LECTURE 12 PARAMETER ESTIMATION FOR STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

1. PARAMETER ESTIMATION FOR SODEs

Consider the estimation of the parameter θ in the following scalar equation

$$\mathrm{d}X(t) = \theta f(X)\,\mathrm{d}t + \sigma b(X(t))\,\mathrm{d}w(t), \tag{1.1}$$

where $w(t)$ is a standard Brownian motion and $a(\cdot)$, $b(\cdot)$ are suitable real-valued functions such that a strong solution X is well-posed.

Usually, the solution $X(t)$ is observed at some realizations $X(t_i, \omega_j)$, $i = 0, 1, 2, \dots, N$ and $j = 1, 2, \dots, M$.

1.1. Estimating σ using the quadratic variation argument.

We can apply the quadratic variation of both sides of the equation and obtain

$$\sigma^2 = \frac{\langle X\rangle_T}{\int_0^T b^2(X)(t)\,\mathrm{d}t}.$$

Thus with given observations, we obtain

$$\widehat{\sigma^2} = \frac{1}{M}\sum_{j=1}^{M} \frac{\sum_{i=0}^{N-1} |X(t_{i+1}, \omega_j) - X(t_i, \omega_j)|^2}{\sum_{i=0}^{N-1} b^2(X(t_i, \omega_j))(t_{i+1} - t_i)}. \tag{1.2}$$

It can be shown that (supposing uniform time step sizes)

$$\sqrt{N}\frac{\langle X\rangle_T - \sum_{i=0}^{N-1} |X(t_{i+1}) - X(t_i)|^2}{\sigma^2(\int_0^T b^4(X)(t)\,\mathrm{d}t)^{1/2}} \xrightarrow{d} \mathcal{N}(0, 2). \tag{1.3}$$

We then apply Slutsky's theorem to obtain the convergence of this estimator.

We will focus only on the estimation of the parameter θ as the estimation of σ is independent of estimating θ.

1.2. Estimating θ using the maximum likelihood method.

Consider the following two scalar diffusion processes driven by the same Brownian motion.

$$dX = A(t, X(t))\,dt + \sigma(t, X(t))\,dw(t), \quad X(0) = X_0, \tag{1.4}$$

$$dY = a(t, X(t))\,dt + \sigma(t, Y(t))\,dw(t), \quad Y(0) = X_0. \tag{1.5}$$

Here the functions A, a and σ satisfy the conditions to ensure existence of a unique strong solution. Assume that the initial conditions are independent of $w(t)$ and $\sigma \geq \sigma_0 > 0$.

Let

$$B(t, x) = \frac{A(t, x) - a(t, x)}{\sigma(t, x)}, \tag{1.6}$$

then by Gisarnov's theorem (e.g., Theorem 8.6.8 of [Øksendal, 2003]),

$$\widetilde{w}(t) = -\int_0^t B(s, Y(s))\,ds + w(t), \quad 0 \leq t \leq T \tag{1.7}$$

Date: December 10, 2019.

is a standard Brownian motion under P_T^Y, where the measure P_T^Y is defined by

$$P_T^Y(A) = \mathbb{E}^{\mathbb{P}}[\mathbf{1}_A \exp(\int_0^T B(t, Y(t))\, dw - \frac{1}{2}\int_0^T B^2(t, Y(t))\, dt)]. \tag{1.8}$$

Here $\mathbb{P}$ or P_T^X is the measure generated by the process X on $C([0, T]; \mathbb{R})$. We can write $Y(t)$ [1] as

$$dY = A(t, Y(t))\, dt + \sigma(t, Y(t))\, d\widetilde{w}(t). \tag{1.9}$$

Then we may define the likelihood (Radon-Nykodym derivative, see e.g. [Liptser and Shiryaev, 2001])

$$\begin{aligned}
\frac{dP_T^X}{dP_T^Y} &= \exp(\int_0^T B(t, Y(t))\, dw - \frac{1}{2}\int_0^T B^2(t, Y(t))\, dt) \\
&= \exp(\int_0^T \frac{A(t, Y(t)) - a(t, Y(t))}{\sigma^2(t, Y(t))}\, dY(t) - \frac{1}{2}\int_0^T \frac{A^2(t, Y(t)) - a^2(t, Y(t))}{\sigma^2(t, Y(t))}\, dt) \qquad (1.10) \\
&= \exp(\int_0^T \frac{A(t, X(t)) - a(t, X(t))}{\sigma^2(t, Y(t))}\, dX(t) - \frac{1}{2}\int_0^T \frac{A^2(t, X(t)) - a^2(t, X(t))}{\sigma^2(t, X(t))}\, dt). \qquad (1.11)
\end{aligned}$$

