CHAPTER 5
CHAPTER 5

Application to Stochastic Differential Equations (SDE)

5.1 Introduction and Summary

In this Chapter we have discussed two areas. First one is regarding first order stochastic differential equation and second one is regarding Kalman-Bucy filter. For this we need to add something in literature for prerequisites.

For the first one:

We consider the processes which satisfy the first order stochastic differential equation.

\[ dX(t) + K(t)X(t) = dB(t) \]  

(5.1.1)

where \( K(t) \) is a non-random function of \( t \), others \( X(t) \) and \( B(t) \) depend on sample points. Here \( B(t) \) is a standard one dimensional Brownian motion. This comes in the following way.

The increment of the process from \( t - \Delta \) to \( t \) time is

\[ X(t) - X(t - \Delta) \propto X(t - \Delta), \text{ state at time } t - \Delta \]

and also \( \propto K(t - \Delta) \)

i.e., rate of increment depends on the present state and on some constraints which are time dependent.
So combining these two we must have

\[ X(t) - X(t - \Delta) = K(t - \Delta)X(t - \Delta)\Delta + \text{normal error} \]

where \( K(t - \Delta)X(t - \Delta) \) is velocity at \( t - \Delta \).

Then

\[ X(t) - X(t - \Delta) = K(t - \Delta)X(t - \Delta)\Delta + (B_t - B_{t-\Delta}) \]

Such things are available in literature (Karlin, S. and Taylor, H. (1975, 1981)).

So the above equation (5.1.1) gives the following integral equation

\[ \Rightarrow X_t = X_0 + \int_0^t K(t - s)X_{t-s}ds + \int_0^t dB_s \quad (5.1.2) \]

It is to be noted that it is not defined through usual differentiation because Brownian paths are non-differentiable with probability one. It is in the sense of Ito.

There is estimator of diffusion parameter which is consistent for large samples and thus one knows the function \( K(t) \). So for the time being we assume that \( K(.) \) is known. Here the solutions are available in literature (Rao (1999)). But our target is to look this solution by the solution obtained from the power series method for solving integral equation. It is really a new way to look sample paths in terms of our solution as done in Chapter 2. But here we are to modify this because of the Brownian part.

Basic tools used here are Markov inequality, Borel-Cantelli lemma and Cauchy-Schwartz inequality.

For second one:
In section (5.3) we consider state and measurement processes of a Kalman-Bucy linear system.

Here p-dimensional 'state' process $X$ and a q-dimentional 'measurement' process $Z$ are assumed to be related by the stochastic differential system

$$dX(t) = A(t)X(t)dt + B(t)u(t)dt + dW(t),$$
$$dZ(t) = C(t)X(t) + dV(t)$$

where $W$ and $V$ are independent p-dimentional and q-dimentional Wiener processes for $0 \leq t \leq T$, $u$ is known input function and $A$, $B$, $C$ are known non-random time dependent matrices of suitble dimensions and $X(0)$ is independent of $W$ and $V$. Also the processes $\{Z(s), 0 \leq s \leq t\}$ is observable whereas $\{X(s), 0 \leq s \leq t\}$ is not. Such models are used in control system.

In real applied problem, often the additive trends are present in the state and measurement processes. So the state space is given with extra additional additive trend terms.

$$dX(t) = f(t)dt + A(t)X(t)dt + B(t)u(t)dt + dW(t)$$

and the observed process instead of $Z$ is given by

$$dY(t) = g(t)dt + dZ(t)$$

For first one, a new method for obtaining solution to stochastic differential equation has been suggested. The power series method has been altered in order to accomodate Brownian part.

For the second, an estimate of trend in Kalman filter is indicated using power series solution and that function estimate belong to $L^2[0, T]$.  

76
5.2 To solve first order SDE by power series method

From our Chapter 2 we see that the solution of

\[ X_t = X_0 + \int_0^t K(t-s)X_{t-s}ds \]

is given by

\[ X_t^{(0)} = a_0\phi_0(t) + a_1\phi_1(t) + a_2\phi_2(t) + \cdots \tag{5.2.1} \]

provided we assume non-random function \( X_t \).

