $\delta_n(\rho) > 0$ such that $(\theta - \theta^*)^T f_n(\theta) \geq \delta_n(\rho)\|\theta - \theta^*\|$ where $\delta_n(\rho)$ satisfies $\sum_{n=1}^\infty a_n \delta_n(\rho) = \infty$ and $d_n \delta_n(\rho)^{-1} \to 0$.*

(A.4) *For each $i = 1, 2, \ldots, d$, and any $\rho > 0$, if $|\theta_{ni} - (\theta^*)_i| > \rho$ eventually, then either $f_{ni}(\theta_n) \geq 0$ eventually or $f_{ni}(\theta_n) < 0$ eventually, where we use $(\theta^*)_i$ to denote the i-th element of θ^*.*

(A.5) *For any $\tau > 0$ and nonempty $S \subset \{1, 2, \ldots, d\}$, there exists a $\rho'(\tau, S) > \tau$ such that for all $\theta \in \{\theta \in \mathcal{R}^d \colon |(\theta - \theta^*)_i| < \tau$ when $i \notin S, |(\theta - \theta^*)_i| \geq \rho'(\tau, S)$ when $i \in S.\}$,*

$$\limsup_{n \to \infty} \left| \frac{\sum_{i \notin S}(\theta - \theta^*)_i f_{ni}(\theta)}{\sum_{i \in S}(\theta - \theta^*)_i f_{ni}(\theta)} \right| < 1.$$

Then the sequence $\{\theta_n\}$ defined by the algorithm (10) converges to θ^*.

Based on Theorem 3.1, we give a convergence result for algorithm (7) by substituting $\nabla L_n(\theta_n) = g(\theta_n) + r_n \nabla P(\theta_n)$, d_n, and $\frac{\epsilon_n}{2c_n \Delta_n}$ into $f_n(\theta_n)$, d_n, and e_n in (10), respectively. We need the following assumptions:

(C.1) There exists $K_1 \in \mathcal{N}$ such that for all $n \geq K_1$, we have a unique $\theta_n^* \in \mathcal{R}^d$ with $\nabla L_n(\theta_n^*) = 0$.

(C.2) $\{\Delta_{ni}\}$ are i.i.d. and symmetrically distributed about 0, with $|\Delta_{ni}| \leq \alpha_0$ a.s. and $E|\Delta_{ni}^{-1}| \leq \alpha_1$.

(C.3) $\sum_{k=1}^n \frac{a_k \epsilon_k}{2c_k \Delta_k}$ converges almost surely.

(C.4) If $\|\theta - \theta^*\| \geq \rho$, then there exists a $\delta(\rho) > 0$ such that
   (i) if $\theta \in G$, $(\theta - \theta^*)^T g(\theta) \geq \delta(\rho)\|\theta - \theta^*\| > 0$.

    (ii) if $\theta \notin G$, at least one of the following two conditions hold
- $(\theta - \theta^*)^T g(\theta) \geq \delta(\rho)\|\theta - \theta^*\|$ and $(\theta - \theta^*)^T \nabla P(\theta) \geq 0$.
- $(\theta - \theta^*)^T g(\theta) \geq -M$ for some constant $M > 0$ and $(\theta - \theta^*)^T \nabla P(\theta) \geq \delta(\rho)\|\theta - \theta^*\| > 0$

(C.5) $a_n r_n \to 0$, $g(\cdot)$ and $\nabla P(\cdot)$ are Lipschitz. (See comments below)

(C.6) $\nabla L_n(\cdot)$ satisfies condition (A5)

THEOREM 3.2 *Suppose that assumptions (C.1–C.6) hold. Then the sequence $\{\theta_n\}$ defined by (7) converges to θ^* almost surely.*

Proof We only need to verify the conditions (A.1–A.5) in Theorem 3.1 to show the desired result:

- Condition (A.1) basically requires the stationary point of the sequence $\{\nabla L_n(\cdot)\}$ converges to θ^*. Assumption (C.1) together with existing results on penalty function methods establishes this desired convergence.
- From the results in (9, 14) and assumptions (C.2–C.3), we can show that condition (A.2) hold.
- Since $r_n \to \infty$, we have condition (A.3) hold from assumption (C.4).
- From (9), assumption (C.1) and (C.5) , we have

$$|(\theta_{n+1} - \theta_n)_i| < |(\theta_n - \theta^*)_i|$$

for large n if $|\theta_{ni} - (\theta^*)_i| > \rho$. Hence for large n, the sequence $\{\theta_{ni}\}$ does not "jump" over the interval between $(\theta^*)_i$ and θ_{ni}. Therefore if $|\theta_{ni} - (\theta^*)_i| > \rho$ eventually, then the sequence $\{f_{ni}(\theta_n)\}$ does not change sign eventually. That is, condition (A.4) holds.
- Assumption (A.5) holds directly from (C.6).

