Lecture 3: Probability Distributions and Statistics of SDEs

Simo Särkkä

Aalto University
Tampere University of Technology
Lappeenranta University of Technology
Finland

November 8, 2012
Consider the **stochastic differential equation (SDE)**

\[ \text{d}x = f(x, t) \, \text{d}t + L(x, t) \, \text{d}\beta. \]

Each \( x(t) \) is **random variable**, and we denote its **probability density** with \( p(x, t) \).

The probability density is solution to a **partial differential equation** called **Fokker–Planck–Kolmogorov equation**.

The mean \( m(t) \) and covariance \( P(t) \) are solutions of certain **ordinary differential equations**.

For **LTI SDEs** we can also compute the **covariance function** of the solution \( C(\tau) = \text{E}[x(t) x(t + \tau)] \).
The probability density $p(x, t)$ of the solution of the SDE

$$\text{d}x = f(x, t) \, \text{d}t + L(x, t) \, \text{d}\beta,$$

solves the Fokker–Planck–Kolmogorov partial differential equation

$$\frac{\partial p(x, t)}{\partial t} = - \sum_i \frac{\partial}{\partial x_i} [f_i(x, t) \, p(x, t)]$$

$$+ \frac{1}{2} \sum_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \left\{ [L(x, t) \, Q \, L^T(x, t)]_{ij} \, p(x, t) \right\}.$$

- In physics literature it is called the Fokker–Planck equation.
- In stochastics it is the forward Kolmogorov equation.
FPK Example: Diffusion equation

Brownian motion can be defined as solution to the SDE

\[ dx = d\beta. \]

If we set the diffusion constant of the Brownian motion to be \( q = 2D \), then the FPK reduces to the diffusion equation

\[ \frac{\partial p}{\partial t} = D \frac{\partial^2 p}{\partial x^2}. \]
Let $\phi(x)$ be an arbitrary twice differentiable function.

The Itô differential of $\phi(x(t))$ is, by the Itô formula, given as follows:

$$d\phi = \sum_i \frac{\partial \phi}{\partial x_i} f_i(x, t) \, dt + \sum_i \frac{\partial \phi}{\partial x_i} [L(x, t) \, d\beta]_i$$

$$+ \frac{1}{2} \sum_{ij} \left( \frac{\partial^2 \phi}{\partial x_i \partial x_j} \right) [L(x, t) \, L^T(x, t)]_{ij} \, dt.$$ 

Taking expectations and formally dividing by $dt$ gives the following equation, which we will transform into FPK:

$$\frac{d \mathbb{E}[\phi]}{dt} = \sum_i \mathbb{E} \left[ \frac{\partial \phi}{\partial x_i} f_i(x, t) \right]$$

$$+ \frac{1}{2} \sum_{ij} \mathbb{E} \left[ \left( \frac{\partial^2 \phi}{\partial x_i \partial x_j} \right) [L(x, t) \, L^T(x, t)]_{ij} \right].$$
The left hand side can now be written as follows:

\[
\frac{dE[\phi]}{dt} = \frac{d}{dt} \int \phi(x) p(x, t) \, dx = \int \phi(x) \frac{\partial p(x, t)}{\partial t} \, dx.
\]

Recall the multidimensional integration by parts formula

\[
\int_C \frac{\partial u(x)}{\partial x_i} v(x) \, dx = \int_{\partial C} u(x) v(x) n_i \, dS - \int_C u(x) \frac{\partial v(x)}{\partial x_i} \, dx.
\]

In this case, the boundary terms vanish and thus we have

\[
\int \frac{\partial u(x)}{\partial x_i} v(x) \, dx = - \int u(x) \frac{\partial v(x)}{\partial x_i} \, dx.
\]
Currently, our equation looks like this:

\[ \int \phi(x) \frac{\partial p(x, t)}{\partial t} \, dx = \sum_i \mathbb{E} \left[ \frac{\partial \phi}{\partial x_i} f_i(x, t) \right] + \frac{1}{2} \sum_{ij} \mathbb{E} \left[ \left( \frac{\partial^2 \phi}{\partial x_i \partial x_j} \right) [L(x, t) Q L^T(x, t)]_{ij} \right]. \]