Here \( \phi_i(t) \) depends on \( X_0(t) \) and kernel function \( K(t) \). But if we assume \( X_0 \) a random function then the solution (5.2.1) also depends on sample point.

Now in order to take account of Brownian part let us define

\[ X_t^{(1)} = X_0 + \int_0^t K(t-s)[a_0\phi_0(t-s) + a_1\phi_1(t-s) + a_2\phi_2(t-s) + \cdots]ds + (B_t - B_0) \]

Then

\[ X_t^{(2)} = X_0 + \int_0^t K(t-s)X_t^{(1)}ds + \int_0^t dB_s \tag{5.2.2} \]

Proceeding like this, we have

\[ X_t^{(n+1)} = X_0 + \int_0^t K(t-s)X_t^{(n)}ds + \int_0^t dB_s \]

Taking sufficiently large \( n \), we have approximate solution to (5.1.2).
5.2.1 Limit of iterative paths from power series solution

In this section as indicated in (5.1) we want to show that $X_i^{(n)} \rightarrow X_i$ in probability and this limit paths are solution of equation (5.1.1). In order to show this let us proceed as in the following

\[
(X_i^{(n+1)} - X_i^{(n)})^2 = \left[ \int_0^t K(t-s) (X_i^{(n)} - X_i^{(n-1)}) \, ds \right]^2, \quad n \geq 1 \quad (5.2.1.1)
\]

Let us write

\[
D_i^{(n)} = E \max_{0 \leq s \leq t} (X_i^{(n)} - X_i^{(n-1)})^2
\]

Also we assume $|K(\cdot)| \leq K$ i.e., bounded kernel function Taking expectation of the maximum in (5.2.1.1) over 0 to $T$ we get

\[
D_i^{(n+1)} \leq KE \left( \int_0^T |X_i^{(n)} - X_i^{(n-1)}| \, ds \right)^2
\]

\[
\leq KE \int_0^T (X_i^{(n)} - X_i^{(n-1)})^2 \, ds \text{ by Cauchy-Schwartz inequality}
\]

and repeating this step again and again we have

\[
\leq KE \int_0^T D_i^{(n)} \, ds
\]

\[
\leq \frac{(KT)^n}{n!} c
\]

where $c = E \int_0^T (X_i^{(0)})^2 \, ds$

Now by Chebyshev’s inequality,

\[
\Pr \left[ \max_{0 \leq s \leq T} |X_i^{(n)} - X_i^{(n-1)}| > 2^{-n} \right] \leq \frac{22^n c (KT)^n}{n!} \quad (5.2.1.2)
\]

78
we see that r.h.s. of (5.2.1.2) is summable. So by Borel-Cantelli lemma

$$\Pr\left[ \max_{0 \leq s \leq t} |X_{t-s}^{(n)} - X_{t-s}^{(n-1)}| > 2^{-n}, \text{ for infinitely many } n \right] = 0$$

So beyond this null set

$$X_t^{(0)} + \sum_{n=0}^{\infty} (X_t^{(n+1)} - X_t^{(n)}) = \lim_{n \to \infty} X_t^{(n)}$$

converges uniformly for $0 \leq t \leq T$ to $X_t$ (say).

Then by Fatou's lemma

$$\int_0^T E\left( X_t^{(n)} - X_t \right)^2 dt \leq \lim_{m \to \infty} \int_0^T E\left( X_t^{(n)} - X_t^{(m)} \right)^2 dt$$

Therefore, by triangle inequality for $L^2$-norms we have

$$\left[ \int_0^T E\left( X_t^{(n)} - X_t \right)^2 dt \right]^{1/2} \leq \sum_{\gamma=n+1}^{m} \left[ \int_0^T E\left( X_t^{(\gamma-1)} - X_t^{(\gamma)} \right)^2 dt \right]^{1/2} \leq \sum_{\gamma=n+1}^{\infty} \left( D_T^{(\gamma)} \right)^{1/2}$$

