

STOCHASTIC APPROXIMATION FOR FUNCTION MINIMIZATION UNDER QUANTIZATION ERROR

László Gerencsér

Computer and Automation Institute of
the Hungarian Academy of Sciences,
Kende 13-17, Budapest, 1111, Hungary
gerencser@sztaki.hu

Zsuzsanna Vágó

Computer and Automation Institute of
the Hungarian Academy of Sciences,
13-17 Kende u., Budapest, 1111, Hungary
vago@novell.olplab.sztaki.hu

Abstract

The SPSA (simultaneous perturbation stochastic approximation) method developed in [12] is applied and analyzed for function minimization under quantization error. Following [7] it is proved that under certain conditions the estimator sequence converges with rate $O(n^{-\beta/2})$ for some $\beta > 0$, where the rate is measured by the L_q -norm of the estimation error for any $1 \leq q < \infty$. The viability of SPSA for the present problem will also be demonstrated by simulation results.

stochastic approximation, Kiefer-Wolfovitz method, SPSA method, optimization, quantization

1 Introduction

Estimation under quantization error is an important problem in signal processing and high accuracy stochastic adaptive control (cf. [11, 8]). The interest in the latter methodology has been motivated by control problems of surgical microrobots, where quantization errors are due to low accuracy of sensors and A/D or D/A conversion.

To put the present paper in perspective consider the problem is quantized linear regression, following [11, 4, 5]. Let α^* be a scalar-valued, unknown physical quantity, which is to be determined by some measurement procedure. Measurements are performed in an environment corrupted by external noise, such as vibration, using a device of low precision. The additive external noise will be denoted by e , and the actual measurement will be written as $q(\alpha^* + e)$, where q is a *quantizer*. This means a mapping from the real line into integer multiples of a fixed, small positive number, say h : if x is a real number with $nh - h/2 \leq x < nh + h/2$, then $q(x) = nh$. At time k we get the measurement

$$y_k = q(\alpha^* + e_k),$$

where the unknown noise process is assumed to be an independent sequence of Gaussian random variables with distribution $N(0, \sigma^2)$, with σ^2 known.

It is easy to see that the parameter α^* is identifiable on the basis of the measurements y_k . Note that, in contrast to plain linear regression, the a priori knowledge of the distribution of the additive noise is an essential ingredient of the method.

In this paper the following related problem is considered: minimize a function $L(\cdot)$ which is defined over an open domain $\theta \in D \subset \mathbb{R}^p$. At any time k and for each θ we have a measurement

$$M(k, \theta) = q(L(\theta) + e_k) \quad (1)$$

where $e_k = e_k(\omega)$ is a sequence of i.i.d. random variables defined over some probability space $(\Omega, \mathcal{F}, \mathcal{P})$. In contrast to [11, 4, 5] the distribution of (e_k) is assumed to be completely *unknown*. We would like minimize $L(\cdot)$ by an iterative procedure using measured values of $L(\theta)$.

We assume that the function $L(\cdot)$ is three times continuously differentiable with respect to θ for $\theta \in D$, and the absolute values of the derivatives up to order three are bounded. It is also assumed that the minimizing value of $L(\theta)$ is unique in D and will be denoted by θ^* .

Obviously, the function value $L(\theta)$ can not be reconstructed from these measurements. However, we have the at first sight surprising result: under reasonable technical conditions the iterative determination of the minimum of $L(\cdot)$ based on measurements of the form given above can be carried out. In fact, it is easy to see that under reasonable conditions on the additive noise e_k the minimizing point θ^* is identifiable, even if $L(\cdot)$ is not. To see this, assume that e_k has a density function $f(\cdot)$ such that the support of f contains an interval $(a, a+h)$ with some a . Then defining $\mu(\alpha) = \text{E}q(\alpha + e_k)$ it is easy to see that $\mu(\cdot)$ is strictly monotone increasing

in α , and defining

$$\bar{L}(\theta) = EM(n, \theta) = \mu(L(\theta)) \quad (2)$$

it is straightforward to see that $\bar{L}(\cdot)$ is minimized at θ^* . Under simple technical conditions on f the auxiliary function $\bar{L}(\cdot)$ is also three times continuously differentiable with respect to θ for $\theta \in D$.