$\square$

    Theorem 3.2 given above is general in the sense that it does not specify the exact type of penalty function $P(\cdot)$ to adopt. In particular, assumption (C.4) seems difficult to satisfy. In fact, assumption (C.4) is fairly weak and does address the limitation of the penalty function based gradient descent algorithm. For example, suppose that a constraint function $q_k(\cdot)$ has a local minimum at θ' with $q_k(\theta') > 0$. Then for every θ with $q_j(\theta) \leq 0, j \neq k$, we have $(\theta - \theta')^T \nabla P(\theta) > 0$ whenever θ is close enough to θ'. As r_n gets larger, the term $\nabla P(\theta)$ would dominate the behavior of the algorithm and result in a possible convergence to θ', a wrong solution. We also like to point out that assumption (C.4) is satisfied if cost function L and constraint functions $q_j, j = 1, \ldots, s$ are convex and satisfy the Slater condition, that is, the minimum cost function value $L(\theta^*)$ is finite and there exists a $\theta \in \mathcal{R}^d$ such that $q_j(\theta) < 0$ for all j (this is the case studied in (8)). Assumption (C.6) ensures that for n sufficiently large each element of $g(\theta) + r_n \nabla P(\theta)$ make a non-negligible contribution to products of the form $(\theta - \theta^*)^T (g(\theta) + r_n \nabla P(\theta))$ when $(\theta - \theta^*)_i \neq 0$. A sufficient condition for (C.6) is that for each i, $g_i(\theta) + r_n \nabla_i P(\theta)$ be uniformly bounded both away from 0 and ∞ when $\|(\theta - \theta^*)_i\| \geq \rho > 0$ for all i.

    Theorem 3.2 in the stated form does require that the penalty function P be differentiable. However, it is possible to extend the stated results to the case where P is Lipschitz but not differentiable at a set of points with zero Lebesgue measure, for example, the absolute value penalty function

$$P(\theta) = \max_{j=1,\ldots,s} \{\max \{0, q_j(\theta)\}\}.$$

In the case where the density function of measurement noise (ϵ_n^+ and ϵ_n^- in (8)) exists and has infinite support, we can take advantage of the fact that iterations of the algorithm visit any zero-measure set with zero probability. Assuming that the set $D \triangleq \{\theta \in \mathcal{R}^d : \nabla P(\theta)$ does not exist$\}$ has Lebesgue measure 0 and the random perturbation Δ_n follows Bernoulli distribution ($P(\Delta_n^i = 0) = P(\Delta_n^i = 1) = \frac{1}{2}$), we can construct a simple proof to show that

$$P\{\theta_n \in D \quad \text{infinitely often}\} = 0$$

if $P\{\theta_0 \in D\} = 0$. Therefore, the convergence result in Theorem 3.2 applies to the penalty functions with

non-smoothness at a set with measure zero. Hence in any practical application, we can simply ignore this technical difficulty and use

$$\nabla P(\theta) = \max\{0, q_{J(\theta)}(\theta)\}\nabla q_{J(\theta)}(\theta),$$

where $J(\theta) = \arg\max_{j=1,\ldots,s} q_j(\theta)$ (note that $J(\theta)$ is uniquely defined for $\theta \notin D$). An alternative approach to handle this technical difficulty is to apply the SP gradient estimate directly to the penalized cost $L(\theta) + rP(\theta)$ and adopt the convergence analysis presented in (15) for nondifferentiable optimization with additional convexity assumptions.

