Linearization and Nonlinear Stochastic Differential Equations with Locally Lipschitz Condition

Omid S. Fard

School of Applied Mathematics
Damghan University of Basic Sciences, Damghan, Iran
Osfard@dubs.ac.ir

Abstract

In this paper, a new numerical method to solve nonlinear stochastic differential equations (SDEs) is presented. It is based upon approximating the drift and diffusion of the SDE by using basic linear functions. In addition, we prove that the solution of the obtained linear SDE is an approximate solution of the nonlinear SDE in the mean square sense.

Mathematics Subject Classification: 65C30, 65C20, 65L20

Keywords: stochastic differential equations, numerical method, linearization

1 Introduction

In this paper, we study the numerical solution of the Itô SDE

\[ dy(t) = f(y(t))dt + g(y(t))dw(t), \quad t_0 \leq t \leq T \]  

with initial value \( y(t_0) = y_0 \). Here \( y(t) \in \mathbb{R}^m \) for each \( t, t_0 \leq t \leq T \) and \( w(t) \) is a d-dimensional Brownian motion, thus \( f : \mathbb{R}^m \rightarrow \mathbb{R}^m \) and \( g : \mathbb{R}^m \rightarrow \mathbb{R}^{m \times d} \), some condition on \( f \) and \( g \) will be defined later. In addition, we assume that the initial condition is chosen independently of the Wiener measure driving the equation, and that all moments of \( y_0 \) are finite. As many SDE systems do not have analytical solutions, it is necessary to solve these systems numerically. The simplest stochastic numerical method is the Euler-Maruyama method:

\[ y_{n+1} = y_n + h_n f(y_n) + \Delta w_n g(y_n) \]
where \( h_n = t_{n+1} - t_n \) and \( \triangle w_n^j = w^j(t_{n+1}) - w^j(t_n) \sim N(0, h) \) for \( j = 1, 2, \ldots, d \). The Euler-Maruyama method has strong order of convergence 0.5 and it converges to the Itô solution (For more details, see [2], [5]). Recently, Higham, Mao, and Stuart have developed this method in continuous form under less restrictive conditions (see [6]). The Euler-Maruyama method is the simplest strong Taylor approximation (see [2], chapter 10). The Taylor approximation is to use truncated forms of the stochastic Taylor series formula (for example see [7], [8], and [2]). The equation (1) can be written in integral form:

\[
y(t) = y_0 + \int_{t_0}^{t} f(y(s))ds + \int_{t_0}^{t} g(y(s))dw(s), \quad t_0 \leq t \leq T. \tag{3}
\]

If in (3) \( f \) and \( g \) are expanded in an Itô stochastic Taylor series about \( y_0 \), then a representation as an infinite series for \( y(t) \) is obtained. Truncating the stochastic at a particular point yield an expression for \( y(t) \) of a certain order (see [2], chapters 10 and 14). Numerical methods can thus be defined for computing approximations to \( y(t) \), although these become difficult to implement if higher derivatives are included. The Milstein scheme (see [2], chapter 10) consists of the first few terms of the stochastic Taylor series:

\[
y_{n+1} = y_n + h_n f(y_n) + \triangle w_n g(y_n) + \frac{1}{2} g(y_n) g'(y_n) (\triangle w_n)^2 - h_n. \tag{4}
\]

Moreover, it converges with strong order 1 as long as \( E(y_0)^2 < \infty \), \( f \) and \( g \) are twice continuously differentiable, and that satisfy a uniform Lipschitz condition (for more details see [2], chapter 10). Obtaining higher order stochastic numerical method from the Taylor series is straight-forward in derivation but involve considerable complexities in implementation not only in the approximation of the higher order stochastic integral but also in the calculation of the derivatives of the coefficient functions \( f \) and \( g \). In this paper, we want to amend the mentioned problem. Our primary objective is to prepare the conditions, which on basis of them we can obtain suitable piecewise continuous linear approximations for functions \( f \) and \( g \) in (1) by using basic linear functions and then we obtain a linear stochastic differential equation corresponding to (1). To get the desired results we explain some theorems and definitions in section 2. In section 3, we consider the conditions, which we are relied the \( p \)th moments of the solution of (1), \( p > 0 \) are finite and bounded and on basis of the conditions, we obtain an upper bound for the solution of (1). In section 4, we define the notion of a closed ball, which in it we show the obtained solution by using of our offered method is converged to the exact solution of (1) in the mean square sense.
2 Existence and Uniqueness of solution of SDE