Now $\sum_{\gamma=n+1}^{\infty} D_T^{(\gamma)} \to 0$ We have

$$\int_0^T E\left( X_t^{(n)} - X_t \right)^2 dt \to 0 \Rightarrow X_t^{(n)} \to X_t \text{ in } L^2 \text{ norm}$$

$$X_t^{(1)} = X_0 + \int_0^t K(t-s) [a_0 \phi_0(t-s) + a_1 \phi_1(t-s) + \cdots] ds + (B_t - B_0)$$

and

$$X_t^{(0)} = a_0 \phi_0(t-s) + a_1 \phi_1(t-s) + \cdots$$

Therefore,

$$\sup (X_t^{(1)} - X_t^{(0)})^2 = \sup \left( X_0 + \int_0^t K(t-s)X_{t-s}^{(0)} ds + (B_t - B_0) - X_t^{(0)} \right)^2$$

79
Thus we have shown that

\[
E \sup_t (X_t^{(1)} - X_t^{(0)})^2 = E \sup_t (B_t - B_0)^2 = c_2
\]

and r.h.s. of above expression can be made arbitrarily small because \( X_s \) is limit of \( X_s^{(n)} \) by our construction method.

Thus we have proved the following theorem

**Theorem 5.2.1.** The iterative sequence using power series method as in (5.2.2) converges to \( X_t \) in \( L^2 \) norm and becomes solution to equation (5.1.1).

**Proof:** Follows from previous derivations.

### 5.3 An Application to Kalman-Bucy Filter

In real applied problem, often the additive trends are present in the state and measurement processes. So the state space is given with extra additional additive trend terms.

\[
dX(t) = f(t)dt + A(t)X(t)dt + B(t)u(t)dt + dW(t)
\]

and the observed process instead of \( Z \) is given by

\[
dY(t) = g(t)dt + dZ(t)
\]

An important problem considered in literature is to estimate \( f \) and \( g \) and to use them in Kalman filter \( \hat{X}(t) \). Here \( n \) observations \( \{Y_i(t), 0 \leq t \leq T\}, 1 \leq \)
\[ i \leq n \] are taken from the earlier processes \( Y \). For the case \( p = q = 1 \) the equations for the filter are

\[
d\hat{X}(t) = \left[ f(t) + A(t)\hat{X}(t) + B(t)u(t) \right] dt + D(t)d\gamma(t)
\]

\[
d\gamma(t) = dY(t) - \left[ g(t) + C(t)\hat{X}(t) \right] dt
\]

where \( f, g, A, B, C \) and \( u \) are smooth functions and \( C(t) \) is non zero on \([0, t]\). The process \( \gamma \) is a standard Wiener process and the Kalman gain \( D(t) = C(t)P(t) \) is free from \( f \) and \( g \). Here \( P \) is the unique solution of differential equation

\[
P'(t) = 2A(t)P(t) - C^2(t)P(t) + 1,
\]

with \( P(0) = \text{var}(X(0)) \). Solving this system of equation (see Davis 1977) for \( \hat{X}(t) \) and substituting the representation for \( Y \) becomes

\[
Y(t) = \int_0^t \left[ h(s) + U(s) \right] ds + \gamma(t) \tag{5.3.1}
\]

where

\[
h(t) = g(t) + C(t)\int_0^t \psi(t, s) \left[ f(s) - D(s)g(s) \right] ds \tag{5.3.2}
\]

and \( \frac{\partial}{\partial t} \psi(t, s) \) depends on \( A(t), D(t), C(t) \) as in Davi's (1977).

Furthermore, \( U \) is given by

\[
U(t) = C(t)\{ \psi(t, 0)m + \int_0^t \psi(t, s) [B(s)u(s)ds + D(s)dY(s)] \}
\]

where, \( m = E(X(0)) \). If \( f \equiv 0 \) then (5.3.2) reduces to

\[
h(t) = g(t) + \int_0^t \Gamma(t, s)g(s)ds \tag{5.3.3}
\]

where \( \Gamma(t, s) = -C(t)\psi(t, s)D(s) \).

