

A Constrained Simultaneous Perturbation Stochastic Approximation Algorithm Based on Penalty Functions*

I-JENG WANG[†] and JAMES C. SPALL[‡]
The Johns Hopkins University
Applied Physics Laboratory

ABSTRACT

In this paper, 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 presents a very general convergence result for the proposed algorithm.

KEYWORDS: stochastic approximation, constrained optimization, penalty function, SPSA.

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:

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

where $L: \mathbb{R}^d \rightarrow \mathbb{R}$ is a real-valued cost function, $\theta \in \mathbb{R}^d$ is the parameter vector, and $G \subset \mathbb{R}^d$ 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 . Therefore, to solve the optimization problem, we have to resort to numerical techniques. Throughout the paper we use θ_n to denote the n th estimate of the solution θ^* .

*This work was supported by the JHU/APL Independent Research and Development Program.

[†]The Johns Hopkins University, Applied Physics Laboratory, 11100 Johns Hopkins Road, Laurel, MD 20723-6099. Tel: (240) 228-6204, E-mail: I-Jeng.Wang@jhuapl.edu.

[‡]The Johns Hopkins University, Applied Physics Laboratory, 11100 Johns Hopkins Road, Laurel, MD 20723-6099. Tel: (240) 228-4960, E-mail: James.Spall@jhuapl.edu.

1.1. Available Techniques

Although a huge amount of literatures is available in constrained deterministic optimization (see, for examples [1, 2, 8]), few concrete results can be found 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:

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

where $\pi_G: \mathbb{R}^p \rightarrow G$ is the set projection operator, and $\hat{g}_n(\theta_n)$ is an estimate of the gradient $g(\theta_n)$; see, for example [3, 6, 7, 11]. 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 [5] and Pflug [10] 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 [6] present several SA algorithms based on the Lagrange multiplier method, the penalty function method and a combination of both. However, 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 . Furthermore, the implementations of these algorithms involve choosing various design parameters that greatly affect the efficiency of these algorithms. More systematic studies of the convergence of these algorithms are necessary to make them practical.

The ultimate objective of the research undertaken in this paper is to establish a general framework and practical guidelines for solving stochastic optimization problems with general constraints. As a first step, we will focus on

the penalty function method and inequality constraints in this paper. Specifically, we present and study the convergence of an algorithm based on the penalty function method and the Simultaneous Perturbation (SP) gradient estimate [14]. Note that our goal here is not to weaken the required conditions for convergence of existing SA algorithms based on the penalty function method (for example, [10]), but to identify the limitations of the penalty function approach and lay out the foundation for further examination of the effects of different penalty function designs on efficiency of the algorithm.

Before we enter the specific topic of the penalty function method, we give a brief discussion on the different types of constraints. We believe that it is crucial to identify the classes of the constraints as different types of constraints normally require different optimization techniques.

1.2. A Concise Taxonomy of Constraints

Depending on motivation behind or causes of the constraints, the constraint set G in a stochastic optimization problem can take many different forms. Just as in classical deterministic optimization it is important to approach each category of constraints differently in solving a stochastic optimization problem.

• Hard vs Soft

Hard Constraints: Many constraints are the consequence of physical limitation on real systems. Therefore no measurements or observations can be taken with parameters violating these constraints. The projection method (2) described above seems to be the only effective method to deal with them.

Soft Constraints: These constraints are usually results of performance considerations. Hence measurements are allowed at the parameters lying outside the constraint set during the course of optimization. Nevertheless, it is required that the optimization algorithms eventually produce estimates inside the constraint set within a reasonable amount of time.

• “Nice” vs Nonlinear

“Nice” Constraints: We classify a constraint as “nice” if there exists an “efficient” algorithm to implement the projection operator π_G on \mathbb{R}^d . Generally speaking, the operator π_G is well defined for a large class of constraint sets. However, only for a small set of constraints, can it be efficiently computed; these includes interval constraints such as $\theta^i \in [l, u]$ and linear constraints (constraints specified by linear equalities or inequalities) such as

$$A\theta \leq b.$$

Nonlinear Constraints: In most cases, the constraints are specified by inequalities or equalities on nonlinear functions of the parameter θ for which the projection operators are very difficult to implement.