Let \( X = (\Omega, F, P) \) be a complete probability space with a \( \{F_t\}_{t \geq 0} \) filtration satisfying the usual conditions. A nonlinear stochastic differential equation of Itô type is defined in the following form:

\[
dy(t) = f(y(t))dt + g(y(t))dw(t), \quad t_0 \leq t \leq T
\]

with initial value \( y(t_0) = y_0 \). Here \( y(t) \in \mathbb{R}^m, f : \mathbb{R}^m \rightarrow \mathbb{R}^{m \times d}, g : \mathbb{R}^m \rightarrow \mathbb{R}^m \), \( w(t) \) is a \( d \)-dimensional Brownian motion on the space \( X \) and \( y_0 \) is an \( F_t \)-measurable \( \mathbb{R}^m \)-valued random variable such that \( E |y_0|^p < \infty, p > 0 \).

Equation (5) is equivalent to the following Itô stochastic integral equation:

\[
y(t) = y_0 + \int_{t_0}^{t} f(y(s))ds + \int_{t_0}^{t} g(y(s))dw(s), \quad t_0 \leq t \leq T
\]

At first, we give the definition of solution (5), (see [1]).

Definition 2.1 An \( \mathbb{R}^m \)-valued stochastic process \( \{y(t)\}_{t_0 \leq t \leq T} \) is called a solution of equation (5) if it has the following properties:

(i) \( \{y(t)\} \) is continuous and \( F_t \)-adapted;
(ii) \( \{f(y(t))\} \in L^1(\mathbb{R}^m, [t_0, T]) \) and \( \{g(y(t))\} \in L^2(\mathbb{R}^m, [t_0, T]) \)
(iii) equation (6) holds for every \( t \in [t_0, T] \) with probability 1.

A solution is said to be unique if any other solution \( \{\tilde{y}(t)\} \) is indistinguishable from \( \{y(t)\} \), that is

\[
P\{y(t) = \tilde{y}(t), \forall t \in [t_0, T]\} = 1.
\]

Here, and throughout the paper, \( | \cdot | \) is used to denote both the Euclidean vector norm and the Frobenius (or trace) matrix norm. Moreover \( M^p([t_0, T], \mathbb{R}^m) \) is the family of processes \( \{f(t)\}_{t_0 \leq t \leq T} \) in \( L^p([t_0, T], \mathbb{R}^m) \) such that \( E \int_{t_0}^{T} |f(t)|^p dt < \infty \).

The following theorem prepares the conditions that guarantee the existence and uniqueness of solution to equation (5).

**Theorem 2.1** Assume that:

(i) (Local Lipschitz condition) for any integer \( n \geq 1 \), there exists a positive constant \( k_n \) such that for all \( t \in [t_0, T] \) and all \( x, y \in \mathbb{R}^m \) and \( x, y \leq n \)

\[
|f(x(t)) - f(y(t))|^2 \leq k_n |x(t) - y(t)|^2
\]

(ii) (Monotone condition) there exists a positive constant \( k \) such that for all \( y \in \mathbb{R}^m \) and for all \( t \in [t_0, T] \),

\[
y^T f(y(t)) + \frac{1}{2} |g(y(t))|^2 \leq k(1 + |y(t)|^2)
\]
then there exists a unique solution \( y(t) \) to equation (5) in \( M^2(\mathbb{R}^m, [t_0, T]) \).

**Proof.** See the proof of theorems 3.4 and 3.5 in [1], pages 57 and 58.

### 3 \( L^p \)-Estimates

The results of this section play the important role in our proposed method. In this section, we assume that \( y(t_0), t_0 \leq t \leq T \) is the unique solution to equation (5) with initial value \( y(t_0) = y_0 \) and we shall investigate the \( p \)th moment of the solution.