Use of non-differentiable penalty functions might allow us to avoid the difficulty of ill-conditioning as $r_n \to \infty$ without using the more complicated penalty function methods such as the augmented Lagrangian method used here. The rationale here is that there exists a constant $\bar{r} = \sum_{j=1}^s \lambda_j^*$ (λ_j^* is the Lagrange multiplier associate with the jth constraint) such that the minimum of $L + rP$ is identical to the solution of the original constrained problem for all $r > \bar{r}$, based on the theory of exact penalties (see; for example, section 4.3 of (1)). This property of the absolute value penalty function allow us to use a constant penalty parameter $r > \bar{r}$ (instead of $r_n \to \infty$) to avoid the issue of ill-conditioning. However, it is difficult to obtain a good estimate for $\bar{r}$ in our situation where the analytical expression of $g(\cdot)$ (the gradient of the cost function $L(\cdot)$) is not available. And it is not clear that the application of exact penalty functions with $r_n \to \infty$ would lead to better performance than the augmented Lagrangian based technique. In Section 4 we will also illustrate (via numerical results) the potential poor performance of the algorithm with an arbitrarily chosen large r.

3.2   *Asymptotic Normality*

When differentiable penalty functions are used, we can establish the asymptotic normality for the proposed algorithms. In the case where $q_j(\theta^*) < 0$ for all $j = 1, \ldots, s$ (that is, there is no active constraint at θ^*), the asymptotic behavior of the algorithm is exactly the same as the unconstrained SPSA algorithm and has been established in (9). Here we consider the case where at least one of constraints is active at θ^*, that is, the set $A \triangleq \{j = 1, \ldots s : q_j(\theta^*) = 0\}$ is not empty. We establish the asymptotic normality for the algorithm with smooth penalty functions of the form

$$P(\theta) = \sum_{j=1}^s p_j(q_j(\theta)),$$

which including both the quadratic penalty and augmented Lagrangian functions.

Assume further that $E[e_n|\mathcal{F}_n, \Delta_n] = 0$ a.s., $E[e_n^2|\mathcal{F}_n] \to \sigma^2$ a.s., $E[(\Delta_n^i)^{-2}] \to \rho^2$, and $E[(\Delta_n^i)^2] \to \xi^2$, where $\mathcal{F}_n$ is the σ-algebra generated by $\theta_1, \ldots, \theta_n$. Let $H(\theta)$ denote the Hessian matrix of $L(\theta)$ and

$$H_p(\theta) = \sum_{j \in A} \nabla^2 \left(p_j(q_j(\theta)) \right).$$

The next proposition establishes the asymptotic normality for the proposed algorithm with the following choice of parameters: $a_n = an^{-\alpha}$, $c_n = cn^{-\gamma}$ and $r_n = rn^\eta$ with $a, c, r > 0$, $\beta = \alpha - \eta - 2\gamma > 0$, and $3\gamma - \frac{\alpha}{2} + \frac{3\eta}{2} \geq 0$.

PROPOSITION 3.3 *Assume that conditions (C.1-6) hold. Let P be orthogonal with $PH_p(\theta^*)P^T = a^{-1}r^{-1}diag(\lambda_1, \ldots, \lambda_d)$ Then*

$$n^{\beta/2}(\theta_n - \theta^*) \xrightarrow{dist} N(\mu, PMP^T), \quad n \to \infty$$

where $M = \frac{1}{4}a^2 r^2 c^{-2}\sigma^2\rho^2 diag[(2\lambda_1 - \beta_+)^{-1}, \ldots, (2\lambda_d - \beta_+)^{-1}]$ with $\beta_+ = \beta < 2\min_i \lambda_i$ if $\alpha = 1$ and

$\beta_+ = 0$ *if* $\alpha < 1$, *and*

$$\mu = \begin{cases} 0 & \text{if } 3\gamma - \frac{\alpha}{2} + \frac{3\eta}{2} > 0, \\ (arH_p(\theta^*) - \frac{1}{2}\beta_+ I)^{-1}T & \text{if } 3\gamma - \frac{\alpha}{2} + \frac{3\eta}{2} = 0, \end{cases}$$

where the lth element of T is

$$-\frac{1}{6}ac^2\xi^2 \left[L_{lll}^{(3)}(\theta^*) + 3 \sum_{i=1, i\neq l}^{p} L_{iii}^{(3)}(\theta^*) \right].$$