• Explicit vs Implicit

Explicit Constraints: In many standard optimization problems, the constraints are specified by equalities and/or inequalities with explicit functions (that is, functions with known explicit closed-form expression). For example, we can specify the constraint set G by

$$G = \{\theta \in \mathbb{R}^d : \|\theta\|^2 \leq \epsilon\}.$$

In the case where the functions defining the constraints have nice properties (e.g., convexity, smoothness, etc.), we can normally use this information to tackle the constraints in optimization.

Implicit Constraints: In some cases the constraints are not placed directly upon the parameter θ but on some other output variables that is the manifestation of the parameter θ . As a result, we are not able to directly verify the feasibility of each arbitrary parameter value without taking a measurement or simulation (which are normally costly and noisy) at the parameter value (and even that will not be enough if there is noise in the loss measurements). These type of constraints normally arise from optimization problems associated with control designs for complex systems. For example, suppose $\{X(t, \theta), t \in [t_0, t_0 + T]\}$ is a stochastic process that represents the output trajectory of a complex system operated at the parameter θ . Then we might have a constraint set defined by

$$G = \left\{ \theta \in \mathbb{R}^d : \frac{1}{T} \int_{t_0}^{t_0+T} E[X^2(t, \theta)] dt \leq M \right\}.$$

Implicit constraints normally present the most difficult challenge in stochastic optimization.

In this paper, we focus on soft, nonlinear inequality constraints. We will treat the explicit constraint case and briefly discuss the extension to implicit constraints.

2. PENALTY FUNCTIONS

We consider a constraint set G defined by

$$G \triangleq \{\theta \in \mathbb{R}^d : q_j(\theta) \leq 0, j = 1, \dots, s\}, \quad (3)$$

where $q_j: \mathbb{R}^d \rightarrow \mathbb{R}$ are continuously differentiable real-valued functions that may or may not be explicitly calculable (see “Explicit vs Implicit constraints” in Section 1). We consider a penalty function $P: \mathbb{R}^p \rightarrow \mathbb{R}$ of the following form:

$$P(\theta) = \sum_{j=1}^s w_j p(q_j(\theta)), \quad (4)$$

where w_j are positive scalar weights and $p(\cdot)$ is a real-valued function on \mathbb{R} satisfying the following conditions:

- (P1) $p(x)$ is differentiable on \mathbb{R} .
- (P2) $p(x) \geq 0$ for all $x \in \mathbb{R}$.
- (P3) $p(x) = 0$ if and only if $x \leq 0$.
- (P4) $p(\cdot)$ is strictly increasing on $[0, \infty)$.

Typical $p(\cdot)$ functions used in constrained optimization include the quadratic penalty function $p(x) = \frac{1}{2}x^2$ for $x > 0$, and the absolute penalty function $p(x) = |x|$ for $x > 0$ (even though $|x|$ is not differentiable at 0). Note that in the case where $p(\cdot)$ is differentiable (for example, $p(\cdot)$ is quadratic) we have

$$\frac{dp(x)}{dx} \begin{cases} = 0 & \text{if } x \leq 0, \\ > 0 & \text{otherwise.} \end{cases}$$

In the sequel, we will denote the derivative of $p(x)$ by $p'(x)$. The differentiability of both $p(\cdot)$ and the constraint function $q_j(\cdot)$ guarantees that the penalty function $P(\cdot)$ is also differentiable. This smoothness of the penalty function P is useful when the gradient descent type of algorithms are used for the constrained optimization problem. Nevertheless, as we will discuss in the following sections, increases in the smoothness of the penalty function (as in the case of quadratic penalty function relative to the absolute penalty function) is potentially problematic for the convergence of the algorithms around the solution of the optimization problem.