2 The SPSA method

The natural approach to minimize $\bar{L}(\theta)$ is to use a stochastic approximation method such as Kiefer-Wolfowitz or SPSA. The latter method is particularly designed for problems where the experimental evaluation of the function is expensive. There are a number of recent papers analyzing the SPSA method. The technical differences among these papers are in the conditions imposed on the *noise*, the truncation procedures and the type of convergence that is obtained. In [12] the noise is a state-independent martingale-difference sequence, and boundedness of the estimator sequence is assumed a priori. A remarkable observation of [2, 3] was that the randomization procedure that is inherent in the SPSA procedure ensures almost sure convergence on a larger product-probability space for *any bounded* and state-independent noise sequence, using a randomly truncated version of the SPSA method. In [7] the state-independent noise process was assumed to satisfy certain *mixing* conditions and the rate of convergence for the moments of the estimation error has been established for the SPSA method with enforced resetting. Finally, in [6] the noise process was assumed to be mixing and *state-dependent*, without hypothesizing the standard martingale assumption. However the noise was assumed to be a sufficiently smooth function of the state. A second order version of the SPSA methods has been given in [13]. SPSA has also been analyzed within the systematic theory of stochastic approximation given in [10].

The technical noise that is relevant for the convergence analysis for the present problem is defined by

$$\varepsilon(k, \theta) = \varepsilon(k, \theta, \omega) = M(k, \theta) - \bar{L}(\theta). \quad (3)$$

This is a state-dependent but *discontinuous* noise-process, therefore the methods of [6] are applicable. It is not clear if existing results for stochastic approximation with discontinuous observations (cf. [1]) are not applicable, either. However, it turns out that the special structure of the problem can be exploited and the analysis given in [7] is applicable with small modifications.

To minimize $\bar{L}(\theta)$ we use an estimator of its gradient using simultaneous random perturbations of the components of θ . Let k denote the iteration time. At time

k we take a random vector over some probability space $(\Omega', \mathcal{F}', \mathcal{P}')$

$$\Delta_k(\omega') = (\Delta_{k1}, \dots, \Delta_{kp})^T,$$

where $\Delta_{ki} = \Delta_{ki}(\omega')$ is a double sequence of i.i.d., random variables. A standard choice is a Bernoulli-sequence so that

$$P'(\Delta_{ki}(\omega') = +1) = 1/2$$

$$P'(\Delta_{ki}(\omega') = -1) = 1/2$$

The size of the perturbation will be denoted by c_k . A standard choice for c_k is $c_k = c/k^\gamma$ with some $\gamma > 0$. Furthermore let D_0 be a compact, convex truncation domain. For each $\theta \in D_0$ we take two measurements:

$$\begin{aligned} M_k^+(\theta) &= q(L(\theta + c_k \Delta_k) + e_{2k-1}) \\ M_k^-(\theta) &= q(L(\theta - c_k \Delta_k) + e_{2k}). \end{aligned}$$

Then the estimator of the gradient given by the SPSA method, denoted by $H(k, \theta)$, is given by

$$\left[\frac{M_k^+(\theta) - M_k^-(\theta)}{2c_k \Delta_{k1}}, \dots, \frac{M_k^+(\theta) - M_k^-(\theta)}{2c_k \Delta_{kp}} \right]^T.$$

Now we define an iterative procedure: let a_k be a fixed sequence of positive numbers with a_k denoting the step-size at time k . In the sequel we shall assume that $a_k = a/k$. We start with an initial estimate θ_1 , and then a sequence of estimated parameters, denoted by $\hat{\theta}_{k+1} = \hat{\theta}_{k+1}(\omega, \omega')$, $k = 0, 1, \dots$ is generated recursively by

$$\hat{\theta}_{k+1} = \hat{\theta}_k - a_{k+1} H(k+1, \hat{\theta}_k). \quad (4)$$

Enforced boundedness of the estimator process is ensured by a resetting mechanism as follows: if $\hat{\theta}_{k+1} \notin D_0$ then redefine the estimator at time $k+1$ to be $\hat{\theta}_0$.

