Statistics of Solutions to Test Models for SPEKF

Chuan Lu

Information and Computing Science

复旦大學

June 2017

Outline

Introduction

Signals from nature can be modeled by Langevin equation:

Langevin Equation

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} = -\gamma(t)u(t) + \mathrm{i}\omega u(t) + \sigma\dot{W}(t) + f(t),$$

where $\dot{W}(t)$ is a Brownian motion, and $f(t)$ is the external force.

A major difficulty in accurate filtering of noisy signals with many degrees of freedom is model error; signal from nature is processed through incomplete physical models, as well as parameterized to inadequate numerical resolution.

B. Gershgorin, *et, al* proposed the Stochastic Parameterization Extended Kalman Filter (SPEKF) to cope with model errors.

Introduction

Test Model

$$\begin{cases} \dfrac{du(t)}{dt} = (-\gamma(t) + \mathrm{i}\omega)u(t) + b(t) + f(t) + \sigma W(t), \\[2ex] \dfrac{db(t)}{dt} = (-\gamma_b + \mathrm{i}\omega_b)(b(t) - \hat{b}) + \sigma_b W_b(t), \\[2ex] \dfrac{d\gamma(t)}{dt} = -d_\gamma(\gamma(t) - \hat{\gamma}) + \sigma_\gamma W_\gamma(t) \end{cases}$$

The initial values are complex random variables, with their first-order and second-order statistics known.

Introduction

Solution

With knowledge of ODEs, solution of the SDE set is

$$\begin{cases} b(t) = \hat{b} + (b_0 - \hat{b})e^{\lambda_b(t-t_0)} + \sigma_b \int_{t_0}^{t} e^{\lambda_b(t-s)} dW_b(s) \\ \gamma(t) = \hat{\gamma} + (\gamma_0 - \hat{\gamma})e^{-d_\gamma(t-t_0)} + \sigma_\gamma \int_{t_0}^{t} e^{-d_\gamma(t-s)} dW_\gamma(s) \\ u(t) = e^{-J(t_0,t)+\hat{\lambda}(t-t_0)} u_0 + \int_{t_0}^{t} (b(s)+f(s))e^{-J(s,t)+\hat{\lambda}(s-t_0)} ds \\ \qquad + \sigma \int_{t_0}^{t} e^{-J(s,t)+\hat{\lambda}(s-t_0)} dW(s) \end{cases}$$

with $\lambda_b = -\gamma_b + \mathrm{i}\omega_b$, $\hat{\lambda} = -\hat{\gamma} + \mathrm{i}\omega$, $J(s,t) = \int_s^t (\gamma(s') - \hat{\gamma}) ds'$.

Introduction

Itô Isometry

$\forall f \in \mathscr{V}(S,T)$, B_t is a standard Brownian motion,

$$\mathrm{E}\left[\left(\int_S^T f(t,\omega)dB_t\right)^2\right] = \mathrm{E}\left[\int_S^T f^2(t,\omega)dt\right].$$

Linear property of Itô integration

$$(1) \quad \int_S^T f dB_t = \int_S^U f dB_t + \int_U^T f dB_t, \text{ a.e.}$$

$$(2) \quad \int_S^T (cf+g)dB_t = c\int_S^T f dB_t + \int_S^T g dB_t, \text{ a.e.}$$

$$(3) \quad \mathrm{E}\left[\int_S^T f dB_t\right] = 0.$$

Outline

Mean of $b(t)$, $\gamma(t)$

With property of Itô integration (3), the last term in $\mathrm{E}[b(t)]$ is 0. Thus we find

$$\mathrm{E}[b(t)] = \hat{b} + (\mathrm{E}[b_0] - \hat{b})e^{\lambda_b(t-t_0)}$$

$$\mathrm{E}[\gamma(t)] = \hat{\gamma} + (\mathrm{E}[\gamma_0] - \hat{\gamma})e^{-d_\gamma(t-t_0)}$$