For the first term on the right, we get via integration by parts:

\[
\mathbb{E} \left[ \frac{\partial \phi}{\partial x_i} f_i(x, t) \right] = \int \frac{\partial \phi}{\partial x_i} f_i(x, t) \, p(x, t) \, dx = - \int \phi(x) \frac{\partial}{\partial x_i} \left[ f_i(x, t) \, p(x, t) \right] \, dx
\]

We now have only one term left.
For the remaining term we use integration by parts twice, which gives

\[
E \left[ \left( \frac{\partial^2 \phi}{\partial x_i \partial x_j} \right) [L(x, t) Q L^T(x, t)]_{ij} \right]
\]

\[
= \int \left( \frac{\partial^2 \phi}{\partial x_i \partial x_j} \right) [L(x, t) Q L^T(x, t)]_{ij} p(x, t) \, dx
\]

\[
= - \int \left( \frac{\partial \phi}{\partial x_j} \right) \frac{\partial}{\partial x_i} \left\{ [L(x, t) Q L^T(x, t)]_{ij} p(x, t) \right\} \, dx
\]

\[
= \int \phi(x) \frac{\partial^2}{\partial x_i \partial x_j} \left\{ [L(x, t) Q L^T(x, t)]_{ij} p(x, t) \right\} \, dx
\]
Fokker-Planck-Kolmogorov PDE: Derivation [5/5]

Our equation now looks like this:

\[
\int \phi(x) \frac{\partial p(x, t)}{\partial t} \, dx = - \sum_i \int \phi(x) \frac{\partial}{\partial x_i} [f_i(x, t) p(x, t)] \, dx \\
+ \frac{1}{2} \sum_{ij} \int \phi(x) \frac{\partial^2}{\partial x_i \partial x_j} \{[L(x, t) Q L^T(x, t)]_{ij} p(x, t)\} \, dx
\]

This can also be written as

\[
\int \phi(x) \left[ \frac{\partial p(x, t)}{\partial t} + \sum_i \frac{\partial}{\partial x_i} [f_i(x, t) p(x, t)] \\
- \frac{1}{2} \sum_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \{[L(x, t) Q L^T(x, t)]_{ij} p(x, t)\} \right] \, dx = 0.
\]

But the function is \( \phi(x) \) arbitrary and thus the term in the brackets must vanish \( \Rightarrow \) Fokker–Planck–Kolmogorov equation.
FPK Example: Benes SDE

The FPK for the SDE

\[ \mathrm{d}x = \tanh(x) \, \mathrm{d}t + \mathrm{d}\beta \]

can be written as

\[
\frac{\partial p(x, t)}{\partial t} = -\frac{\partial}{\partial x} (\tanh(x) p(x, t)) + \frac{1}{2} \frac{\partial^2 p(x, t)}{\partial x^2}
\]

\[
= (\tanh^2(x) - 1) p(x, t) - \tanh(x) \frac{\partial p(x, t)}{\partial x} + \frac{1}{2} \frac{\partial^2 p(x, t)}{\partial x^2}.
\]
Mean and Covariance of SDE [1/2]

- Using Itô formula for $\phi(x, t)$, taking expectations and dividing by $dt$ gives

$$\frac{d \mathbb{E}[\phi]}{dt} = \mathbb{E} \left[ \frac{\partial \phi}{\partial t} \right] + \sum_i \mathbb{E} \left[ \frac{\partial \phi}{\partial x_i} f_i(x, t) \right]$$

$$+ \frac{1}{2} \sum_{ij} \mathbb{E} \left[ \left( \frac{\partial^2 \phi}{\partial x_i \partial x_j} \right) [\mathbf{L}(x, t) \mathbf{Q} \mathbf{L}^T(x, t)]_{ij} \right]$$