3 The main result

Assume that the stability condition for the associated ODE is given as Condition 2.4 in [7] is satisfied. Then using the arguments of the cited paper we get the following result:

Theorem 1 *Let $\beta = \min(4\gamma, 1 - 2\gamma) > 0$. Assume that the smallest eigenvalue of the Hessian-matrix of L at $\theta = \theta^*$, denoted by α , satisfies $\alpha > \beta/2$. Then under the conditions above*

$$\hat{\theta}_k - \theta^* = O_M(k^{-\beta/2}).$$

Here the notation $O_M(\cdot)$ means that the $L_q(\Omega \times \Omega', \mathcal{F} \times \mathcal{F}', \mathcal{P} \times \mathcal{P}')$ -norm of the left hand side decreases with the rate given on the right hand side for any $q \geq 1$.

The proof given in [7] is based on a sophisticated stochastic averaging principle. A critical term that can not be handled the old way is

$$\delta_s^*(J^\varepsilon) = \sup_{\substack{\sigma \leq t \leq q' \wedge \tau(\sigma) \\ s \leq \sigma \leq q' s}} \left| \sum_{k=s}^t \frac{a}{k} \frac{u_k}{2c_k} \right| \quad (5)$$

where $u_k = (\varepsilon(2k - 1, \hat{\theta}_{k-1} + c_k \Delta_k) - \varepsilon(2k, \hat{\theta}_{k-1} - c_k \Delta_k)) \Delta_k^{-1}$, $q' > 1$ and $\tau(\sigma)$ is the first hitting time after σ . An analogous quantity has been introduced following equation (3.18) of [7] in continuous-time and for state-independent noise.

To estimate the higher order moments of $\delta_s^*(J^\varepsilon)$ note that (u_k) is martingale-difference process, due to assumed independence of the sequences e_k and Δ_k . Note, however, that for a general state-dependent error-process the process (u_k) may be neither a martingale-difference process nor an L -mixing process. Thus in the general case the arguments of [7] are not directly applicable.

Since (u_k) is a martingale-difference process, applying Doob's maximal inequality and Burkholder's inequality (cf. Theorems 2.2 and 2.10 of [9]) we may proceed as in [7], and thus finally get the required inequality

$$\delta_s^*(J^\varepsilon) = O_M(s^{-1/2+\gamma}),$$

which is the key estimate required for the completion of the proof.

4 Tuning the noise

The effect of the additive noise e_k on the asymptotic properties of $\hat{\theta}_k$ is an interesting problem. Intuitively, for very small additive noise the performance of the algorithm will be poor. Assuming that intensity of the noise can be controlled, we may ask what is the optimal intensity to get best performance.

An insight to this problem can be obtained by using the asymptotic theory given in [12] developed under conditions different from those of the present paper. Letting

$$A = \frac{\partial^2}{\partial \theta^2} \bar{L}(\theta) \Big|_{\theta=\theta^*}, \quad \sigma^2 = E\varepsilon^2(k, \theta^*) \quad (6)$$

we have under the conditions of [12] that

$$k^{\beta/2}(\hat{\theta}_k - \theta^*) \rightarrow \mathcal{N}(0, S) \quad (7)$$

where S is the solution of the Lyapunov-equation

$$(-aA + \frac{\beta}{2}I)S + S(-aA + \frac{\beta}{2}I)^T + \frac{a^2}{c^2}\sigma^2 I = 0.$$

This Lyapunov equation is explicitly solvable by a simple coordinate transformation (cf. [12]): write $PAP^T = \text{diag}(\lambda_i)$ with some orthonormal P , then we immediately get that

$$PSP^T = \text{diag} \sigma_i \quad (8)$$

with

$$\sigma_i = \frac{a^2}{c^2} \sigma^2 (2a\lambda_i - \beta)^{-1}. \quad (9)$$

Thus we get for $\text{tr } S = \text{tr } PSP^T$

$$\text{tr } S = \sum_i \frac{a^2}{c^2} \sigma^2 (2a\lambda_i - \beta)^{-1}. \quad (10)$$