Variance of $b(t)$, $\gamma(t)$

According to definition,

$$
\begin{aligned}
\mathrm{Var}(b(t)) &= \mathrm{E}\left[(b(t)-\mathrm{E}[b(t)])(b(t)-\mathrm{E}[b(t)])^*\right] \\
&= e^{-2\gamma_b(t-t_0)}\mathrm{Var}(b_0) + \mathrm{E}\left[\sigma_b^2 \int_{t_0}^{t} e^{\lambda_b(t-s)}dW_b(s)\left(\int_{t_0}^{t} e^{\lambda_b(t-s)}dW_b(s)\right)^*\right] \\
&= e^{-2\gamma_b(t-t_0)}\mathrm{Var}(b_0) + \sigma_b^2\mathrm{E}\left[\int_{t_0}^{t} e^{-2\gamma_b(t-s)}ds\right]
\end{aligned}
$$

The last step takes use of Itô isometry. Compute the integration in last term, we have

$$
\mathrm{Var}(b(t)) = e^{-2\gamma_b(t-t_0)}\mathrm{Var}(b_0) + \frac{\sigma_b^2}{2\gamma_b}(1-e^{-2\gamma_b(t-t_0)})
$$

$$
\mathrm{Var}(\gamma(t)) = e^{-2d_\gamma(t-t_0)}\mathrm{Var}(\gamma_0) + \frac{\sigma_\gamma^2}{2d_\gamma}(1-e^{-2d_\gamma(t-t_0)})
$$

Covariance of $b(t)$, $\gamma(t)$

According to definition,

$$\mathrm{Cov}(b(t), b(t)^*) = \mathrm{E}\left[(b(t) - \mathrm{E}[b(t)])(b(t)^* - \mathrm{E}[b(t)^*])\right]$$

$$= \mathrm{E}\left[(b_0 - \mathrm{E}[b_0])(b_0^* - \mathrm{E}[b_0^*])e^{2\lambda_b(t-t_0)}\right] + \sigma_b \mathrm{E}\left[(b_0 - \mathrm{E}[b_0])\int_{t_0}^{t} e^{\lambda_b(t-s)} dW_b(s)\right]$$

$$+ \sigma_b \mathrm{E}\left[(b_0^* - \mathrm{E}[b_0^*])e^{\lambda_b(t-t_0)}\int_{t_0}^{t} e^{\lambda_b(t-s)} dW_b(s)\right] + \sigma_b^2 \mathrm{E}\left[\left(\int_{t_0}^{t} e^{\lambda_b(t-s)} dW_b(s)\right)^2\right]$$

With property of Itô integration (3), the second and third term are both 0; with Itô isometry we know the last term is also 0. Accordingly,

$$\mathrm{Cov}(b(t), b(t)^*) = \mathrm{E}[(b_0 - \mathrm{E}[b_0])(b_0^* - \mathrm{E}[b_0^*])]e^{2\lambda_b(t-t_0)} = \mathrm{Cov}(b_0, b_0^*)e^{2\lambda_b(t-t_0)}$$

$$\mathrm{Cov}(b(t), \gamma(t)) = \mathrm{E}[(b(t) - \mathrm{E}[b(t)])(\gamma(t) - \mathrm{E}[\gamma(t)])] = \mathrm{Cov}(b_0, \gamma_0)e^{(\lambda_b - d_\gamma)(t-t_0)}$$

Outline

Mean of $u(t)$

Using the same properties, it is easy to find

$$\mathrm{E}[u(t)] = e^{\hat{\lambda}(t-t_0)}\mathrm{E}\left[e^{-J_0(t_0,t)}u_0\right] + \int_{t_0}^{t} e^{\hat{\lambda}(t-s)}\mathrm{E}\left[b(s)e^{-J(s,t)}\right]ds$$

$$+\sigma\int_{t_0}^{t} e^{\hat{\lambda}(t-s)}f(s)\mathrm{E}\left[e^{-J(s,t)}\right]ds.$$

We find it necessary to compute expectations of terms like

$$\mathrm{E}[ze^{bx}],$$

where z is a complex-valued Gaussian random variable and x is a real-valued Gaussian variable. We propose two lemmas here.

Mean of $u(t)$

Lemma (1)

$$E\left[ze^{ibx}\right] = (E[z] + ibCov(z,x))e^{ibE[x]-\frac{1}{2}b^2 Var(x)},$$

with z being a complex-valued Gaussian, and x a real-valued Gaussian.