Proof For large enough n, we have

$$E[\hat{g}_n(\theta_n)|\theta_n] = H(\bar{\theta}_n)(\theta_n - \theta^*) + b_n(\theta_n),$$
$$\nabla P(\theta_n) = H_p(\bar{\theta}'_n)(\theta_n - \theta^*),$$

where $b_n(\theta_n) = E[\hat{g}_n(\theta_n) - g(\theta_n)|\theta_n]$. Rewrite the algorithm into

$$\theta_{n+1} - \theta^* = (I - n^{-\alpha+\eta}\Gamma_n)(\theta_n - \theta^*) + n^{-(\alpha-\eta+\beta)/2}\Phi_n V_n$$
$$+ n^{-\alpha+\eta-\beta/2}T_n,$$

where

$$\Gamma_n = an^{-\eta}H(\bar{\theta}_n) + arH_p(\bar{\theta}'_n),$$
$$V_n = n^{-\gamma}[\hat{g}_n(\theta_n) - E(\hat{g}_n(\theta_n)|\theta_n)]$$
$$\Phi_n = -aI,$$
$$T_k = -an^{\beta/2-\eta}b_n(\theta_n).$$

Following the techniques used in (9) and the general Normality results from (16) we can establish the desired result. □

Note that based on the result in Proposition 3.3, the convergence rate at $n^{1/3}$ is achieved with $\alpha = 1$ and $\gamma = \frac{1}{6} - \eta > 0$.

4 Numerical Experiments

We test our algorithm on a constrained optimization problem described in (17, p. 352):

$$\min_{\theta} L(\theta) = \theta_1^2 + \theta_2^2 + 2\theta_3^2 + \theta_4^2 - 5\theta_1 - 5\theta_2 - 21\theta_3 + 7\theta_4$$

subject to

$$q_1(\theta) = 2\theta_1^2 + \theta_2^2 + \theta_3^2 + 2\theta_1 - \theta_2 - \theta_4 - 5 \leq 0$$
$$q_2(\theta) = \theta_1^2 + \theta_2^2 + \theta_3^2 + \theta_4^2 + \theta_1 - \theta_2 + \theta_3 - \theta_4 - 8 \leq 0$$
$$q_3(\theta) = \theta_1^2 + 2\theta_2^2 + \theta_3^2 + 2\theta_4^2 - \theta_1 - \theta_4 - 10 \leq 0.$$

The minimum cost $L(\theta^*) = -44$ under constraints occurs at $\theta^* = [0, 1, 2, -1]^T$ where the constraints $q_1(\cdot) \leq 0$ and $q_2(\cdot) \leq 0$ are active. The Lagrange multiplier is $[\lambda_1^*, \lambda_2^*, \lambda_3^*]^T = [2, 1, 0]^T$. The problem had

not been solved to satisfactory accuracy with deterministic search methods that operate directly with constraints (claimed by (17)). Further, we increase the difficulty of the problem by adding i.i.d. zero-mean Gaussian noise to $L(\theta)$ and assume that only noisy measurements of the cost function L are available (without the gradient). The initial point is chosen at $[0, 0, 0, 0]^T$; and the standard deviation of the added Gaussian noise is 4.0.

We consider three different penalty functions:

- **Quadratic penalty function:**

$$P(\theta) = \frac{1}{2} \sum_{j=1}^{s} [\max\{0, q_j(\theta)\}]^2 . \tag{11}$$

In this case the gradient of $P(\cdot)$ required in the algorithm is

$$\nabla P(\theta) = \sum_{j=1}^{s} \max\{0, q_j(\theta)\} \nabla q_j(\theta). \tag{12}$$

- **Augmented Lagrangian:**

$$P(\theta) = \frac{1}{2r^2} \sum_{j=1}^{s} [\max\{0, \lambda_j + rq_j(\theta)\}]^2 - \lambda^2. \tag{13}$$

In this case, the actual penalty function used will vary over iteration depending on the specific value selected for r_n and λ_n. The gradient of the penalty function required in the algorithm for the nth iteration is

$$\nabla P(\theta) = \frac{1}{r_n} \sum_{j=1}^{s} \max\{0, \lambda_{nj} + r_n q_j(\theta)\} \nabla q_j(\theta). \tag{14}$$

To properly update λ_n, we adopt a variation of the multiplier method (1):

$$\lambda_{nj} = \max\{0, \lambda_{nj} + r_n q_j(\theta_n), M\}, \tag{15}$$

where λ_{nj} denotes the jth element of the vector λ_n, and $M \in \mathcal{R}^+$ is a large constant scalar. Since (15) ensures that $\{\lambda_n\}$ is bounded, convergence of the minimum of $\{L_n(\cdot)\}$ remains valid. Furthermore, $\{\lambda_n\}$ will be close to the true Lagrange multiplier as $n \to \infty$.