Above assumption is important in order to identify \( g(s) \).
Then equation (5.3.3) is linear Volterra integral equation of second kind. From representation (5.3.1) it follows

\[ l(h) = \int_0^T \pi(s) dY(s) - \frac{1}{2} \int_0^T \pi^2(s) ds \]

where \( l(h) \) is log likelihood and

\[ \pi(s) = h(s) + U(s), \ 0 \leq s \leq T \]

Then according to observations of \( n \) many paths the log-likelihood is given by

\[ l(h) = \sum_{i=1}^n l_i(h) \]

where

\[ l_i(h) = \int_0^T [h(s) + U_i(s)] dY_i(s) - \frac{1}{2} \int_0^T [h(s) + U_i(s)]^2 ds \]

Then in the following we describe the estimation of \( h(s) \).

Now \( l(h) = 0 \)

\[ \Rightarrow \int_0^T [h(s) + U_i(s)] dY_i(s) = \frac{1}{2} \int_0^T [h(s) + U_i(s)]^2 ds \]

\[ \Rightarrow \int_0^T h(s)dY_i(s) + \int_0^T U_i(s)dY_i(s) = \frac{1}{2} \left[ \int_0^T h^2(s)ds + \int_0^T U^2_i(s)ds + 2 \int_0^T h(s)U_i(s)ds \right] \]

\[ \Rightarrow \int_0^T h(s)dY_i(s) - \int_0^T h(s)U_i(s)ds = \frac{1}{2} \int_0^T h^2(s)ds + \frac{1}{2} \int_0^T U^2_i(s)ds - \int_0^T U_i(s)dY_i(s) \]

From this solution for \( \hat{h}(s) \) is obtained because the above system of equations are approximately in the form

\[ A \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_k \end{pmatrix} + B \begin{pmatrix} h_1^2 \\ h_2^2 \\ \vdots \\ h_k^2 \end{pmatrix} = C \quad (5.3.4) \]
This system (5.3.4) of equation can be solved by iteration method. Then above system comes as we approximate the following expression, \( \int_0^T h(s)dY_t(s) \) in the following way.

Let us take \( 0 = s_0 < s_1 < \cdots < s_k = T \) s.t. \( s_j - s_{j-1} = h \) for \( j = 1, 2, \cdots, k \) and then we have,

\[
\sum_{j=1}^{k} h(s_j) [Y_t(s_j) - Y_t(s_{j-1})] \simeq \int_0^T h(s)Y_t(s)
\]

This is a way of solving \( h(s) \) but we have also due to others

Theorem: Above solution can be chosen to be consistent.


So in previous section we have described a method of getting an estimate of \( h(.) \). So based on observations of \( Y_t(t), 0 \leq t \leq T \) we have an estimate \( \hat{h}(t) \). This estimate is used to find out an estimate of \( g(.) \), where \( g \) should be obtained from

\[
\hat{h}(t) = \hat{g}(t) + \int_0^t \Gamma(t, s)\hat{g}(s)ds
\]

or in otherwords,

\[
\hat{g}(t) = \hat{h}(t) - \int_0^t \Gamma(t, s)\hat{g}(s)ds \tag{5.3.5}
\]

In this section solution of \( \hat{g}(.) \) was obtained in terms of consistent \( \hat{h}(.) \) function and kernel function \( \Gamma(t, s) \).

Also it is to be noted that Davis (1977, page 125) ensured the existence of unique solution \( \hat{g} \in L^2[0, T] \).

In this situations we prove the following theorem.

Theorem 5.3.1. Our solution \( \hat{g}(t) \) as described by power series method is a \( L^2[0, T] \) function.
Proof: Davis (1977, page 125) proved that given \( \hat{h} \in L^2[0,T] \) there exists a unique solution to equation (5.3.5).

But the way we obtained solution by power series method is the unique solution. This implies that the solution obtained by power series method is also \( L^2[0,T] \) function.

Remark: The solution obtained by power series technique is unique and the solution obtained is \( L^2 \) consistent, because Davis showed existence and uniqueness.

5.4 Concluding Remarks

In this Chapter we have considered two areas related stochastic differential equations.

Firstly, we have used the power series solution of ordinary integral equation, then have accommodated Brownian part iteratively and finally the limiting paths thus obtained are solution to SDE.

In future non-linear stochastic integral equation are to be investigated in the same line.

Secondly, we have suggested that the power series solution gives the estimates of trend function of Kalman Bucy system in dimension 1. But this is yet to be generalised in higher dimension.