Corollary (1)

Under the condition of Lemma 1,

$$E\left[ze^{bx}\right] = (E[z] + bCov(z,x))e^{bE[x]+\frac{1}{2}b^2 Var(x)}.$$

Proof of lemma 1 takes advantage of characteristic function of multivariable Gaussian distribution.

Proof

Let $z = y + \mathrm{i}w,\ y, w \in \mathbb{R}$. Denote $\mathbf{v} = (x, y, w)^\top$, then $\mathbf{v}$ satisfies the multivariable Gaussian distribution, with its characteristic function

$$\phi_{\mathbf{v}}(\mathbf{s}) = \exp(\mathrm{i}\mathbf{s}^\top \mathrm{E}[\mathbf{v}] - \frac{1}{2}\mathbf{s}^\top \Sigma \mathbf{s}).$$

Let $g(\mathbf{v})$ being the PDF of $\mathbf{v}$, then one knows from that char. func. being Fourier transform of PDF,

$$\phi_{\mathbf{v}}(\mathbf{s}) = \frac{1}{(2\pi)^3}\int e^{\mathrm{i}\mathbf{s}^\top \mathbf{v}} g(\mathbf{v}) d\mathbf{v}.`'$$

According to the differential property of Fourier transform (Proposition 2.10),

$$\frac{\partial \phi_{\mathbf{v}}(\mathbf{s})}{\partial s_2} = \frac{1}{(2\pi)^3}\int \mathrm{i}y_0 e^{\mathrm{i}\mathbf{s}^\top \mathbf{v}} g(\mathbf{v}) d\mathbf{v} = \mathrm{i}\mathrm{E}\left[y_0 e^{\mathrm{i}\mathbf{s}^\top \mathbf{v}}\right].$$

Let $\mathbf{v} = (b, 0, 0)^\top$,

$$\mathrm{E}[y_0 e^{ibx_0}] = -\mathrm{i}\frac{\partial \phi_{\mathbf{v}}(s)}{\partial s_2}\bigg|_{\mathbf{s}=(b,0,0)^\top}, \quad \mathrm{E}[w_0 e^{ibx_0}] = -\mathrm{i}\frac{\partial \phi_{\mathbf{v}}(s)}{\partial s_3}\bigg|_{\mathbf{s}=(b,0,0)^\top}$$

Proof

From PDF of multivariable Gaussian distribution, one knows

$$\frac{\partial \phi_{\mathbf{v}}(\mathbf{s})}{\partial s_2} = (\mathrm{iE}[y_0] - \mathrm{Var}(y_0)s_2 - \mathrm{Cov}(x_0, y_0)s_1 - \mathrm{Cov}(y_0, w_0)s_3)\phi_{\mathbf{v}}(s)$$

$$\frac{\partial \phi_{\mathbf{v}}(\mathbf{s})}{\partial s_3} = (\mathrm{iE}[w_0] - \mathrm{Var}(w_0)s_3 - \mathrm{Cov}(x_0, w_0)s_1 - \mathrm{Cov}(y_0, w_0)s_2)\phi_{\mathbf{v}}(s)$$

Compute the partial derivatives at $\mathbf{s} = (b, 0, 0)^\top$,

$$\mathrm{E}\left[y_0 e^{ibx_0}\right] = (\mathrm{E}[y_0] + \mathrm{iCov}(x_0, y_0)b)\exp(\mathrm{i}b\mathrm{E}[x_0] - \frac{1}{2}\mathrm{Var}(x_0)b^2)$$

$$\mathrm{E}\left[w_0 e^{ibx_0}\right] = (\mathrm{E}[w_0] + \mathrm{iCov}(x_0, w_0)b)\exp(\mathrm{i}b\mathrm{E}[x_0] - \frac{1}{2}\mathrm{Var}(x_0)b^2)$$

Then

$$\mathrm{E}\left[ze^{ibx}\right] = (\mathrm{E}[z] + \mathrm{i}b\mathrm{Cov}(z, x))e^{\mathrm{i}b\mathrm{E}[x] - \frac{1}{2}b^2\mathrm{Var}(x)}.$$

□

Mean of $u(t)$

Lemma (2)

$$E\left[zwe^{bx}\right] = [E[z]E[w] + Cov(z,w^*) + b(E[z]Cov(w,x)) + E[w]Cov(z,x) + b^2 Cov(z,x)Cov(w,x)]\, e^{bE[x]+\frac{b^2}{2}Var(x)}.$$

with z,w being complex-valued Gaussian, and x real-valued Gaussian.