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

$$\min_{\theta} L(\theta) + rP(\theta), \quad (5)$$

where r is a positive real number normally referred to as the penalty parameter. To guarantee the solution of the unconstrained optimization problem (5) coincides with the solution of the original problem (1), we normally need

to choose the penalty parameter r larger than a threshold \bar{r} that depends on both $L(\cdot)$ and $P(\cdot)$. Given that we do not have complete information on L , we “slowly” increase the penalty parameter r during the course of optimization. In other words, we consider a sequence of unconstrained optimization problems defined by

$$\min_{\theta \in \mathbb{R}^d} L(\theta) + r_n P(\theta) \quad (6)$$

with a sequence of penalty parameters $\{r_n\}$ satisfying $r_n > 0$ and $\lim_{n \rightarrow \infty} r_n = \infty$. We will denote the penalized cost function $L(\cdot) + r_n P(\cdot)$ by $L_n(\cdot)$. Assuming that $p(\cdot)$ is differentiable, a standard result in nonlinear programming (see, for example [2]) gives us

Proposition 1. *Suppose for each large n , there exists a unique $\theta_n^* \in \mathbb{R}^d$ such that $\nabla L_n(\theta_n^*) = 0$, and $p(\cdot)$ satisfies conditions (P1–P4). Then $\lim_{n \rightarrow \infty} \theta_n^* = \theta^*$.*

In this paper, we consider a stochastic approximation algorithm that minimize the sequence of the penalized cost function $\{L_n(\cdot)\}$ based on the SP gradient estimate. The algorithm is of the following form

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

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. Later on, we will discuss the possibility of estimating the gradient $\nabla P(\theta)$ on line based on the SP gradient estimate.

3. MAIN RESULT

In this section, we present the specific form of the algorithm for solving the constrained stochastic optimization problem and give a convergence theorem under very general assumptions. The algorithm we consider is defined by

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

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 $\lim_{n \rightarrow \infty} r_n = \infty$, $\nabla P(\theta)$ is the gradient of $P(\theta)$ at θ , and $\{a_n\}$ is a positive scalar sequence satisfying $a_n \rightarrow 0$ 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

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

where $\Delta_n \in \mathbb{R}^p$ is a random perturbation vector, $c_n \rightarrow 0$ is a positive sequence, ϵ_n^+ and ϵ_n^- are noise in the measurements. and $\frac{1}{\Delta_n}$ denotes the vector $[\frac{1}{\Delta_n^1}, \dots, \frac{1}{\Delta_n^p}]$. For analysis, we rewrite the algorithm (7) into

$$\begin{aligned} \theta_{n+1} = & \theta_n - a_n g(\theta_n) - a_n r_n \nabla P(\theta_n) + a_n d_n \\ & - a_n \frac{e_n}{2c_n \Delta_n}, \end{aligned} \quad (9)$$

where d_n and e_n are defined by

$$\begin{aligned} d_n &\triangleq g(\theta_n) - \frac{L(\theta_n + c_n \Delta_n) - L(\theta_n - c_n \Delta_n)}{2c_n \Delta_n}, \\ e_n &\triangleq \epsilon_n^+ - \epsilon_n^-, \end{aligned}$$

respectively.

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:

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

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 [15] for an algorithm in the generic form (10).

Theorem 1. Assume the following conditions hold:

- (A.1) For each n large enough ($\geq N$ for some $N \in \mathbb{N}$), there exists a unique θ_n^* such that $f_n(\theta_n^*) = 0$. Furthermore, $\lim_{n \rightarrow \infty} \theta_n^* = \theta^*$.
- (A.2) $d_n \rightarrow 0$, 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 $d_n \delta_n(\rho)^{-1} \rightarrow 0$.
- (A.4) For each $i = 1, 2, \dots, d$, 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.
- (A.5) For any $\tau > 0$ and nonempty $S \subset \{1, 2, \dots, d\}$, there exists a $\rho'(\tau, S) > \tau$ such that for all $\theta \in \{\theta \in \mathbb{R}^d : |(\theta - \theta^*)_i| < \tau \text{ when } i \notin S, |(\theta - \theta^*)_i| \geq \rho'(\tau, S) \text{ when } i \in S\}$,

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

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

Based on Theorem 1, we give a convergence result for algorithm (7). We need the following assumptions:

- (C.1) There exists $K_1 \in \mathbb{N}$ such that for all $n \geq K_1$, we have a unique $\theta_n^* \in \mathbb{R}^d$ with $f_n(\theta_n^*) = g(\theta_n^*) + r_n \nabla P(\theta_n^*) = 0$.
- (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) $\sum_{k=1}^n \frac{a_k e_k}{2c_k \Delta_k}$ converges almost surely.
- (C.4) If $\|\theta - \theta^*\| \geq \rho$, then there exists a $\delta_n(\rho) > 0$ 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 $(\theta - \theta^*)^T \nabla P(\theta) \geq 0$.
 - $(\theta - \theta^*)^T g(\theta) \geq -M$ and $(\theta - \theta^*)^T \nabla P(\theta) \geq \delta(\rho) \|\theta - \theta^*\| > 0$
- (C.5) $a_n r_n \rightarrow 0$, $g(\cdot)$ and $\nabla P(\cdot)$ are Lipschitz. (See comments below)
- (C.6) $f_n(\cdot) = g(\cdot) + r_n \nabla P(\cdot)$ satisfies condition (A5)

Theorem 2. Suppose that assumptions (C.1–C.6) hold. Then the sequence $\{\theta_n\}$ defined by (7) converges to θ^* almost surely.

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

- (C.1) together with Proposition 1 yield condition (A.1).
- From the results in [14, 16] and assumptions (C.2–C.3), we can show that condition (A.2) hold.
- Since $r_n \rightarrow \infty$, 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).

□

Theorem 2 given above is very 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. In situations similar to this, application of an alternative method based on a combination of the Lagrangian and penalty function methods, called *augmented Lagrangian method* [1], might be required. However, currently there exists no general convergence result for the augmented Lagrangian methods in a stochastic setting. We also like to point out that assumption (C.4) is satisfied if cost function L and constraint functions $q_j, j = 1, \dots, s$ are convex and satisfy the Slater condition, that is, the minimum cost function value $L(\theta^*)$ is finite and there exists a $\theta \in \mathbb{R}^d$ such that $q_j(\theta) < 0$ for all j (this is that case studied in [10]). 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 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 point with zero measure, for example, the absolute value penalty function. In this case, condition (C.4) can be modified by replacing the inner product $(\theta - \theta^*)^T \nabla P(\theta)$ with directional derivative of $P(\theta)$ along $(\theta - \theta^*)$. Relaxation of condition (C.5) where the differentiability of P is required explicitly is more involved. We can take advantage of the fact that iterations of the algorithm visit any zero-measure set with zero probability. Nevertheless, some technical details need to be worked out. Another possible approach suggested in [4] to circumvent this difficulty of non-smoothness is to approximate $P(\theta)$ by a smooth expected function

$$\bar{P}(\theta) = \int P(\theta + \delta\theta) h(\delta\theta) d\delta\theta,$$

where $\delta\theta$ is a random vector with density function $h(\theta)$. The algorithm is then implemented with the SP gradient estimate of $\bar{P}(\theta)$ that is differentiable.

3.2. Choices of Penalty Functions

The choice of penalty function $p(\cdot)$ obviously has a great impact on convergence of the proposed algorithm. The quadratic penalty function $p(x) = \frac{1}{2}x^2$ is a typical choice due to its differentiability. Nevertheless, an implementation of the proposed algorithm with the quadratic penalty function may lead to divergence. This can be partially justified by the possibility that condition (C.4) cannot be satisfied with a quadratic penalty function: consider a case where $q_j(\theta) > 0, q_i(\theta) \leq 0$ for $i \neq j$, and $q_j(\theta^*) = 0$, then $\nabla P(\theta) = w_j q_j(\theta) \nabla q_j(\theta)$ can be arbitrarily small even if $\|\theta - \theta^*\|$ is large. Moreover, it has been shown [1] that, with the quadratic penalty function, the penalized cost function $L_n = L + r_n \nabla P$ 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). Therefore, even if the algorithm does converge, it might converge extremely slowly.

