Journal of Applied Sciences

ISSN 1812-5654

Journal of Applied Sciences 6 (11): 2506-2509, 2006
ISSN 1812-5654

A Strong Convergent Stochastic Algorithm for the Global Solution of Autonomous Linear Equation in n-spaces

A.C. Okoroafor
Department of Mathematics, Abia State University

Abstract: In this study, we proposed the method of response surface exploration for approximating the global equilibrium of the autonomous differential equation

$$\frac{du(t)}{dt} + Au(t) = 0 \qquad u(0) = u_0, u \in R^n, t \in R$$

It is shown that the gradient type stochastic approximation sequence arising from this method converges strongly to $u^* \in R^n$, where the initial value problem attains its global equilibrium.

Key words: Autonomous linear differential equation, monotone operator, equilibrium point stochastic algorithm

INTRODUCTION

It is well known that for any continuous linear operator A in a Banach Space X, the ordinary differential equation

$$\frac{du}{dt} + Au = 0 \qquad u(0) = u_0 \tag{1}$$

is solvable uniquely. At the point where

$$\frac{du}{dt} = 0 \tag{2}$$

the trajectory U(t) through U(0) approaches a single point (isolated point), say, c as t tends to infinity. c is in fact the equilibrium point of the system. An early fundamental result in the existence of global solution of autonomous differential equation, due to (Martin, 1970) states that the initial value problem (1) is solvable uniquely on $[0, \infty)$ if A is a continuous accretive operator.

Examples of how such initial value problems (1) arise are found in models involving either the heat or wave or the Schrodinger equation.

Considerable efforts have been devoted to developing constructive techniques for the determination of the equilibrium of the system (1) (Chidume, 1997).

Closely related to this problem in R^n is the solution of systems of linear equations

$$\sum a_{ij} u_j = b_i \tag{3}$$

Several significant problems where (3) arise are found in models involving autonomous differential equations in n-Spaces (Ladde *et al*., 1992)

Observe that a real $n \times n$ matrix $A = \{a_{ij}\}$ defines a linear operator on R^n so that if $(A\,u, u) \geq 0$ for any $u \in R^n$ then $f(u) = 1/2(Au,u)-bu$ is a convex, real valued function, which satisfies the growth condition $f(u) \to \infty$ as $\|u\| \to \infty$. It follows that f assumes a minimum

$$\left\langle u^* : \frac{\partial f(u^*)}{\partial u} = b \right\rangle$$

which coincides with the equilibrium point of the system

$$\frac{du}{dt} + Au = b, \quad u(0) = u_0$$

and also with the solution of (3). More over any minimizing sequence of f converges to

$$\left\langle u^* : \frac{\partial f(u^*)}{\partial u} = b \right\rangle$$

Hence the problem of determining the equilibrium point c of (1) when $A = \{a_{ij}\}$ and b = 0, reduces to that of locating the point u^* at which f attains its minimum.

Iterative methods for the solution of (3) have been studied extensively by various authors (Blum, 1972).

The purpose of this study to present a stochastic sequence that converges strongly to the solution (1). This is a modification of the method used by (Okoroafor and Ekere, 1999) in approximating the orbit and attractor of a dynamical system.

The advantages of this algorithm are not only its speed of convergence, compared with some other known methods, but also its precision and its ability to overcome the problem posed by sparcity in A.

PRELIMINARIES

Our R^n has the usual Euclidean norm $\|x\|^2 = x'x$ and inner product

$$(x, y) = \sum_{j=1}^{n} x_j y_j$$

where x' denotes the transpose of $x \in R^n$.

For a fixed $a \in R^n$ and random vectors $z_1, z_2 \in R^n$, $\|z\|$, $\langle z_1, z_2\rangle$, $\langle a, z_1\rangle$ are random variables in the usual sense.

Moreover, we associate with each random vector $z \in R^n$ the expectation operator E such that E_z defined by the requirement that $E\langle a, z_1\rangle = \langle a, Ez\rangle$ if $E\|z\| < \infty$.

For any $x \in R^n$ let

$$f(x) = \frac{1}{2}(Ax, x) - (b, x)$$

Since f is differentiable at x, then, by Taylor's expansion, there exists x_c on the line segment between x and $\hat{x}$ such that

$$f(x) - f(\hat{x}) = \left(\frac{\partial f(\hat{x})}{\partial x}, x - \hat{x}\right) + \frac{1}{2}(x - \hat{x})\, H(x_c)(x - \hat{x}) \qquad (4)$$

where H(x) is the Hessian of at x_c, so that for $D(f) = \{x \in R^n: f(x) < \infty\} \neq \phi$ if we set

$$y(x_j) = f(x^k + t_j) - f(x^k),$$

$x^k \in D(f)$ for a fixed k and $j = 1, ..., m$, $n + 2 < m < \frac{1}{2} n(n+1)$ where $t_j = (t_{1j}, ..., t_{nj}) \in R^n$ then (1) is identifiable with:

$$y(x_j) = \left(\frac{\partial f(x^k)}{\partial x}, t_j\right) + \frac{1}{2}\sum_{i=1}^{n}\sum_{i'=1}^{n} t_{ij} t_{i'j} \frac{\partial f(x^k)}{\partial x_i \partial x_{i'}} + e(x_j) \qquad (5)$$