The proof of this lemma is the same as Lemma 1.

Mean of $u(t)$

We now make use of Lemma 1 to obtain mean of $u(t)$.

$$
\begin{aligned}
\mathrm{E}[u(t)] &= e^{\hat{\lambda}(t-t_0)}(\mathrm{E}[u_0] - \mathrm{Cov}(u_0, J(t_0,t)))e^{-\mathrm{E}[J(t_0,t)]+\frac{1}{2}\mathrm{Var}(J(t_0,t))} \\
&\quad + \int_{t_0}^{t} e^{\hat{\lambda}(t-s)}(\hat{b} + e^{\lambda_b(s-t_0)}(\mathrm{E}[b_0] - \hat{b} - \mathrm{Cov}(b_0, J(s,t))))e^{-\mathrm{E}[J(s,t)]+\frac{1}{2}\mathrm{Var}(J(s,t))}ds \\
&\quad + \int_{t_0}^{t} e^{\hat{\lambda}(t-s)} f(s) e^{-\mathrm{E}[J(s,t)]+\frac{1}{2}\mathrm{Var}(J(s,t))}ds
\end{aligned}
$$

The terms $\mathrm{Cov}(u_0, J(s,t))$, $\mathrm{Cov}(b_0, J(s,t))$, $\mathrm{E}[J(s,t)]$ and $\mathrm{Var}(J(s,t))$ can be found using Itô isometry.

Variance of $u(t)$

Denote $u(t) = A + B + C$,

$$\begin{cases} A = e^{-J(t_0,t)+\hat{\lambda}(t-t_0)} u_0, \\ B = \int_{t_0}^{t} (b(s) + f(s)) e^{-J(s,t)+\hat{\lambda}(t-s)} ds, \\ C = \sigma \int_{t_0}^{t} e^{-J(s,t)+\hat{\lambda}(t-s)} dW(s). \end{cases}$$

By definition we find $\mathrm{Var}(u(t)) = \mathrm{E}\left[|u(t)|^2\right] - |\mathrm{E}[u(t)]|^2$, with

$$\mathrm{E}\left[|u(t)|^2\right] = \mathrm{E}\left[|A|^2\right] + \mathrm{E}\left[|B|^2\right] + \mathrm{E}\left[|C|^2\right] + 2\mathrm{Re}\{\mathrm{E}[A^*B]\}.$$

We can obtain $\mathrm{E}\left[|A|^2\right]$ by Lemma 2, and $\mathrm{E}\left[|B|^2\right]$ by Itô isometry. Noticing that

$$\mathrm{Cov}(J(s,t), J(r,t)) = \mathrm{Var}(J(s,t)) + \mathrm{Cov}(J(s,t), J(r,s)).$$

$\mathrm{E}\left[|C|^2\right]$ and $\mathrm{Re}\{\mathrm{E}[A^*B]\}$ can also be computed by Itô isometry and property of Itô integration.

Covariances

By definition,

$$\begin{aligned}
\mathrm{Cov}(u(t), u^*(t)) &= \mathrm{E}\left[u(t)^2\right] - \mathrm{E}[u(t)]^2 \\
\mathrm{Cov}(u(t), \gamma(t)) &= \mathrm{E}[u(t)(\gamma(t) - \hat{\gamma})] + \mathrm{E}[u(t)](\hat{\gamma} - \mathrm{E}[\gamma(t)]) \\
\mathrm{Cov}(u(t), b(t)) &= \mathrm{E}[u(t)b^*(t)] - \mathrm{E}[u(t)]\mathrm{E}[b(t)]^* \\
\mathrm{Cov}(u(t), b^*(t)) &= \mathrm{E}[u(t)b(t)] - \mathrm{E}[u(t)]\mathrm{E}[b(t)].
\end{aligned}$$

Each term can be obtained by Lemma 3 and Itô isometry.