**Theorem 3.1** Let \( p \geq 2 \) and \( y_0 \in L^p(\Omega, \mathbb{R}^m) \). Assume that there exists a constant \( \alpha > 0 \) such that for all \( y(t) \in \mathbb{R}^m \) and \( t \in [t_0, T] \),

\[
y^T f(y(t)) + \frac{p-1}{2} \ | g(y(t)) |^2 \leq \alpha (1 + \ | y \ |^2)
\]

Then

\[
E \ | y(t) |^p \leq 2^{\frac{p}{2}} (1 + E \ | y_0 |^p) e^{\alpha (t-t_0)}
\]

**Proof.** See the proof of theorem 4.1 in [1], page 59.

Now, we turn to consider the case of \( 0 < p < 2 \). This is rather easy if we note that the Holder inequality implies

\[
E \ | y(t) |^p \leq (E \ | y(t) |^2)^{\frac{p}{2}}
\]

In other words, the estimate for \( E \ | y(t) |^p \) can be done via the estimate for the second moment. Therefore, we have the following corollary.

**Corollary 3.2** Let \( 0 < p < 2 \) and \( y_0 \in L^2(\Omega, \mathbb{R}^m) \). Assume that there exists a constant \( \alpha > 0 \) such that for all \( y(t) \in \mathbb{R}^m \) and \( t \in [t_0, T] \),

\[
y^T f(y(t)) + \frac{1}{2} \ | g(y(t)) |^2 \leq \alpha (1 + \ | y(t) |^2)
\]

Then

\[
E \ | y(t) |^p \leq (1 + E \ | y_0 |^2)^{\frac{p}{2}} e^{\alpha (t-t_0)}, \quad t \in [t_0, T]
\]

### 4 Continuous Piecewise Linear Approximation

In this section, we try to find a way to be able to approximate the function \( f \) and \( g \) in (5) by two continuous piecewise linear functions. One of the important tools in the approximation topic is the existence of a closed ball. We need a closed ball to deal with an approximation topic. Since \( y(\cdot) \) is a random
variables, in general, it can not be uniformly bounded. Hence, to overcome this problem, we do the following process.

As we showed in sections 2 and 3, if $E \mid y_0 \mid^p < \infty$ for all $0 < p < \infty$, there exists a constant and finite $C_1 = C_1(p, y_0) \in \mathbb{R}$, such that if $y(t)$ is a unique solution to equation (5) and the solution belongs to $M^2([t_0, T], \mathbb{R}^m)$ under the conditions of theorem (3.1) and corollary (3.2), then

$$E(\sup_{t_0 < t < T} | y(t) |^p) \leq C_1 e^{p\alpha(T - t_0)} < \infty,$$

where $\alpha$ is the constant in theorem (3.1) and corollary (3.2). In particular, for $p = 2$, we have

$$E(\sup_{t_0 < t < T} | y(t) |^2) \leq C_1 e^{\beta} < \infty, \forall t \in [t_0, T], \quad (14)$$

where $\beta = 2\alpha(T - t_0)$.

Therefore, by using the Chebyshev inequality, we can conclude that, for any $\sigma > 0$, there exists a real positive constant $M_1 = M_1(\sigma)$ and such that:

$$P(A_\sigma) = P\{\omega : \sup_{t_0 < t < T} | y(t) | \leq M_1\} \geq 1 - \sigma.$$

Hence, we can conclude that if we choose $\sigma > 0$ sufficiently small or, equivalently, $M_1$ sufficiently large, then for every $t \in [t_0, T]$ and almost every $\omega \in A_\sigma$, we have:

$$| y(t) | \leq M_1.$$

In other words, if $\sigma > 0$ be chosen sufficiently small, we can choose a closed ball $B = B(M_1)$ such that we expect the values of $y(.)$ to be in $B$, with probability more than $1 - \sigma$, and we hope the values of $y(.)$ to be out of $B$, with probability less than $\sigma$.

Therefore, by considering $\sigma > 0$ sufficiently small, there exists a closed ball $B \in \mathbb{R}^m$ such that

$$\forall t \in [t_0, T] \text{ and almost } \omega \in A_\sigma, \quad y(t) \in B.$$

The following lemma plays an important role in this section.

**Lemma 4.1.** Let $f : N \rightarrow M$ be a function from the metric space $N$ to another metric space $M$. Let also, $f$ be uniformly continuous on $A$ in $D_f$ (domain of $f$), and $M$ be a complete space. Then, there exists a uniformly continuous function $\overline{f}$ on $\overline{A}$ (closure of $A$), such that the restriction of $\overline{f}$ on $A$ is equal to the restriction of $f$ on $A$ (i.e., $\overline{f}|_A \equiv f|_A$).