- If we select the function as $\phi(x, t) = x_u$, then we get

$$\frac{d \mathbb{E}[x_u]}{dt} = \mathbb{E} [f_u(x, t)]$$

- In vector form this gives the differential equation for the mean:

$$\frac{d \mathbf{m}}{dt} = \mathbb{E} [\mathbf{f}(x, t)]$$
If we select \( \phi(x, t) = x_u x_v - m_u(t) m_v(t) \), then we get differential equation for the components of covariance:

\[
\frac{d}{dt} \mathbb{E}[x_u x_v - m_u(t) m_v(t)] = \mathbb{E}[(x_v - m_v(t)) f_u(x, t)] + \mathbb{E}[(x_u - m_u(v)) f_v(x, t)] + \mathbb{E}[L(x, t) Q L^T(x, t)]_{uv}.
\]

The final mean and covariance differential equations are

\[
\frac{d\mathbf{m}}{dt} = \mathbb{E}[\mathbf{f}(x, t)]
\]
\[
\frac{d\mathbf{P}}{dt} = \mathbb{E}[\mathbf{f}(x, t)(x - \mathbf{m})^T] + \mathbb{E}[(x - \mathbf{m}) \mathbf{f}^T(x, t)] + \mathbb{E}[L(x, t) Q L^T(x, t)]
\]

Note that the expectations are w.r.t. \( \rho(x, t) \)!
To solve the equations, we need to know \( p(x, t) \), the solution to the FPK.

In linear-Gaussian case the first two moments indeed characterize the solution.

Useful starting point for Gaussian approximations of SDEs.
\[ dx(t) = \tanh(x(t)) \, dt + dB(t), \quad x(0) = 0, \]
Higher Order Moments

It is also possible to derive differential equations for the higher order moments of SDEs.

But with state dimension $n$, we have $n^3$ third order moments, $n^4$ fourth order moments and so on.

Recall that a given scalar function $\phi(x)$ satisfies

$$\frac{d}{dt} \mathbb{E}[\phi(x)] = \mathbb{E} \left[ \frac{\partial \phi(x)}{\partial x} f(x) \right] + \frac{q}{2} \mathbb{E} \left[ \frac{\partial^2 \phi(x)}{\partial x^2} L^2(x) \right].$$

If we apply this to $\phi(x) = x^n$:

$$\frac{d}{dt} \mathbb{E}[x^n] = n \mathbb{E}[x^{n-1} f(x, t)] + \frac{q}{2} n(n - 1) \mathbb{E}[x^{n-2} L^2(x)].$$

This, in principle, is an equation for higher order moments.

To actually use this, we need to use moment closure methods.
Consider a linear stochastic differential equation
\[ \mathbf{d} \mathbf{x} = \mathbf{F}(t) \mathbf{x}(t) \, dt + \mathbf{u}(t) \, dt + \mathbf{L}(t) \, d\beta(t), \quad \mathbf{x}(t_0) \sim \mathcal{N}(\mathbf{m}_0, \mathbf{P}_0). \]

The mean and covariance equations are now given as
\[
\begin{align*}
\frac{d\mathbf{m}(t)}{dt} &= \mathbf{F}(t) \mathbf{m}(t) + \mathbf{u}(t) \\
\frac{d\mathbf{P}(t)}{dt} &= \mathbf{F}(t) \mathbf{P}(t) + \mathbf{P}(t) \mathbf{F}^T(t) + \mathbf{L}(t) \mathbf{Q} \mathbf{L}^T(t),
\end{align*}
\]

The general solutions are given as
\[
\begin{align*}
\mathbf{m}(t) &= \Psi(t, t_0) \mathbf{m}(t_0) + \int_{t_0}^{t} \Psi(t, \tau) \mathbf{u}(\tau) \, d\tau \\
\mathbf{P}(t) &= \Psi(t, t_0) \mathbf{P}(t_0) \Psi^T(t, t_0) \\
&+ \int_{t_0}^{t} \Psi(t, \tau) \mathbf{L}(\tau) \mathbf{Q}(\tau) \mathbf{L}^T(\tau) \Psi^T(t, \tau) \, d\tau
\end{align*}
\]
Mean and covariance of LTI SDEs

In LTI SDE case
\[ d\mathbf{x} = F\mathbf{x}(t) \, dt + L \, d\beta(t), \]
we have similarly
\[ \frac{d\mathbf{m}(t)}{dt} = F\mathbf{m}(t), \]
\[ \frac{d\mathbf{P}(t)}{dt} = F\mathbf{P}(t) + \mathbf{P}(t) F^T + L Q L^T. \]