Outline

Parameters

As a real-world example, we choose external forcing

$$f(t) = \frac{3}{2} e^{\frac{1}{10} it},$$

and parameters of the equation set are given

$$\begin{cases} d = 1.5, & d_\gamma = 0.01d \\ \sigma = 0.1549, & \omega = 1.78 \\ \sigma_\gamma = 5\sigma, & \gamma_b = 0.1d \\ \sigma_b = 5\sigma, & \omega_b = \omega \\ \hat{b} = 0, & \hat{\gamma} = 0 \end{cases}$$

We assume that initial values satisfy

$$\begin{cases} \mathrm{Re}(u_0), \mathrm{Im}(u_0) \sim \mathcal{N}(0, \frac{1}{\sqrt{2}}), \text{ i.i.d.} \\ \mathrm{Re}(b_0), \mathrm{Im}(b_0) \sim \mathcal{N}(0, \frac{1}{\sqrt{2}}), \text{ i.i.d.} \\ \gamma_0 \sim \mathcal{N}(0, 1) \end{cases}$$

Euler-Maruyama Scheme

Itô integration can be simulated by E-M scheme:

$$X_j = X_{j-1} + f(X_{j-1})\Delta t + g(X_{t-1})(W(\tau_j) - W(\tau_{j-1})),$$

with

$$W(\tau_j) - W(\tau_{j-1}) = \sum_{k=jR-R+1}^{jR} \mathrm{d}W_k,$$

and R being the step length of E-M scheme,

$$\mathrm{d}W = \sqrt{\Delta t} \times \mathtt{randn}().$$

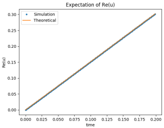
Revised Euler Scheme

Since we need to simulate a SDE set with three relevant equations, we choose the Revised Euler Scheme from numerical solution to ODE:

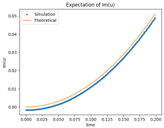
$$x_n = x_{n-1} + \Delta t \phi\left(t_{n-1} + \frac{\Delta t}{2},\ x_{n-1} + \frac{\Delta t}{2}\phi(t_{n-1},\ x_{n-1})\right),$$

it is a second-order Runge-Kutta scheme with a locally quadratic convergence.

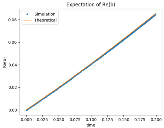
Simulate 10^6 times with $R = 1$, the results are as follows:



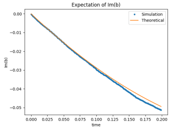
(a) E(Re(u))



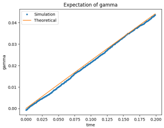
(b) E(Im(u))



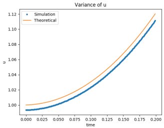
(c) E(Re(b))



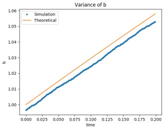
(d) E(Im(b))



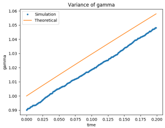
(e) E(γ)



(f) Var(u)



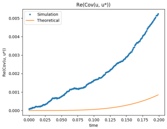
(g) Var(b)



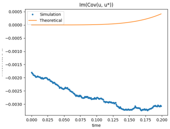
(h) Var(γ)

Figure: Simulation of Expectations and Variances, $n = 10^6$

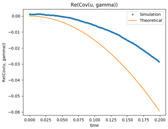
Result



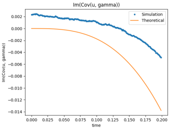
(a) $\mathrm{Re}(\mathrm{Cov}(u,u^*))$



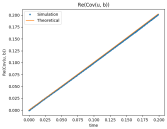
(b) $\mathrm{Im}(\mathrm{Cov}(u,u^*))$



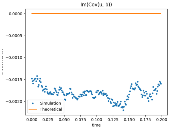
(c) $\mathrm{Re}(\mathrm{Cov}(u,\gamma))$



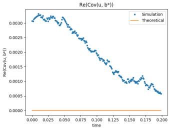
(d) $\mathrm{Im}(\mathrm{Cov}(u,\gamma))$