For the problem of estimating θ, let $A(t, x) = \theta f(x)$ and $a(t, x) = \theta_0 f(x)$. Then by taking the logarithm of the likelihood and letting the derivative of the log-likelihood be zero, we have

$$\widehat{\theta} = \frac{\int_0^T \frac{f(X(t))}{\sigma^2(X(t))}\, dX(t)}{\int_0^T \frac{f^2(X(t))}{\sigma^2(X(t))}\, dt}. \tag{1.12}$$

Multiplying $f(X)/\sigma^2(X)$ over both sides of the equation of X, we obtain that

$$\theta = \frac{\int_0^T \frac{f(X(t))}{\sigma^2(X(t))}\, dX(t) + \int_0^T \frac{f(X(t))}{\sigma(X(t))}\, dw(t)}{\int_0^T \frac{f^2(X(t))}{\sigma^2(X(t))}\, dt}. \tag{1.13}$$

Thus, we have

$$\widehat{\theta} - \theta = \frac{\int_0^T \frac{f(X(t))}{\sigma(X(t))}\, dw(t)}{\int_0^T \frac{f^2(X(t))}{\sigma^2(X(t))}\, dt}. \tag{1.14}$$

Under certain conditions on f and σ, we may obtain that $\widehat{\theta} - \theta$ converges in distribution to a normal random variable with zero mean and a certain variance. For convergence and its rate, we will defer discussions to the next section, where we will use the Ornstein–Uhlenbeck process as an example.

1.3. **Estimating θ using the Gaussian (or nonlinear) filter.** The parameter θ can be also estimated using the filtering technique.

Example 1.1 (Parameter estimation)**.**

$$dY = \theta G(t)\, dt + D(t)\, dB(t), Y(0) = 0, \tag{1.15}$$

Suppose that we have $Y(t)$, $G(t)$, $D(t)$, we want to find out what θ is.

[1] Here we then have a weak solution $(Y(t), \widetilde{w}(t))$ to Equation (1.1).

Solution. Here $d\theta = 0$ (state model) and Y is the observation. Then from Lecture 5, we have

$$\frac{dS}{dt} = -\frac{G^2(t)}{D^2(t)} S^2(t), \tag{1.16}$$

and

$$\widehat{\theta}_t = \mathbb{E}[\theta | F_t^Y] = \frac{\widehat{\theta}_0 S_0^{-1} + \int_0^t G(s) D^{-2}(s) dY(s)}{S_0^{-1} + \int_0^t G^2(s) D^{-2}(s)\, ds}. \tag{1.17}$$

See Lecture 5 for the derivation of this formula.

Remark 1.2. *Let $S_0^{-1} = 0$ gives the maximum likelihood estimation.*

1.4. **A least-square method.** We may also consider the following problem with multiple parameters

$$dX = \sum_{j=1}^{N} \theta_j f_j(x)\, dt + \sum_{j=1}^{N} \sqrt{\theta_j} g_j(x)\, dw_j(t), \quad \theta_j > 0. \quad X(0) = x. \tag{1.18}$$

Suppose that $\phi \in C_b^2(\mathbb{R}^d)$, then we have from Ito's formula that

$$\mathbb{E}[\phi(X(t))] - \mathbb{E}[\phi(X(0))] = \int_0^t \mathcal{L}\phi(X(s))\, ds, \tag{1.19}$$

where the operator $\mathcal{L}$ is the generator of the process $X(t)$ and

$$\mathcal{L}v = \sum_{j=1}^{n} \theta_j f_j^\top \nabla v + \frac{1}{2} \mathrm{Tr}\Big(\sum_{j=1}^{n} \theta_j g_j g_j^\top H_X v\Big).$$

Denote $\mathcal{L}_j = f_j^\top \nabla v + \frac{1}{2}\mathrm{Tr}(g_j g_j^\top \nabla\nabla v)$. Then

$$\sum_{j=1}^{n} \theta_j \int_0^t \mathcal{L}_j \phi(X(s))\, ds = \mathbb{E}[\phi(X(t))] - \mathbb{E}[\phi(x)]. \tag{1.20}$$

This equation is under-determined. If we can obtain samples for various x, e.g., $x_i \in \mathbb{R}^d$, $i = 1, 2, \ldots, M$. We then obtain an equation of the form $A\theta = b$, where $A_{i,j} = \int_0^t \mathcal{L}_j \phi(X_{0,x_i}(s))\, ds$ and $b_i = \mathbb{E}[\phi(X_{0,x_i}(t))] - \mathbb{E}[\phi(x_i)]$. We can then obtain a least-square solution $\widehat{\theta} = A^+ b$, where A^+ is the pseudoinverse of A.