**Proof:** See [9, chapter 4].

Now, we can discuss about the approximation. First, We define the closed set $B_c \in \mathbb{R}^m$ such that

$$B_1 = (D_f \cap D_g \cap B),$$
where $D_f$ and $D_g$ are the domain of $f$ and $g$, respectively. Condition (7) and (8) imply that $f$ and $g$ in (5) are two uniformly continuous function on $B_1$ and so, lemma 4.1 holds for $f$ and $g$. In other words, if $\overline{f}$ and $\overline{g}$ are the unique extension of $f$ and $g$, respectively, on $B_1$, then $\overline{f}$ and $\overline{g}$ are uniformly continuous on $B_1$.

The next theorem implies that if we consider functions $\overline{f}$ and $\overline{g}$ as two deterministic functions, then for given $\epsilon > 0$, we can find uniformly approximation of these functions by the linear functions as $a_1^k y(t) + b_1^k$ and $(a_2^k y(t) + b_2^k, ..., a_d^k y(t) + b_d^k)$, respectively for $\overline{f}$ and $\overline{g}$, where $a_i^k, a_{2i}^k \in \mathbb{R}^{m \times m}$ and $b_i^k, b_{2i}^k \in \mathbb{R}^m$, for $i = 1, 2, ..., d$. In other words, for every $\epsilon > 0$ given, we can divide the set $B_1$ to the finite disjoint cells $I_1, I_2, ..., I_n$, $n = n(\epsilon)$, such that by definition of functions $h_{1n(z)}$ and $h_{2n(z)}$ as follows:

$$h_{1n(z)}(y(t)) = a_1^k y(t) + b_1^k$$
and

$$h_{2n(z)}(y(t)) = (a_2^k y(t) + b_2^k, ..., a_d^k y(t) + b_d^k)$$

We have

$$| h_{1n(\epsilon)} - \overline{f} |_{B_1} \vee | h_{2n(\epsilon)} - \overline{g} |_{B_1} < \epsilon,$$
where

$$| h_{1n(\epsilon)} - \overline{f} | = \sup\{ | h_{1n(\epsilon)}(y(t)) - \overline{f}(y(t)) | : y(t) \in B_1, t \in [t_0, T] \}$$

and

$$| h_{2n(\epsilon)} - \overline{g} | = \sup\{ | h_{2n(\epsilon)}(y(t)) - \overline{g}(y(t)) | : y(t) \in B_1, t \in [t_0, T] \}$$

**Theorem 4.1.** Let $f$ be a continuous function whose domain $D_f$ is a compact cell in $\mathbb{R}^p$ and whose values belong to $\mathbb{R}^q$. Then $f$ can be uniformly approximated on $D_f$ by continuous piecewise linear function.

**Proof.** See the proof of theorem 24.5 in [4], page 169.

Based on above discussions, $\overline{f}(.)$ and $\overline{g}(.)$ satisfy the conditions of theorem 4.1, therefore, since $f(.)$ and $g(.)$ are the restrictions of $\overline{f}(.)$ and $\overline{g}(.)$, respectively, on $B_1 = B \cap D_f \cap D_g$, we can conclude that:

$$| h_{1n(\epsilon)} - f |_{B_1} \vee | h_{2n(\epsilon)} - g |_{B_1} < \epsilon,$$

(15)

Hence, if we replace $h_{1n(\epsilon)}$ and $h_{2n(\epsilon)}$ respectively with $f$ and $g$ in equation (5), we have

$$dy(t) \cong h_{1n(\epsilon)}(y(t))dt + h_{2n(\epsilon)}(y(t))dw(t),$$
and by definition $y_{n(\epsilon)}(t)$ as the approximation solution to equation (5), we obtain

$$dy_{n(\epsilon)}(t) = h_{1n(\epsilon)}(y(t))dt + h_{2n(\epsilon)}(y_{n(\epsilon)}(t))dw(t) \text{ for all } \omega \in A_\sigma,$$

(16)

and we define $y_{n(\epsilon)} \equiv 0$, whenever $\omega \in A_C^\sigma$.

Remark. As we have seen before, if we define $A_\sigma = \{ \omega : \sup_{t_0 \leq t \leq T} |y(t)| \leq M_1 \}$, then $P(A_\sigma) > 1 - \sigma$ and so $P(A_C^\sigma) < \sigma$. On the other hand, we have $E |X| = \int_{\Omega} |X| dP(\omega)$ and we can write

$$E |X| = \int_{\Omega} |X| dP(\omega) = E_{A_\sigma} |X| + E_{A_C^\sigma} |X|,$$

where $E_C |X| = \int_C |X| dP(\omega)$, for every $C \subseteq \Omega$.