$y(x_j)$, for each j is the outcome of an observation corresponding to the trial point $x_j \in E^n$ so that $y(x_1),..., y(x_n)$ are real-valued independent observable random variables performed on $x_1, x_2, ..., x_n$ for a fixed x^k.

$$e(x_1), e(x_2), ..., e(x_m)$$

are non-observable random errors.

We can show in the foregoing that if $t_j \in R^n$ is chosen such that (Pazman, 1987)

$$\sum_{j=1}^{m} t_{ij} = 0 \text{ and } \frac{1}{m}\sum_{j=1} t_{ij}^2 = 1 \quad i = 1, 2, ..., n \qquad (6)$$

then

Theorem 1

Let $\{e(x_j)\}$ be a sequence of identically distributed random variables satisfying $Ee(x_j) = 0$ for each j and $x_1,..., x_m$ are chosen in the neighbourhood of x^k such that $t_j = x_j - x^k$ for a fixed k, then, the relationship between $y(x_j)$ and t_j for $j = 1, ..., m$ is adequately represented by

$$y(x_j) = \left\langle \frac{\partial f(x^k)}{\partial x}, t_j \right\rangle + e(x_j),$$

Proof

Assume

$$y(x_j) = \left\langle \frac{\partial f(x^k)}{\partial x}, t_j \right\rangle + e(x_j),$$

is adequate to represent the relationship between $y(x_j)$ and t_j. Then the least square estimates of $\frac{\partial f(\dot{x}^k)}{\partial x}$ is

$$d(x^k) = M^{-1}\sum_{j=1}^{m} t_j y(x_j), \qquad (7)$$

where $M = \sum_{j=1}^{m} t_j t_j'$ so that

$$Ed(x^k) = M^{-1}\sum_{j=1}^{m} t_j\, Ey(x_j) = \frac{\partial f(x^k)}{\partial x}$$

Assume the contrary, then by Taylor's expansion

$$y(x_j) = \left(\frac{\partial f(x^k)}{\partial x}, t_j\right) + \frac{1}{2}\sum_{i=1}^{n}\sum_{i'=1}^{n} t_{ij} t_{i'j} \frac{\partial f(x^k)}{\partial x_i \partial x_{i'}} + e(x_j)$$

So that

$$Ed(x^k) = \frac{\partial f(x^k)}{\partial x} + \frac{1}{2}M^{-1}\sum_{j=1}^{m}\sum_{i=1}^{n}\sum_{i'=1}^{n} t_{ij}t'_{ij}\frac{\partial^2 f(x^k)}{\partial x_i \partial x'_i}$$

$$= \frac{\partial f(x^k)}{\partial x} + \frac{1}{2}M^{-1}\sum_{i=1}^{n}\sum_{i'=1}^{n}\frac{\partial^2 f(x^k)}{\partial x_i \partial x'_i}\sum_{j=1}^{m} t_{ij}^2 t'_{ij}$$

by Fubini's theorem

$$= \frac{\partial f(x^k)}{\partial x}$$ by Eq. (6) and (Box and Wilson, 1951)

Hence under Eq. (6) we can estimate $\frac{\partial f(x^k)}{\partial x}$ with least square error under the assumption that f is a linear function at x^k. Hence the result.

THE ITERATIVE SCHEME

Let

$$d^k = d(x^k)$$

and $\frac{\partial f(x^k)}{\partial x} = f^k$ we assume further the so called Gauss-Markov condition:

$$E\, e(x_j)e(x_i) = \delta^2\sigma ij,\ i.j = 1,\dots,\ n$$

where σ^2 is usually unknown and $0 < \sigma^2 < \infty$.
Thus for each k, d^k is a random vector and { d^k } is a stochastic sequence.
It follows that the sequence $\{x^k\}$ defined by $\rho^{2k} < \infty$.

$$\varphi(x^{\cdot}) = x^k - \rho^k d^k \tag{8}$$

is a sequence of independent random variables and

$$E\varphi(x^k) = x^k - \rho^k f^k$$

so that

$$E\left\|\varphi(x^k) - E\varphi(x^k)\right\| = \rho^k E\left\|d^k - f^k\right\| = 0$$

and

$$E\left\|\varphi(x^k) - E\varphi(x^k)\right\|^2 = \rho^{2k} E\left\|d^k - f^k\right\|^2 \tag{9}$$

$$= M^{-1}\sigma^2\rho^{2k} < \infty \text{ for each k}$$

CONVERGENCE THEOREM

Theorem 2

Suppose f: $R^n \to R$ is defined as in where $x' Ax \geq 0$ for any $x \in R^n$ and $\{\hat{\rho}^k\}$ is a real sequence satisfying

$$\sum\left(\frac{\rho^k}{k}\right)^2 < \infty$$

Then the sequence $\{x^k\}_{k=0}^{\infty}$ generated by $x^0 \in R^n$, $\varphi(x^k) = x^k - \rho^k d^k$ converges with probability one to the unique minimum point x* of f a.s.