The explicit solutions are
\[ \mathbf{m}(t) = \exp(F(t - t_0)) \mathbf{m}(t_0) \]
\[ \mathbf{P}(t) = \exp(F(t - t_0)) \mathbf{P}(t_0) \exp(F(t - t_0))^T \]
\[ + \int_{t_0}^{t} \exp(F(t - \tau)) L Q L^T \exp(F(t - \tau))^T \, d\tau. \]
Let the matrices $C(t)$ and $D(t)$ solve the LTI differential equation

$$
\begin{pmatrix}
\frac{dC(t)}{dt} \\
\frac{dD(t)}{dt}
\end{pmatrix} =
\begin{pmatrix}
F & LQLT^T \\
0 & -F^T
\end{pmatrix}
\begin{pmatrix}
C(t) \\
D(t)
\end{pmatrix}
$$

Then $P(t) = C(t)D^{-1}(t)$ solves the differential equation

$$
\frac{dP(t)}{dt} = FP(t) + P(t)F^T + LQL^T
$$

Thus we can solve the covariance with matrix exponential as well:

$$
\begin{pmatrix}
C(t) \\
D(t)
\end{pmatrix} = \exp\left\{ \begin{pmatrix}
F & LQLT^T \\
0 & -F^T
\end{pmatrix} t \right\}
\begin{pmatrix}
C(t_0) \\
D(t_0)
\end{pmatrix}.
$$
Let’s now consider steady state solution of LTI SDEs
\[ \text{d}x = F \text{d}t + L \text{d}\beta \]

At the steady state, the time derivatives of mean and covariance should be zero:
\[ \frac{d\mathbf{m}(t)}{dt} = F \mathbf{m}(t) = 0 \]
\[ \frac{d\mathbf{P}(t)}{dt} = F \mathbf{P}(t) + \mathbf{P}(t) F^T + L Q L^T = 0. \]

The first equation implies that the stationary mean should be identically zero \( m_{\infty} = 0 \).

The second equation gives the Lyapunov equation, a special case of algebraic Riccati equations (AREs):
\[ F \mathbf{P}_{\infty} + \mathbf{P}_{\infty} F^T + L Q L^T = 0. \]
The general solution of LTI SDE is

\[ x(t) = \exp \left( \mathbf{F} \left( t - t_0 \right) \right) x(t_0) + \int_{t_0}^{t} \exp \left( \mathbf{F} \left( t - \tau \right) \right) \mathbf{L} \, d\beta(\tau). \]

If we let \( t_0 \to -\infty \) then this becomes:

\[ x(t) = \int_{-\infty}^{t} \exp \left( \mathbf{F} \left( t - \tau \right) \right) \mathbf{L} \, d\beta(\tau) \]

The covariance function is now given as

\[
E[x(t) x^T(t')] = \int_{-\infty}^{\min(t',t)} \exp \left( \mathbf{F} \left( t - \tau \right) \right) \mathbf{L} \, d\beta(\tau) \left( \exp \left( \mathbf{F} \left( t' - \tau' \right) \right) \mathbf{L} \right)^T \, d\tau.
\]
But we already know the following:

\[ P_\infty = \int_{-\infty}^{t} \exp\left( F(t - \tau) \right) L Q L^T \exp\left( F(t' - \tau) \right)^T d\tau, \]

which, by definition, should be independent of \( t \).

If \( t \leq t' \), we have

\[
\mathbb{E}[x(t) x^T(t')] = \int_{-\infty}^{t} \exp\left( F(t - \tau) \right) L Q L^T \exp\left( F(t' - t + t - \tau) \right)^T d\tau \\
= \int_{-\infty}^{t} \exp\left( F(t - \tau) \right) L Q L^T \exp\left( F(t' - t) \right)^T d\tau \exp\left( F(t' - t) \right)^T \\
= P_\infty \exp\left( F(t' - t) \right)^T.
\]
If $t > t'$, we get similarly

\[
E[x(t) x^T(t')] = \int_{-\infty}^{t'} \exp\left(\mathbf{F}(t - \tau)\right) \mathbf{L} \mathbf{Q} L^T \exp\left(\mathbf{F}(t' - \tau)\right)^T d\tau
\]

\[
= \exp\left(\mathbf{F}(t - t')\right) \int_{-\infty}^{t'} \exp\left(\mathbf{F}(t - \tau)\right) \mathbf{L} \mathbf{Q} L^T \exp\left(\mathbf{F}(t' - \tau)\right)^T d\tau
\]

\[
= \exp\left(\mathbf{F}(t - t')\right) \mathbf{P}_{\infty}.
\]