- **Absolute value penalty function:**

$$P(\theta) = \max_{j=1,\dots,s} \{\max\{0, q_j(\theta)\}\} . \tag{16}$$

The gradient of $P(\cdot)$ when it exists is

$$\nabla P(\theta) = \max\{0, q_{J(\theta)}(\theta)\} \nabla q_{J(\theta)}(\theta), \tag{17}$$

where $J(\theta) = \arg\max_{j=1,\dots,s} q_j(\theta)$.

For all the simulation we use the following parameter values: $a_n = 0.1(n + 100)^{-0.602}$ and $c_n = n^{-0.101}$. These parameters for a_n and c_n are chosen following a practical implementation guideline recommended in (18). For the augmented Lagrangian method, λ_n is initialized as a zero vector. For the quadratic penalty function and augmented Lagrangian, we use $r_n = 10n^{0.1}$ for the penalty parameter. For the absolute value penalty function, we consider two possible values for the constant penalty: $r_n = r = 3.01$ and $r_n = r = 10$.
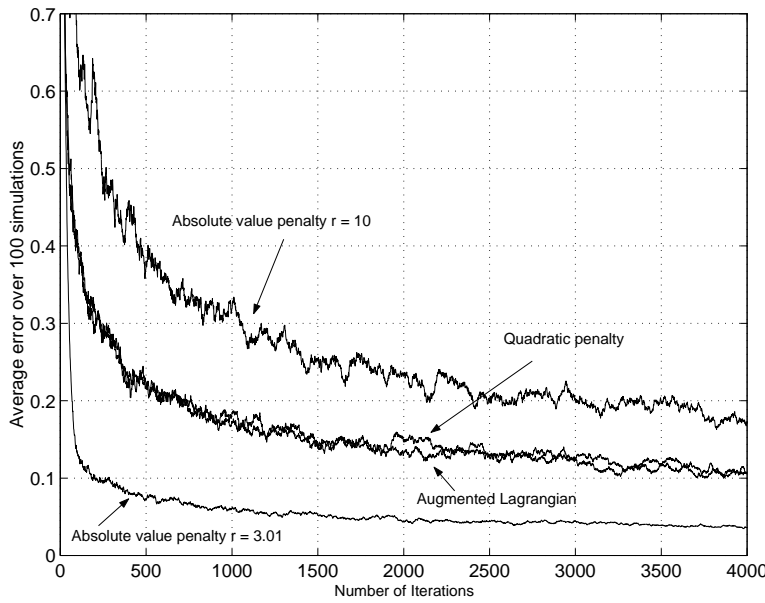
Figure 1. Error to the optimum ($\|\theta_n - \theta^*\|$) averaged over 100 independent simulations.

Note that in our experiment, $\bar{r} = \sum_{j=1}^{s} \lambda_j^* = 3$. Hence the first choice of r at 3.01 is theoretically optimal but not practical since there is no reliable way to estimate $\bar{r}$. The second choice of r represent a more typical scenario where an upper bound on $\bar{r}$ is estimated.

   Figure 1 plots the averaged errors (over 100 independent simulations) to the optimum over 4000 iteration of the algorithms. The simulation result in Figure 1 seems to suggest that the proposed algorithm with the quadratic penalty function and the augmented Lagrangian led to comparable performance (the augmented Lagrangian method performed slightly better than the standard quadratic technique). This suggests that a more effective update scheme for λ_n than (15) is needed for the augmented Lagrangian technique. The absolute value function with $r = 3.01 (\approx \bar{r} = 3)$ has the best performance. However, when an arbitrary upper bound on $\bar{r}$ is used ($r = 10$), the performance is much worse than both the quadratic penalty function and the augmented Lagrangian. This illustrates a key difficulty in effective application of the exact penalty theorem with the absolute penalty function.

5   Conclusions and Remarks

We present a stochastic approximation algorithm based on penalty function method and a simultaneous perturbation gradient estimate for solving stochastic optimization problems with general inequality constraints. We also present a general convergence result and the associated asymptotic Normality for the proposed algorithm. Numerical results are included to demonstrate the performance of the proposed algorithm with the standard quadratic penalty function and a more complicated penalty function based on the augmented Lagrangian method.