We recall that if $f \in L^1(P)$, then for every $\epsilon > 0$, there exists $\sigma > 0$ such that $\int_C |f| dP < \epsilon$, whenever $P(C) < \sigma$.

We can prove that the proposed method converges strongly if $f$ and $g$ are locally Lipschitz and the exact and numerical solution have bounded $p$th moments for some $p > 2$. The bounded moment assumption will not, of course, hold in general as solutions to the SDE may explode in a finite time.

Assumption 4.2. For each $R \geq 1$, there exists a positive constant $k_R$, depending only on $R$, such that for all $t \in [t_0, T]$ and all $x, y \in \mathbb{R}^m$ with $|x| \vee |y| \leq R$,

$$|f(x(t)) - f(y(t))|^2 \vee |g(x(t)) - g(y(t))|^2 \leq k_R |x(t) - y(t)|^2,$$

(17)

for some $p > 0$ there is a constant $A$ such that,

$$E(\sup_{t_0 \leq t \leq T} |y(t)|^p) \vee E(\sup_{t_0 \leq t \leq T} |y_{n(\epsilon)}(t)|^p) \leq A, \hspace{1em} \forall \epsilon > 0.$$

(18)

Note that inequality (17) is a local Lipschitz assumption. From the Mean Value Theorem, any $f$ and $g$ in $C^1$ will satisfy (17). The equality (18) states that the $p$th moments of the exact and numerical solution are bounded for some $p > 0$. It is clearly the inequality (18) obtains from Theorem 3.1 and the Holder inequality.

We now prove that assumption 4.2 is sufficient to ensure strong convergence of the proposed method.

Theorem 4.3. Under assumption 4.2, for any $T > 0$ the solution of the proposed method, $y_{n(\epsilon)}(t)$, for each $\epsilon > 0$, satisfies

$$\lim_{\epsilon \to 0} E(\sup_{t_0 \leq t \leq T} |y_{n(\epsilon)}(t) - y(t)|^2) = 0.$$
Proof. We first show that
\[ \lim_{\epsilon \to 0} E_{A,\sigma} \left( \sup_{t_0 \leq t \leq T} | y_n(\epsilon)(t) - y(t) |^2 \right) = 0, \]
for every \( 0 < \sigma < 1 \).
Let \( \sigma > 0 \) be arbitrary. For given \( \gamma > 0 \), put
\[ \epsilon_0 = \sqrt{\frac{\gamma}{8T(T - t_0 - 4) \exp(4kR(T + 4))}} \]
and, let \( \epsilon < \epsilon_0 \) be given. We divide the set \( B_1 \) to finite disjoint cells \( I_1, I_2, ..., I_n \), \( n = n(\epsilon) \) such that (15) holds.
We define
\[ \tau_R = \inf \{ t \geq 0 : | y_n(\epsilon)(t) | \geq R \}, \rho_R = \inf \{ t \geq 0 : | y(t) | \geq R \}, \theta_R = \tau_R \wedge \rho_R \]
and
\[ e(t) = y_n(\epsilon)(t) - y(t). \]
Recall the Young inequality: for \( r^{-1} + q^{-1} = 1 \),
\[ ab \leq \frac{\delta}{r} a^r + \frac{1}{q \delta^{\frac{1}{r}}} \cdot \forall a, b, \delta > 0, \]
we thus have for any \( \delta > 0 \)
\[ E_{A,\sigma} \left( \sup_{t_0 \leq t \leq T} | e(t) |^2 \right) = E_{A,\sigma} \left( \sup_{t_0 \leq t < T} | e(t) |^2 1_{\tau_R \leq T, \rho_R > T} \right) \]
\[ + E_{A,\sigma} \left( \sup_{t_0 \leq t < T} | e(t) |^2 \cdot 1_{\tau_R > T \lor \rho_R < T} \right) \]
\[ \leq E_{A,\sigma} \left( \sup_{t_0 \leq t < T} | e(t \wedge \theta_R) |^2 \cdot 1_{\tau_R > T} \right) + \frac{2\delta}{p} E_{A,\sigma} \left( \sup_{t_0 \leq t < T} | e(t) |^p \right) \]
\[ + \frac{1 - \frac{2}{p}}{\delta^{\frac{1}{p}}} P(\tau_R \leq T \lor \rho_R \leq T). \]
Now
\[ P(\tau_R \leq T) = E_{A,\sigma} \left( 1_{\tau_R \leq T} | y_n(\epsilon)(\tau_R) |^p \right) \leq \frac{1}{R^p} \left( \sup_{t_0 \leq t < T} | y_n(\epsilon)(\tau_R) | \right) \leq \frac{A}{R^p}, \]
using (18). A similar result can be derived for \( \rho_R \), so that
\[ P(\tau_R \leq T \lor \rho_R \leq T) \leq P(\tau_R \leq T) + P(\rho_R \leq T) \leq \frac{2A}{R^p}. \]
Using these bounds along with