Thus the covariance function of LTI SDE is simply

\[
C(\tau) = \begin{cases} 
\mathbf{P}_{\infty} \exp(\mathbf{F} \tau)^T & \text{if } \tau \geq 0 \\
\exp(-\mathbf{F} \tau) \mathbf{P}_{\infty} & \text{if } \tau < 0.
\end{cases}
\]
Let’s reconsider Fourier domain solutions of LTI SDEs

\[ \text{d}x = \mathbf{F} \, x(t) \, \text{d}t + \mathbf{L} \, \text{d}\beta(t) \]

We already analyzed them in white noise formalism, which required computation of

\[ W(i\omega) = \int_{-\infty}^{\infty} w(t) \exp(-i\omega t) \, \text{d}t, \]

Every stationary Gaussian process \( x(t) \) has a representation of the form

\[ x(t) = \int_{0}^{\infty} \exp(i\omega t) \, \text{d}\zeta(i\omega), \]

\( \omega \mapsto \zeta(i\omega) \) is some complex valued Gaussian process with independent increments.
The mean squared difference $E[|\zeta(\omega_{k+1}) - \zeta(\omega_k)|^2]$ corresponds to the mean power on the interval $[\omega_k, \omega_{k+1}]$.

The spectral density then corresponds to a function $S(\omega)$ such that

$$E[|\zeta(\omega_{k+1}) - \zeta(\omega_k)|^2] = \frac{1}{\pi} \int_{\omega_k}^{\omega_{k+1}} S(\omega) \, d\omega,$$

By using this kind of integrated Fourier transform the Fourier analysis can be made rigorous.

For more information, see, for example, Van Trees (1968).
Another is to consider ODE with smooth Gaussian process $u$:

$$\frac{dx}{dt} = F x(t) + L u(t),$$

We can take

$$C_u(\tau; \Delta) = Q \frac{1}{\sqrt{2\pi \Delta^2}} \exp\left(-\frac{1}{2 \Delta^2 \tau^2}\right)$$

which in the limit $\Delta \to 0$ gives the white noise.

Spectral density of the ODE solution is then

$$S_x(\omega; \Delta) = (F - (i \omega) I)^{-1} L Q \exp\left(-\frac{\Delta^2}{2} \omega^2\right) L^T (F + (i \omega) I)^{-T}.$$
In the limit $\Delta \to 0$ to get the spectral density corresponding to the white noise input:

$$S_x(\omega) = \lim_{\Delta \to 0} S_x(\omega; \Delta) = (F - (i\omega)I)^{-1} L Q L^T (F + (i\omega)I)^{-T},$$

The limiting covariance function is then

$$C_x(\tau) = \mathcal{F}^{-1}[(F - (i\omega)I)^{-1} L Q L^T (F + (i\omega)I)^{-T}].$$

Because $C_x(0) = P\_\infty$, we also get the following interesting identity:

$$P\_\infty = \frac{1}{2\pi} \int_{-\infty}^{\infty} (F - (i\omega)I)^{-1} L Q L^T (F + (i\omega)I)^{-T} \, d\omega.$$
The probability density of SDE solution $x(t)$ solves the Fokker–Planck–Kolmogorov (FKP) partial differential equation.

The mean $m(t)$ and covariance $P(t)$ of the solution solve a pair of ordinary differential equations.

In non-linear case, the expectations in the mean and covariance equations cannot be solved without knowing the whole probability density.

For higher moment moments we can derive (theoretical) differential equations as well—can be approximated with moment closure.

In linear case, we can solve the probability density and all the moments.

The covariance functions for LTI SDEs can be solved by considering stationary solutions to the equations.