   In this paper, we consider the explicit constraints where the analytical expressions of the constraints are available. It is also possible to apply the same algorithm with appropriate gradient estimate for $P(\theta)$ to problems with implicit constraints where constraints can only be measured or estimated with possible errors. The success of this approach would depend on efficient techniques to obtain unbiased gradient estimate of the penalty function. For example, if we can measure or estimate a value of the penalty function $P(\theta_n)$ at arbitrary location with zero-mean error, then the SP gradient estimate can be applied. Of course, in this situation further assumptions on r_n need to be satisfied (in general, we would at least need $\sum_{n=1}^{\infty} (a_n r_n / c_n)^2 < \infty$). However, in a typical application, we most likely can only measure the value of constraint $q_j(\theta_n)$ with zero-mean error. Additional bias would be present if the standard finite-difference or the SP techniques were applied to estimate $\nabla P(\theta_n)$ directly in this situation. A novel technique to

obtain unbiased estimate of $\nabla P(\theta_n)$ based on a reasonable number of measurements is required to make the algorithm proposed in this paper feasible in dealing with implicit constraints.

References

[1] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, MA, 1995.

[2] P. Dupuis and H. J. Kushner, "Asymptotic behavior of constrained stochastic approximations via the theory of large deviations," *Probability Theory and Related Fields*, vol. 75, pp. 223–274, 1987.

[3] H. Kushner and D. Clark, *Stochastic Approximation Methods for Constrained and Unconstrained Systems.* Springer-Verlag, 1978.

[4] J. Kiefer and J. Wolfowitz, "Stochastic estimation of the maximum of a regression function," *Ann. Math. Statist.*, vol. 23, pp. 462–466, 1952.

[5] H. Kushner and G. Yin, *Stochastic Approximation Algorithms and Applications.* Springer-Verlag, 1997.

[6] P. Sadegh, "Constrained optimization via stochastic approximation with a simultaneous perturbation gradient approximation," *Automatica*, vol. 33, no. 5, pp. 889–892, 1997.

[7] J. Hiriart-Urruty, "Algorithms of penalization type and of dual type for the solution of stochastic optimization problems with stochastic constraints," in *Recent Developments in Statistics* (J. Barra, ed.), pp. 183–2219, North Holland Publishing Company, 1977.

[8] G. C. Pflug, "On the convergence of a penalty-type stochastic optimization procedure," *Journal of Information and Optimization Sciences*, vol. 2, no. 3, pp. 249–258, 1981.

[9] J. C. Spall, "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," *IEEE Transactions on Automatic Control*, vol. 37, pp. 332–341, March 1992.

[10] I-J. Wang and J. C. Spall, "A constrained simultaneous perturbation stochastic approximation algorithm based on penalty function method," in *Proceedings of the 1999 American Control Conference*, vol. 1, pp. 393–399, San Diego, CA, June 1999.

[11] D. P. Bertsekas, *Constrained Optimization and Lagragne Mulitplier Methods*, Academic Press, NY, 1982.

[12] E. K. Chong and S. H. Żak, *An Introduction to Optimization.* New York, New York: John Wiley and Sons, 1996.

[13] J. C. Spall and J. A. Cristion, "Model-free control of nonlinear stochastic systems with discrete-time measurements," *IEEE Transactions on Automatic Control*, vol. 43, no. 9, pp. 1178–1200, 1998.

[14] I-J. Wang, *Analysis of Stochastic Approximation and Related Algorithms.* PhD thesis, School of Electrical and Computer Engineering, Purdue University, August 1996.

[15] Y. He, M. C. Fu, and S. I. Marcus, "Convergence of simultaneous perturbation stochastic approximation for nondifferentiable optimization," *IEEE Transactions on Automatic Control*, vol. 48, no. 8, pp. 1459–1463, 2003.

[16] V. Fabian, "On asymptotic normality in stochastic approximation," *The Annals of Mathematical Statistics*, vol. 39, no. 4, pp. 1327–1332, 1968.

[17] H.-P. Schwefel, *Evolution and Optimum Seeking.* John Wiley and Sons, Inc., 1995.

[18] J. C. Spall, "Implementation of the simultaneous perturbation algorithm for stochastic optimization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 3, pp. 817–823, 1998.