2. Parameter estimation for bilinear SPDEs

Example 2.1 (Stochastic heat equation driven by additive noise)**.** *Consider the following evolution equation*

$$\mathrm{d}u(t,x) - \theta \Delta u(t,x)\, \mathrm{d}t = \sigma\, \mathrm{d}W^Q(t,x), \qquad t > 0,\ x \in G \subset \mathbb{R}^d \tag{2.1}$$

with zero initial condition $u(0,x) = 0$, and $\Delta u = \sum_{k=1}^d \partial_{x_i}^2 u$ denotes the Laplace operator, $\theta, \sigma \in \mathbb{R}_+$. Here G is a smooth bounded domain in $\mathbb{R}^d$ and we consider zero Dirichlet boundary condition.

If $-\Delta$ and Q have only point spectrum and the same set of eigenfunctions (in this case, the equation is called a *fully diagonalizable equation*), say $Qe_k = q_k e_k$ and $-\Delta e_k = \nu_k e_k$, then we have

$$\mathrm{d}u_k(t) + \theta \nu_k u_k\, \mathrm{d}t = \sigma \sqrt{q_k}\, \mathrm{d}w_k(t), \quad k \geq 1. \tag{2.2}$$

Here $u_k = (u, e_k)$ and $\{e_k\}$ is a complete orthonormal system in $L^2(G)$.

From the discussion in the last section, the MLE gives

$$\widehat{\theta}_N^k = -\frac{\sigma\int_0^T \nu_k^\alpha q_k^\rho u_k\,\mathrm{d}u_k(t)}{\int_0^T \nu_k^{1+\alpha} q_k^\rho u_k^2(t)\,\mathrm{d}t} = \theta - \frac{\sigma\int_0^T \nu_k^\alpha q_k^{\rho+1/2} u_k\,\mathrm{d}w_k(t)}{\int_0^T \nu_k^{1+\alpha} q_k^\rho u_k^2(t)\,\mathrm{d}t}, \tag{2.3}$$

where $\alpha = 1$ and $\rho = -1$. But the derivation of an estimation of θ (1.13) shows that we can take arbitrary $\alpha \in \mathbb{R}$ and $\rho \in \mathbb{R}$.

As we have an explicit form of the difference $\widehat{\theta}_N^k - \theta$

$$\widehat{\theta}_N^k - \theta = -\frac{\sigma\int_0^T \nu_k^\alpha q_k^{\rho+1/2} u_k\,\mathrm{d}w_k(t)}{\int_0^T \nu_k^{1+\alpha} q_k^\rho u_k^2(t)\,\mathrm{d}t}, \tag{2.4}$$

we may work out the convergence and its rate which is stated in the following theorem; see details of calculation in Chapter 6.1 of [Lototsky and Rozovsky, 2017].

Theorem 2.2. *Assume $q_k \equiv 1$ and $\alpha = 1$ and $\rho = -1$. For each $k \geq 1$, the estimator θ_N^k is asymptotically normal with rate $\sqrt{T}$.*

$$\lim_{T\to\infty} \sqrt{T}(\widehat{\theta}_N^k(T) - \theta) \overset{d}{=} \mathcal{N}(0, \frac{2\theta}{k^2}). \tag{2.5}$$

For each $T > 0$,

$$\lim_{k\to\infty} k(\widehat{\theta}_N^k(T) - \theta) \overset{d}{=} \mathcal{N}(0, \frac{2\theta}{T}). \tag{2.6}$$

Exercise 2.3. *Show the convergence of $\widehat{\theta}_N^k$ to θ is also in probability.*

We need $(\widehat{\theta}_N^k)^+ = \widehat{\theta}_N^k \vee 0$ to ensure that the estimator of θ is positive.