As we mentioned earlier, we consider the explicit constraints for the most part of the paper. It is also possible to apply the same algorithm with appropriate gradient estimate for $P(\theta)$ to problems with implicit constraints. The success of this approach would very much 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 $\sum_{n=1}^{\infty} \left(\frac{a_n r_n}{c_n} \right)^2 < \infty$). 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. One possible way to resolve this “bias” problem, as presented in [10], is to use the knowledge of $p'(\cdot)$ and estimate $\nabla P(\theta_n)$ by

$$\sum_{\{j: q_j(\theta_n) > 0\}} w_j p'(\bar{q}_j(\theta_n)) \hat{\nabla} q_j(\theta_n),$$

where $\bar{q}_j(\theta_n)$ is the average of multiple measurements of $q_j(\theta_n)$, and $\hat{\nabla} q_j(\theta_n)$ is the SP estimate of $\nabla q_j(\theta_n)$ based on measurements of $q_j(\cdot)$ at randomly perturbed points around θ_n . Under appropriate assumptions, the above gradient estimate is unbiased [10]. The major drawback of the above gradient estimation is that the number of

measurements required for computing $\bar{q}_j(\theta_n)$ increase to ∞ with n . A novel technique to obtain unbiased estimate of $\nabla P(\theta_n)$ based on a smaller number of measurements is required to make the algorithm proposed in this paper feasible in dealing with implicit constraints. Note that, with the absolute penalty function, the standard SP gradient estimate is unbiased almost everywhere (except at the boundary of the constraint set) if the constraints can be measured with zero-mean error.

4. SIMULATION RESULTS

We test our algorithm on a constrained optimization problem described in [12, p.352]:

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

subject to

$$\begin{aligned} q_1(\theta) &= 2\theta_1^2 + \theta_2^2 + \theta_3^2 + 2\theta_1 \\ &\quad -\theta_2 - \theta_4 - 5 \leq 0 \\ q_2(\theta) &= \theta_1^2 + \theta_2^2 + \theta_3^2 + \theta_4^2 + \theta_1 - \theta_2 \\ &\quad + \theta_3 - \theta_4 - 8 \leq 0 \\ q_3(\theta) &= \theta_1^2 + 2\theta_2^2 + \theta_3^2 + 2\theta_4^2 \\ &\quad -\theta_1 - \theta_4 - 10 \leq 0. \end{aligned}$$

The minimum occurs at $\theta^* = [0, 1, 2, -1]^T$ where only the constraint $q_1(\cdot) \leq 0$ is active. The problem had not been solved to satisfactory accuracy with deterministic search methods that operate directly with constraints (claimed by [12]). 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 gradient). The initial point is chosen at $[0, 0, 0, 0]^T$; and the standard deviation of the added Gaussian noise is 1 (roughly $\frac{1}{2}$ of the initial error).

We consider a class of penalty function $p(x) = \frac{x^\beta}{\beta}, 2 \geq \beta \geq 1$. In the case where $\beta = 1$, that is $P(\theta) = \sum_j w_j \max(0, q_j(\theta))$, we ignore the technical difficulty that $P(\theta)$ is not differentiable at θ with $q_j(\theta) = 0$ and implement the algorithm with $\nabla P(\theta) = \sum_j w_j q'(\theta)$ (See comments in section 3 on fact that differentiability is likely not needed). For all the simulation we use the following parameter values: $a_n = 0.1n^{-0.602}$, $c_n = n^{-0.101}$, and $r_n = 0.5 \log(n^{0.5})$. These parameters (except for r_n) are chosen following a practical implementation guideline recommended in [13]. We discovered from simulations that, with any penalty function in the considered class with $\beta > 1$, the proposed algorithm can diverge quickly.