\[ E_{A_n} (\sup_{t_0 < t < T} | e(t) |^2) \leq \]

\[ E_{A_n} (\sup_{t_0 < t < T} | y_n(t \land \theta_R) - y(t \land \theta_R) |^2) + \frac{2^{p+1} \delta A}{p} + \frac{(p - 2)2A}{p \delta^{\frac{p-2}{2}} R^p} \]

Now, we bound the first term on the right-hand side of (19) using an approach similar to a finite time convergence proof for the globally Lipschitz case. Using

\[ y(t \land \theta_R) = y_0 + \int_{t_0}^{t \land \theta_R} f(y(s))ds + \int_{t_0}^{t \land \theta_R} g(y(s))dw(s), \]

and

\[ y_n(t \land \theta_R) = y_0 + \int_{t_0}^{t \land \theta_R} h_{1n}(y_n(s))ds + \int_{t_0}^{t \land \theta_R} h_{2n}(y_n(s))dw(s), \]

from (19) and Cauchy-Schwartz, we have

\[ |y_n(t \land \theta_R) - y(t \land \theta_R)|^2 \leq \int_{t_0}^{t \land \theta_R} h_{1n}(y_n(s)) - f(y(s))ds + \int_{t_0}^{t \land \theta_R} h_{2n}(y_n(s))ds \]

\[ -g(y(s))dw(s) |^2 \]

\[ \leq 2(T \int_{t_0}^{t \land \theta_R} | h_{1n}(y_n(s)) - f(y(s)) |^2 ds + \int_{t_0}^{t \land \theta_R} | h_{2n}(y_n(s)) - g(y(s)) |^2 dw(s) |^2 \]

Note that we can define \( \epsilon_1 = f(y_n(s)) - h_{1n}(y_n(t)), \epsilon_2 = g(y_n(s)) - h_{2n}(y_n(t)) \) and by (16) we have \( |\epsilon_1(t) \lor \epsilon_2(t)| < \epsilon \leq \epsilon_0 \) for all \( t \in [t_0, T] \), therefore

\[ |y_n(t \land \theta_R) - y(t \land \theta_R)|^2 \leq 2(T \int_{t_0}^{t \land \theta_R} | f(y_n(s)) - f(y(s)) - \epsilon_1(s) |^2 ds \]

\[ + | \int_{t_0}^{t \land \theta_R} g(y_n(s)) - g(y(s)) - \epsilon_2(s) |^2 dw(s) |^2 \]

\[ \leq 2(2T \int_{t_0}^{t \land \theta_R} | f(y_n(s)) - f(y(s)) |^2 ds + 2T \int_{t_0}^{t \land \theta_R} | \epsilon_1(s) |^2 ds \]

\[ + 2(| \int_{t_0}^{t \land \theta_R} (g(y_n(s)) - g(y(s)))dw(s) - \int_{t_0}^{t \land \theta_R} \epsilon_2(s) |^2 ds \]

\[ \leq (2Tk_R \int_{t_0}^{t \land \theta_R} | y_n(s) - g(y(s)) |^2 ds + 2T(t \land \theta_R - t_0) \epsilon^2 \]

\[ + 2(2 | \int_{t_0}^{t \land \theta_R} (g(y_n(s)) - g(y(s)))dw(s) |^2 + 2 | \int_{t_0}^{t \land \theta_R} \epsilon_2(s)dw(s) |^2 \]