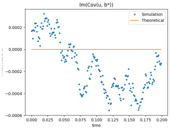
(e) $\mathrm{Re}(\mathrm{Cov}(u,b))$



(f) $\mathrm{Im}(\mathrm{Cov}(u,b))$



(g) $\mathrm{Re}(\mathrm{Cov}(u,b^*))$

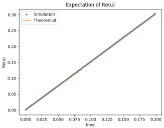


(h) $\mathrm{Im}(\mathrm{Cov}(u,b^*))$

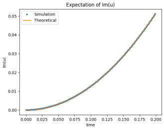
Figure: Simulation of Covariances, $n = 10^6$

Results

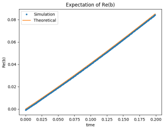
Simulate 10^7 times with $R = 1$, the results are as follows:



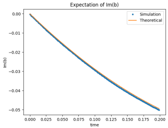
(a) E(Re(u))



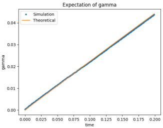
(b) E(Im(u))



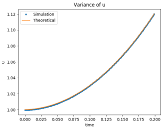
(c) E(Re(b))



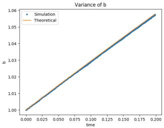
(d) E(Im(b))



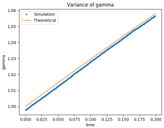
(e) E(γ)



(f) Var(u)

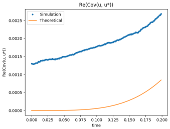


(g) Var(b)

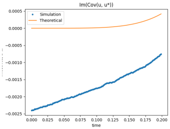


(h) Var(γ)

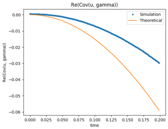
Figure: Simulation of Expectations and Variances, $n = 10^7$



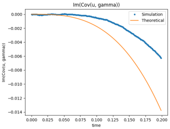
(a) $\mathrm{Re}(\mathrm{Cov}(u, u^*))$



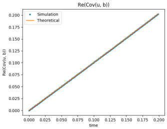
(b) $\mathrm{Im}(\mathrm{Cov}(u, u^*))$



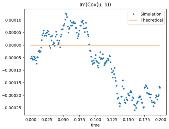
(c) $\mathrm{Re}(\mathrm{Cov}(u, \gamma))$



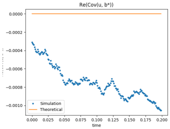
(d) $\mathrm{Im}(\mathrm{Cov}(u, \gamma))$



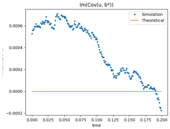
(e) $\mathrm{Re}(\mathrm{Cov}(u, b))$



(f) $\mathrm{Im}(\mathrm{Cov}(u, b))$



(g) $\mathrm{Re}(\mathrm{Cov}(u, b^*))$



(h) $\mathrm{Im}(\mathrm{Cov}(u, b^*))$

Figure: Simulation of Covariances, $n = 10^7$

Discussion

From the simulation results we find that the simulation of expectations fit the theoretical results satisfyingly.
For the results of Variances and Covariances,

- When $n = 10^6$, the error of variances is about $O(10^{-2}) \sim O(10^{-3})$, error of covariances is about $O(10^{-2}) \sim O(10^{-3})$.
- When $n = 10^7$, the error of $\mathrm{Var}(u),\ \mathrm{Var}(b)$ is about $O(10^{-5})$, and error of $\mathrm{Var}(\gamma)$ is about $O(10^{-3}) \sim O(10^{-4})$, error of $\mathrm{Cov}(u,u^*),\ \mathrm{Cov}(u,\gamma)$ is about $O(10^{-3})$, error of $\mathrm{Cov}(u,b),\ \mathrm{Cov}(u,b^*)$ is $O(10^{-4})$.

Discussion

One can find the simulation errors decrease with the growth of simulations. Thus one could believe that when $n \to \infty$, the simulation would converge to the theoretical result. In fact, error of simulation comes from that initial values are all random variables; according to Law of large numbers, the errors would converge to 0 when $n \to \infty$.

We can get third-order and fourth-order statistics of u, b, γ with the same methods, which could be written in a form of multiple integrals of statistics of initial values.

Thank you!