Above we use the information of one single frequency k, averaging over $k = 1, 2, \ldots, N$ leads to another estimation and we expect higher-order convergence of the averaging as N modes of solutions have more information. Denote

(2.7)
$$\widehat{\theta}_N = \sum_{k=1}^n \frac{a_k}{\sum_{k=1}^n a_k}\widehat{\theta}_N^k, \quad \widehat{\theta}_N^k = \frac{b_k}{a_k} \quad a_k = \int_0^T \nu_k^{1+\alpha} q_k^\rho u_k^2(t)\,\mathrm{d}t, \quad b_k = -\sigma\int_0^T \nu_k^\alpha q_k^\rho u_k\,\mathrm{d}u_k(t).$$

The estimator $\widehat{\theta}_N$ can be written as

$$\widehat{\theta}_N = -\frac{\sigma\sum_{k=1}^N\int_0^T \nu_k^\alpha q_k^\rho u_k\,\mathrm{d}u_k(t)}{\sum_{k=1}^N\int_0^T \nu_k^{1+\alpha} q_k^\rho u_k^2(t)\,\mathrm{d}t} = \theta - \frac{\sigma\sum_{k=1}^N\int_0^T \nu_k^\alpha q_k^{\rho+1/2} u_k\,\mathrm{d}w_k(t)}{\sum_{k=1}^N\int_0^T \nu_k^{1+\alpha} q_k^\rho u_k^2(t)\,\mathrm{d}t}. \tag{2.8}$$

When $\alpha = 1$, $\rho = -1$, $q_k \equiv= 1$, the proof convergence is sketched as follows. By the ergodicity of the process u_k,

$$\frac{1}{T}\int_0^T \sum_{k=1}^N q_k^{-1}\nu_k^2 u_k^2\,\mathrm{d}t \to \sum_{k=1}^N q_k^{-1}\frac{\sigma^2\nu_k^2}{2\mu_k(\theta)}, \quad T\to\infty.$$

We can also show that $\sigma\sum_{k=1}^N\int_0^T \nu_k q_k^{-1} u_k\,\mathrm{d}w_k(t)$ converges in distribution to a normal random variable with mean zero (see Chapter 6.3 of [Lototsky and Rozovsky, 2017] for a proof) using the martingale central limit theorem A.2 in Appendix.

Theorem 2.4. *Assume that, $\nu_k \to \infty$, $q_k \to 0$, and α, ρ are such that*

$$\nu_n^{\alpha-1} q_n^{2\rho+2} \leq M,\ n \geq 1, \qquad \sum_{k\geq 1}\nu_k^\alpha q_k^{2\rho+2} = \infty, \tag{2.9}$$

for some $M \in \mathbb{R}$. Then, $\widehat{\theta}_N$ is a consistent estimator of θ, i.e. $\widehat{\theta}_N \to \theta$ with probability one. Moreover, if

$$\sum_{k\geq 1} \nu_k^{2\alpha-1} q_k^{4\rho+4} = \infty, \tag{2.10}$$

then, $\widehat{\theta}_N$ is also asymptotically normal

$$\frac{\sum_{k=1}^N \nu_k^{\alpha} q_k^{2\rho+2}}{\sqrt{\sum_{k=1}^N \nu_k^{2\alpha-1} q_k^{4\rho+4}}} (\widehat{\theta}_N - \theta) \xrightarrow[N\to\infty]{d} \mathcal{N}(0, \frac{2\theta}{T}). \tag{2.11}$$

When $q_k \equiv 1$ and $\nu_k = k^2$ ($G = (0, \pi) \subsetneq \mathbb{R}^1$), then the convergence rate is

$$\frac{\sqrt{\sum_{k=1}^N \nu_k^{2\alpha-1}}}{\sum_{k=1}^N \nu_k^{\alpha}} \approx N^{(2(2\alpha-1)+1)/2} N^{-(2\alpha+1)} = N^{-\frac{3}{2}}. \tag{2.12}$$

Exercise 2.5. *What is the convergence rate for θ_N when $T \to \infty$?*

Example 2.6 (Stochastic heat equation driven by additive noise)**.**

$$\mathrm{d}u(t,x) - (A_0 u(t,x) + \theta A_1 u(t,x))\,\mathrm{d}t = \sigma\,\mathrm{d}W^Q(t,x), \quad (t,x) \in (0,T] \times (0,\pi). \tag{2.13}$$

with zero initial and boundary conditions, and $\theta \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$.