So, from Doob’s martingale inequality [1], we have for any \( \tau \leq T \)

\[ E_{A_n} (\sup_{t_0 < t < \tau} | y_n(t \land \theta_R) - y(t \land \theta_R) |^2) \leq 4Tk_R \int_{t_0}^{\tau \land \theta_R} E_{A_n} (\sup_{t_0 < s < \tau} | y_n(s) - y(s) |^2)ds \]

\[ + 4T(t \land \theta_R - t_0)^2 + 16k_R \int_{t_0}^{\tau \land \theta_R} E_{A_n} (\sup_{t_0 < r < s} | y_n(r) - y(r) |^2)ds + 16T(t \land \theta_R - t_0) \epsilon^2 \]

\[ + 4\epsilon^2(T - t_0 + 4) + 4k_R(T + 4) \int_{t_0}^{\tau \land \theta_R} E_{A_n} (\sup_{t_0 < r < s} | y_n(r) - y(r) |^2)ds \]
Applying the Gronwall inequality, we obtain

\[ E_{A_\sigma}(\sup_{t_0 < t < \tau} | y_n(t) - y(t) |^2) \leq 4e^2T(T - t_0 + 4)e^{4kR(T + 4)}. \]

Inserting this into (19) gives

\[ E_{A_\sigma}(\sup_{t_0 < t < T} | e(t) |^2) \leq 4e^2T(T - t_0 + 4)e^{4kR(T + 4)} + \frac{2^{p+1}\delta A}{p} + \frac{(p - 2)2A}{p^{\delta - 2} R^p}. \]

For any \( \gamma > 0 \), we can choose \( \delta \) so that \( \frac{2^{p+1}\delta A}{p} < \frac{\gamma}{2} \), then choose \( R \) so that

\[ \frac{(p - 2)2A}{p^{\delta - 2} R^p} < \frac{\gamma}{2} \]

therefore,

\[ E_{A_\sigma}(\sup_{t_0 < t < T} | e(t) |^2) \leq 4e^2T(T - t_0 + 4)e^{4kR(T + 4)} + \frac{\gamma}{2} < \gamma, \]

and since \( \gamma > 0 \) is arbitrary, we obtain

\[ \lim_{\epsilon \to 0} E_{A_\sigma}(\sup_{t_0 < t < T} | e(t) |^2) = 0. \]

Now, to prove \( \lim_{\epsilon \to 0} E(\sup_{t_0 \leq t \leq T} | y_n(t) - y(t) |^2) = 0 \), again assume that \( \gamma > 0 \) is given.

Since \( \sup_{t_0 \leq t \leq T} | y(t) |^2 < \infty \), there is some \( \sigma > 0 \) such that

\[ \int_C \sup_{t_0 \leq t \leq T} | y(t) |^2 dP(\omega) < \frac{\gamma}{2}, \]

for every \( C \subseteq \Omega \) with \( P(C) < \sigma \).

Specially,

\[ \int_{\bar{A}_\sigma^C} \sup_{t_0 \leq t \leq T} | y(t) |^2 dP(\omega) < \frac{\gamma}{2}. \]

On the other hand, by the first part of the proof

\[ \lim_{\epsilon \to 0} E_{A_\sigma}(\sup_{t_0 \leq t \leq T} | y_n(t) - y(t) |^2) = 0. \]

So, there is some \( \epsilon_1 = \epsilon_1(\sigma) > 0 \) such that \( E_{A_\sigma}(\sup_{t_0 \leq t \leq T} | y_n(t) - y(t) |^2) < \frac{\gamma}{2} \), for each \( 0 < \epsilon < \epsilon_1 \).

Then we get

\[ E\left(\sup_{t_0 \leq t \leq T} | y_n(t) - y(t) |^2\right) = E_{A_\sigma}\left(\sup_{t_0 \leq t \leq T} | y_n(t) - y(t) |^2\right) + E_{A_\sigma}\left(\sup_{t_0 \leq t \leq T} | y_n(t) - y(t) |^2\right) \leq E_{A_\sigma}\left(\sup_{t_0 \leq t \leq T} | y_n(t) - y(t) |^2\right) + E_{A_\sigma}\left(\sup_{t_0 \leq t \leq T} | y(t) |^2\right) < \frac{\gamma}{2} + \frac{\gamma}{2} = \gamma, \]

and the proof is complete.
References


Received: June 1, 2007