Proof

Let $T^k = \rho^k\|d^k\text{-}f^k\|$ so that $\{T^k\}$ is a sequence of independent random variable such that
$Et^k = 0$ for each k
Then by Eq. 9

$$\sum_{k=1}^{\infty} E\frac{|T^k|}{k^2} = M^{-1}\sigma^2\sum\frac{\rho^{2k}}{k^2} < \infty$$

Hence by the strong law of large numbers (Whittle, 1976)
$T^k \to 0$ a.s. and then $\|x^k\text{-}\rho^k d^k\text{-}(x^k\text{-}\rho^k f^k)\| \to 0$ a.s.
so that $\{\varphi(x^k)\}$ and $\{E\varphi(x^k)\}$ converge to the same limit point x*
But $\{E\varphi(x^k)\}$ is a minimizing sequence of f,

so that if $\left\{x^*: \frac{\partial f(x^*)}{\partial x} = 0\right\}$ then $x^* = E\,\varphi(x^*)$ and

$$\lim_{k\to\infty}\left\|\varphi(x^k) - x^*\right\| = 0$$

Thus the result.

Remark; Theorem 2 remains valid if

ρ^k is replaced by ρ^{*k}
Such that

$f(x^k\text{-}\rho^{*k} d^k) = \min f(x^k\text{-}\rho d^k)$

so that,

$\rho^{*2k} < r < \infty$

Hence

$$\sum_{k=1}^{\infty}\frac{\rho^{*2k}}{k^2} < \sum_{k=1}^{\infty}\frac{r}{k^2} = r\sum_{k=1}^{\infty}\frac{1}{k^2} < \infty$$

Thus for the sequence satisfying

$$f(x^k - \rho^{*k} d^k) = \min f(x^k - \rho d^k) \quad , \quad \{\rho^k\}_{k=1}^{\infty}$$

generated by

$$x^0 \in R^n, \ \varphi(x^k) = x^k - \rho^k d^k$$

converges to the unique minimum point

x^* of f a.s.

The choice of

$$\rho^{*k}$$

for each k optimizes the sequence

$$\{\rho^k\}$$

Consider the iterative scheme

$$x^{k+1} = \varphi(x^k) = x^k - \rho^k d^k \tag{10}$$

It is easy to see that the trajectories of (1) can be associated with the asymptotic behaviour of Eq. 10.

So that

Corollary

Beginning at some initial condition x^0, the sequence of paths produced by $\{\varphi(x^k)\}_{k=1}^{\infty}$, through its definition by successive iteration on the function φ is associated with the trajectory of the initial value problem (1) and converges strongly to the global solution of the autonomous linear differential equation.

Starting at a given initial condition x^0 a search for the solution x* is conducted along the line $x^{k+1} = x^k - \rho^k d^k$ as follows:

- Compute as $\frac{\partial}{\partial x} f(x^k) \underset{-}{\sim} d^k$ in (7)
- Compute $\hat{\rho}^k$ such that
  $$f(x^k - \hat{\rho}^k d^k) = \min_{\rho} f(x^k - \rho d^k)$$
- $x^{k+1} = x^k - \hat{\rho}^k d^k$ is $\|x^{k+1} - x^k\| \le \delta, \ \delta \ge 0$.

  Yes. Then $x^{k+1} = x^*$

  If no: set k = k + 1 and return to 1.

Here, the search for the global solution x* is along the minimum error gradient direction which is supposed to lead to the solution faster.

REFERENCES

Blum, E.K., 1972. Numerical analysis and Computation. Addison-Wesley.

Box, G.E. and K.B. Wilson, 1951. On the experimental attainment of optimum conditions. J. Roy. Stat. Soc., 3: 1-45.

Chidume, C.E., 1997. Iterative approximation of fixed point of Lipschitzian strictly pseudocontractive mapping. Proc. Am. Math. Soc. 99' pp: 283-288.

Martin, R.H., 1970. A global existence theorem for autonomous differential equations in banach spaces. Proc. Am. Math. Soc., 26: 307-314.

Okoroafor, A.C. and Ekere, 1999. A Stochastic Approximation Method for the Attractor of a Dynamical system. In Direction of Mathematics. Ekhaguere, G.O.S. and O.O. Ugbebor (Eds.), pp: 131-142.

Pazman, A., 1987. Fundamentals of Optimal Experimental Designs. D. Reidel Publishing company.

Whittle, P., 1976. Probability. John Willey and Sons.