The procedure of finding θ is similar to the one we use in the last example and the estimator is

$$\widehat{\theta}_N = -\frac{\displaystyle\int_0^T (Q^{-1}A_1 u^N, du_N) - \int_0^T (Q^{-1}A_1 u^N, A_0 u^N),\,\mathrm{d}t}{\displaystyle\int_0^T \left\| Q^{-1/2} A_1 u^N \right\|^2 \mathrm{d}t}, \tag{2.14}$$

while

$$\theta = -\frac{\displaystyle\int_0^T (Q^{-1}A_1 u^N, du_N) - \int_0^T (Q^{-1}A_1 u^N, A_0 u^N),\,\mathrm{d}t + \sigma \int_0^T Q^{-1}A_1 u^N\,\mathrm{d}W^Q(t)}{\displaystyle\int_0^T \left\| Q^{-1/2} A_1 u^N \right\|^2 \mathrm{d}t}. \tag{2.15}$$

Here $u^N = P^N u = \sum_{k=1}^n u_k e_k$, where e_k's are the common eigenfunctions of A_0, A_1 and the operator Q. If we set $A_0 e_k = \rho_k e_k$ ($\rho_k > 0$), $A_1 e_k = \nu_k e_k$ ($\nu_k > 0$) and $Q e_k = q_k e_k$, we then have

$$\widehat{\theta}_N - \theta = -\frac{\displaystyle\sigma \int_0^T Q^{-1}A_1 u^N\,\mathrm{d}W^Q(t)}{\displaystyle\int_0^T \left\| Q^{-1/2} A_1 u^N \right\|^2 \mathrm{d}t} = -\frac{\displaystyle\sigma \sum_{k=1}^N \int_0^T \nu_k q_k^{-1} u_k\,\mathrm{d}w_k(t)}{\displaystyle\int_0^T \sum_{k=1}^N q_k^{-1} \nu_k^2 u_k^2\,\mathrm{d}t}. \tag{2.16}$$

Denote $I_{N,\theta} = \sigma^2 \sum_{k=1}^N \frac{\nu_k^2}{\mu_k(\theta)}$, where $\mu_k(\theta) = \nu_k + \theta\rho_k$.

Theorem 2.7. *Let $q_k \equiv 1$. Assume that $\theta \in \mathbb{R}$ and $u(0,x) = 0$. Assume also that $\mu_k(\theta) \nearrow +\infty$ when $k \to \infty$ and $\sup_k \frac{|\nu_k|}{\mu_k(\theta)} < +\infty$. When $\lim_{N\to\infty} I_{N,\theta} = +\infty$, then $\lim_{N\to\infty} \widehat{\theta}_N = \theta$ with probability one.*

$$\lim_{N\to\infty} \sqrt{I_{N,\theta}} (\widehat{\theta}_N - \theta) \overset{d}{=} \mathcal{N}(0, \frac{2}{T}). \tag{2.17}$$

When $\lim_{N\to\infty} I_{N,\theta} < +\infty$,

$$\lim_{N\to\infty} \widehat{\theta}_N = -\frac{\displaystyle\sum_{k=1}^{\infty} \nu_k q_k^{-1} \int_0^T u_k \,\mathrm{d}w_k(t)}{\displaystyle\sum_{k=1}^{\infty} \int_0^T q_k^{-1} \nu_k^2 u_k^2 \,\mathrm{d}t}. \tag{2.18}$$

The proof of this theorem can be found at Chapter 6.3 of [Lototsky and Rozovsky, 2017].

2.1. **General theory.** Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ satisfy the usual conditions and let H be a separable Hilbert space equipped with the inner product $(\,\cdot\,,\,\cdot\,)$ and the norm $\|\cdot\|$.

Let's consider the following evolution equation

$$\mathrm{d}u(t) + (\theta A + B)\,u(t)\,\mathrm{d}t = (Mu(t) + \sigma)\,\mathrm{d}W^Q(t), \tag{2.19}$$

with the initial condition $u(0) = u_0 \in H$, and $\sigma, \theta > 0$ and

- the leading operator A is a linear, positive defined, self-adjoint operator in H;
- $\mathcal{B}$ is a linear or nonlinear operator in H;
- $\mathcal{M}$ is an operator acting in a suitable Hilbert space;
- W^Q is a Q-cylindrical Brownian motion in H.