Write $y^* = L(\theta^*)$ and

$$A^* = \frac{\partial^2}{\partial \theta^2} L(\theta) \Big|_{\theta=\theta^*}$$

and let the noise intensity be denoted by λ . Introducing the notations

$$\mu(y, \lambda) = E q(y + \lambda e) \quad \sigma^2(y, \lambda) = \sigma^2 q(y + \lambda e)$$

we have

$$A = \mu_y(y^*, \lambda) A^*, \quad \sigma^2 = \sigma^2(y, \lambda).$$

Thus we get for $\text{tr } S(\lambda)$ the expression

$$\sum_i \frac{a^2}{c^2} \sigma^2(y, \lambda)^2 (2a\lambda_i \mu_y(y^*, \lambda) - \beta)^{-1}.$$

The condition $a\alpha > \beta/2$ in Theorem 1 is equivalent to the condition $2a\lambda_i \mu_y(y^*, \lambda) - \beta > 0$ for all i . Now if λ gets small then $\mu(y^*, \lambda)$ gets flat with respect to y , hence $\mu_y(y^*, \lambda)$ will get close to the value where one of the conditions above is violated. Obviously, in the vicinity of this value $\text{tr } S(\lambda)$ tends to infinity. Similarly, it can be shown that when λ tends to infinity then also $\text{tr } S(\lambda)$ tends to infinity. The empirical optimization of $\text{tr } S(\lambda)$ is an interesting open problem.

The functions $\mu(y, \lambda), \sigma^2(y, \lambda)$ are not arbitrary. Namely, empirical studies show that the function

$$I(\lambda) = \frac{\sigma^2(y^*, \lambda)}{(\mu_y(y^*, \lambda))^2}$$

is unimodal in λ for Gaussian additive noise (cf. [4]).

Simulation results have been carried out for quadratic functions in 50 dimensions. The additive noise is zero mean normal with variance σ^2 . The quantization step is 0.1. We present the results of two experiments, in which the results for a moderate variance $\sigma = 0.5$ is compared with the results for a large and a small variance.

5 Conclusion

We have shown that SPSA is applicable for the function minimization when measurements are corrupted both by additive state-independent noise and quantization. The remarkable feature of the method that it is applicable under very weak assumptions on the distribution of the noise. The paradoxical role of additive noise has been analyzed both theoretically and experimentally

Acknowledgement

The first author expresses his thanks to James C. Spall and John L. Maryak for initiating and cooperating in this research, and to Serge Lafontaine and Ian W. Hunter for bringing his attention to problems in the mathematics of microrobotics and to the work of E. Masry.

References

- [1] J.A. Bucklew, Th.G. Kurtz, and W.A. Sethares. Weak convergence and local stability properties of fixed step size recursive algorithms. *IEEE Trans. Informat. Theory*, 30:966–978, 1993.
- [2] H.F. Chen, T.E. Duncan, and B. Pasik-Duncan. A stochastic approximation algorithm with random differences. In J.Gertler, J.B. Cruz, and M. Peshkin, editors, *Proceedings of the 13th Triennial IFAC World Congress, San Francisco, USA*, pages 493–496, 1996. Volume editors: R.Bitmead, J. Petersen, H.F. Chen and G. Picci.
- [3] H.F. Chen, T.E. Duncan, and B. Pasik-Duncan. A Kiefer-Wolfowitz algorithm with randomized differences. *IEEE Trans. Automat. Contr.*, 44:442–453, 1999.
- [4] L. Finesso, L. Gerencsér, and I. Kmeacs. Estimation of parameters from quantized noisy observations. In *European Control Conference, Karlsruhe*, page to appear, 1999.
- [5] L. Finesso, L. Gerencsér, and I. Kmeacs. A randomized EM-algorithm for estimating quantized linear Gaussian regression. In *Proceedings of the Conference on Decision and Control, CDC 38. IEEE*, 1999. to appear.
- [6] L. Gerencsér. SPSA with state-dependent noise—a tool for direct adaptive control. In *Proceedings of the Conference on Decision and Control, CDC 37. IEEE*, 1998.
- [7] L. Gerencsér. Rate of convergence of moments for a simultaneous perturbation stochastic approximation method for function minimization. *IEEE Trans. Automat. Contr.*, 44:894–906, 1999.