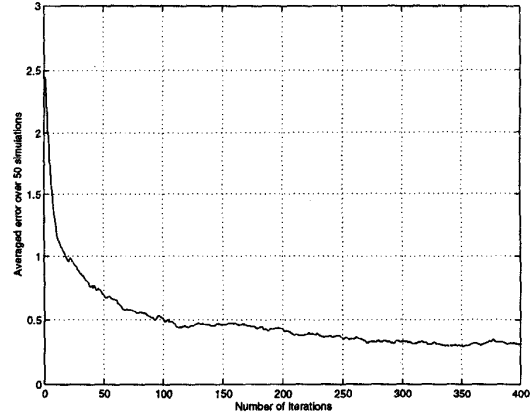


Figure 1: Average error for absolute function

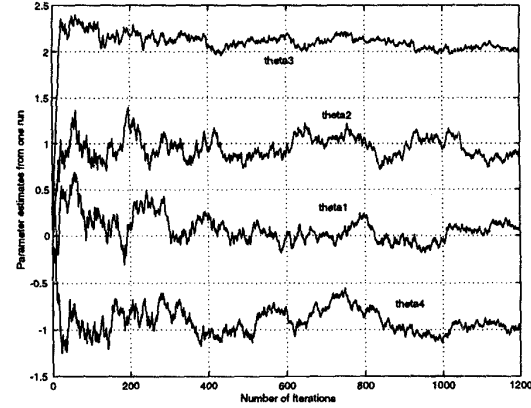


Figure 2: A sample path for the proposed algorithm

The algorithm with the quadratic penalty ($\beta = 2$) is especially unstable and sensitive to noise. Despite the technical difficulty with the absolute value penalty function ($\beta = 1$), the algorithm is stable and always converges to a small neighborhood of θ^* . Here we only include the simulation plots for the algorithm with absolute value function. Figure 4 shows the averaged error of the algorithm over 50 runs. Figure 4 shows a typical sample path of the vector sequence $\{\theta_n\}$. As shown in Figure 4, the estimate oscillates around the optimum with small error (compared with the magnitude of noise).

5. CONCLUSIONS

In this paper, 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 presents a very general convergence result for the proposed algorithm. Numerical results are included to demonstrate the performance of the proposed algorithm.

As discussed in the last section, much research is needed to establish a general framework and guidelines for dealing with general constraints in stochastic optimization. We summarize here some important directions for future research:

- Based on the proposed general convergence result, we wish to further study specific types of penalty functions and analyze convergence of the corresponding SA algorithms.
- We wish to establish a rigorous framework for analyzing the asymptotic convergence rates of SA algorithms based on penalty function methods. This rate analysis should provide guidelines for selecting appropriate penalty functions to improve convergence speed.
- We also wish to develop an efficient unbiased gradient estimator for the penalty function and apply the proposed algorithm to problems with implicit constraints.
- We wish to test our algorithm on a large-scale optimization problem considered in [9] for optimization of complex systems of systems.

References

- [1] D. P. Bertsekas, *Nonlinear Programming*. Belmont, Massachusetts: Athena Scientific, 1995.
- [2] E. K. Chong and S. H. Żak, *An Introduction to Optimization*. New York, New York: John Wiley and Sons, 1996.
- [3] 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.
- [4] L. Gerencsér, G. Kozmann, and Z. Vágó, "The use of the SPSA method in ECG analysis," *Manuscript*, 1998.
- [5] 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.
- [6] H. Kushner and D. Clark, *Stochastic Approximation Methods for Constrained and Unconstrained Systems*. Springer-Verlag, 1978.
- [7] H. Kushner and G. Yin, *Stochastic Approximation Algorithms and Applications*. Springer-Verlag, 1997.
- [8] D. G. Luenberger, *Linear and Nonlinear Programming*. Reading, Massachusetts: Addison Wesley, 2 ed., 1984.
- [9] R. R. Luman, *Quantitative Decision Support for Upgrading Complex Systems of Systems*. PhD thesis, The School of Engineering and Applied Science, The George Washington University, November 1997.
- [10] 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.
- [11] P. Sadegh, "Constrained optimization via stochastic approximation with a simultaneous perturbation gradient approximation," *Automatica*, vol. 33, no. 5, pp. 889–892, 1997.
- [12] H.-P. Schwefel, *Evolution and Optimum Seeking*. John Wiley and Sons, Inc., 1995.
- [13] J. C. Spall, "Implementation of the simultaneous perturbation algorithm for stochastic optimization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 3, 1998. To Appear.
- [14] J. C. Spall, "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," *IEEE Transactions on Automatic Control*, vol. 37, no. 3, pp. 332–341, March 1992.
- [15] 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. 8, 1998. To Appear.
- [16] I.-J. Wang, *Analysis of Stochastic Approximation and Related Algorithms*. PhD thesis, Purdue University, 1996.