We consider diagonalizable SPDEs (not key assumption but it does simplifies problems)

- the operators A and Q have pure point spectrum, and a common system of eigenfunctions $\{e_k\}_{k\geq 1}$ that forms a complete orthonormal system in H.

Thus

$$\dot{W}^Q(t) = \sum_{k=1}^{\infty} \sqrt{q_k} e_k \dot{w}_k(t),$$

where w_k are independent standard Brownian motions.

Estimating σ for SPDEs is essentially the same as for SODEs. We will focus on estimating the parameter $\theta \in \Theta \subset \mathbb{R}^+$ and assume hat the positive constant σ is known.

We apply the projection operator $\mathcal{P}^N$ to (2.19), and obtain

$$du^N + (\theta A u^N(t) + \Psi^N)dt = \sigma dW^{Q,N}(t), \quad u^N(0) = P^N u_0, \tag{2.20}$$

where $\Psi^N = \mathcal{P}^N B(u)$, and $W^{Q,N} = \mathcal{P}^N W^Q$. and $\mathcal{P}^N$ is the projection operator from H onto $H^N = \operatorname{span}\{e_k : k = 1, \ldots, N\}$.

Denote $u_N = \mathcal{P}^N u = \sum_{k=1}^N u_k e_k$, where $u_k = (u, e_k),\ k \geq 1$, are the Fourier coefficients (or modes) of the solution u with respect to $\{e_k\}_{k\geq 1}$. Denote by $\mathbb{P}_\theta^{T,N}$ the probability measure on $C([0,T]; H^N)$ generated by the solution u^N of (2.20). We assume that the family of measures $\{\mathbb{P}_\theta^{T,N}(\cdot)\}_{\theta\in\Theta}$ are mutually absolutely continuous (without justification, no need actually). Using the Radon-Nykodym derivative or Likelihood Ratio

$$\begin{aligned}\frac{\mathrm{d}\mathbb{P}_\theta^N}{\mathrm{d}\mathbb{P}_{\theta_0}^N}(u_N) = \exp\Bigg(&-(\theta-\theta_0)\sigma^{-2} \int_0^T (Q^{-1} A u^N, \mathrm{d}u^N) - \frac{1}{2}(\theta^2 + \theta_0^2)\sigma^{-2} \int_0^T \left\| Q^{-1/2} A u_N \right\|^2 \mathrm{d}t \\ &- (\theta-\theta_0)\sigma^{-2} \int_0^T (Q^{-1} A u^N, \Psi^N)\,\mathrm{d}t \Bigg).\end{aligned} \tag{2.21}$$

We can formally compute the maximum likelihood estimator (MLE) for the parameter of interest θ by maximizing the log-likelihood ratio $\log(\mathrm{d}\mathbb{P}_\theta^N / \mathrm{d}\mathbb{P}_{\theta_0}^{T,N}(u_N))$ with respect to θ, and obtain the following estimator

$$\theta_N^\sharp = -\frac{\displaystyle\int_0^T (Q^{-1}Au^N, du_N) + \int_0^T (Q^{-1}Au^N, P^N B(u))\,\mathrm{d}t}{\displaystyle\int_0^T \left\|Q^{-1/2}Au^N\right\|^2 \mathrm{d}t}. \tag{2.22}$$

The above methodology can be applied to the following Navier-Stokes equation with additive noise. For details, see [Cialenco and Glatt-Holtz, 2011] or the review paper [Cialenco, 2018].

Example 2.8 (2D Navier–Stokes Equations forced with additive noise)**.** *Consider the equation that describes the flow of a viscous, incompressible fluid*

$$\mathrm{d}u - \theta\Delta u\,\mathrm{d}t + (u\cdot\nabla u)u\,\mathrm{d}t + \nabla P\,\mathrm{d}t = \sigma\,\mathrm{d}W^Q(t), \tag{2.23}$$

$$\nabla\cdot u = 0, \tag{2.24}$$

$$u(0) = u_0, \tag{2.25}$$

where u represent the velocity field and P the pressure. Here we may consider either zero boundary conditions on a bounded domain in $\mathbb{R}^2$, or periodic domain with periodic boundary conditions.

For an *estimation of multiple parameters* using the MLE, see Chapter 6.3 of [Lototsky and Rozovsky, 2017].