- [8] L. Gerencsér, Zs. Vágó, I. Hunter, S. Lafontaine, and A. Horváth. High sampling rate identification of continuous time processes. In *Proceedings of the 1994 SYSID Conference*, volume 3, pages 651–656, 1994.
- [9] P. Hall and C.C. Heyde. *Martingale limit theory and its applications*. Academic Press, 1980.
- [10] H.J. Kushner and G. Yin. *Stochastic Approximation Algorithms and Applications*. Springer Verlag. New York, 1997.
- [11] E. Masry and S. Cambanis. Consistent estimation of continuous-time signals from nonlinear transformations of noisy samples. *IEEE Trans. on Information Theory*, IT-27:84–96, 1981.
- [12] J.C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Trans. Automat. Contr.*, 37:332–341, 1992.
- [13] J.C. Spall. Adaptive stochastic approximation by the simultaneous perturbation method. In *Proceedings of the 1998 IEEE CDC*, pages 3872 – 3879, 1998.

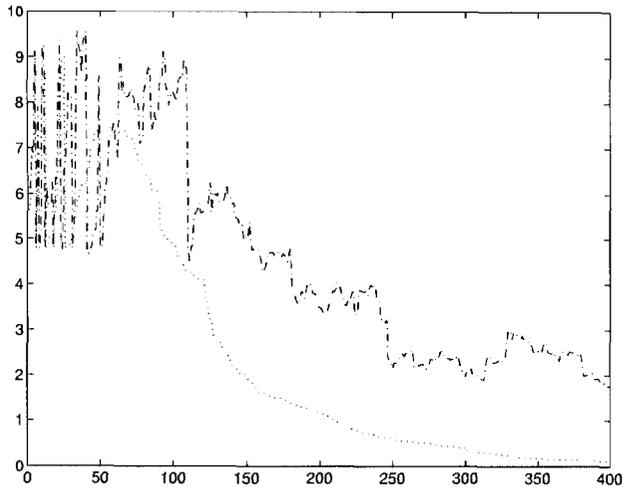


Figure 1: The value of $L(\hat{\theta}_k)$ for $\sigma = 0.5$ (dotted line) and for $\sigma = 5$ (dash-dot line)

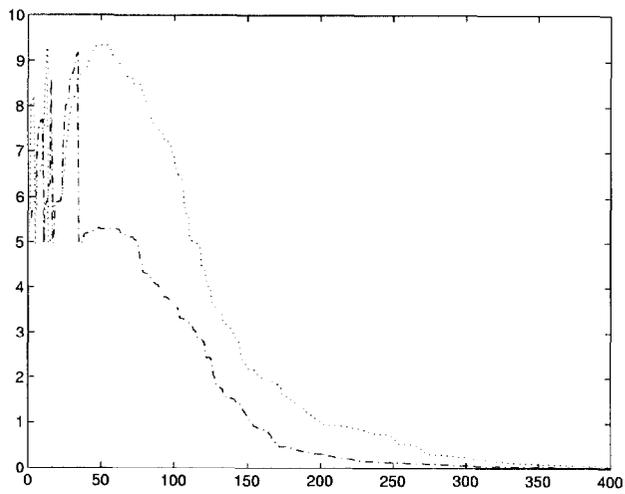


Figure 2: The value of $L(\hat{\theta}_k)$ for $\sigma = 0.05$ (dotted line) and for $\sigma = 0.5$ (dash-dot line)