3. The p-variation method

Measurements of the solution $u(t,x)$ are only at some *discrete time* points t_i and/or some *discrete spatial* points x_j, over one path/realization $\omega\in\Omega$.

Let $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq 0},\mathbb{P})$ be a stochastic basis satisfying the usual conditions. We consider the following stochastic partial differential equation on $H = L^2(G)$, where $G=(0,\pi)$.

$$\begin{cases}\partial_t u(t,x) = \theta u_{xx}(t,x) + \sigma\dot{W}(t,x), & x\in G, \quad t>0,\\ u(0,x) = 0, \end{cases} \tag{3.1}$$

where θ,σ are some positive constants, and $\dot{W}(t,x)$ is a space-time white noise and we also assume zero boundary conditions $u(t,0)=u(t,\pi)=0,\ t>0$. Here $\theta,\sigma\in\mathbb{R}_+$ are the (unknown) parameters of interest.

Here we assume that the observation is at a fixed time instant and at discrete space locations. Specifically, for a fixed instant of time $t>0$, and given interval $[a,b]\subset G$, the solution u is observed at points $(t,x_j),\ j=1,\ldots,m$, with $x_j = a+(b-a)j/m, \quad j = 0,1,\ldots,m$.

Introduce the partition $\mathcal{T}^m(a,b) = \{a_j \mid a_j = a+(b-a)j/m,\ j=0,1,\ldots,m\}$ for the uniform partition of size m of a given interval $[a,b]\subset\mathbb{R}$. Assume that $t>0$ is a fixed time instant, and consider the partition $\mathcal{T}^m(a,b)$ of the fixed interval $[a,b]\subset G$. Suppose that the solution u of (3.1) is observed at the grid points $\{(t,x_j)\mid x_j\in\mathcal{T}^m(a,b), j=1,\ldots,m\}$. Consider the following estimators for θ and σ^2 respectively

$$\widetilde{\theta}_{m,t} := \frac{(b-a)\sigma^2}{2\sum_{j=1}^m (u(t,x_j)-u(t,x_{j-1}))^2},\ (\text{known } \sigma)\ , \tag{3.2}$$

$$\widetilde{\sigma}^2_{m,t} := \frac{2\theta}{b-a}\sum_{j=1}^m (u(t,x_j)-u(t,x_{j-1}))^2,\ (\text{known } \theta)\ . \tag{3.3}$$

These estimators are consistent and asymptotically normal.

Theorem 3.1. *Assume that u is the solution of* (3.1) *and that σ is known. The estimator* (3.2) *of θ is (strongly) consistent, i.e. $\lim_{m\to\infty}\widetilde{\theta}_{m,t}=\theta$ with probability one, and asymptotically normal,*

$$\sqrt{m}(\widetilde{\theta}_{m,t}-\theta)\xrightarrow[m\to\infty]{d}\mathcal{N}(0,2\theta^2). \tag{3.4}$$

Assuming that θ is known, the estimator (3.3) *is a (strongly) consistent and asymptotically normal estimator of σ^2, with*

$$\sqrt{m}(\widetilde{\sigma}^2_{m,t}-\sigma^2)\xrightarrow[m\to\infty]{d}\mathcal{N}(0,2\sigma^4). \tag{3.5}$$

Proof. For every fixed $t>0$, there is a Brownian motion $B(x)$ on $[0,\pi]$, and a Gaussian process $R(x)\in C^\infty(0,\pi)$ such that

$$u(t,x)=\frac{\sigma}{\sqrt{2\theta}}B(x)+R(x),\quad x\in[0,\pi]. \tag{3.6}$$

Indeed, we may take

$$B(x)=\xi_0+\sum_{k\ge1}\frac{1}{k}\xi_k e_k(x),\qquad R(x)=-\frac{\sigma x}{\sqrt{2\theta\pi}}\xi_0+\frac{\sigma}{\sqrt{2\theta}}\sum_{k\ge1}\frac{a_k-1}{k}\xi_k e_k(x),$$

$$\xi_k=\sqrt{\frac{2\theta k^2}{(1-e^{-2\theta k^2 t})\sigma^2}}u_k(t),\qquad a_k=\sqrt{1-e^{-2\theta k^2 t}}.$$

Here ξ_k's are i.i.d. standard Gaussian random variables. From Lecture 2, we know that B is a standard Brownian motion on $[0,\pi]$ and that R is smooth.

We then can calculate the convergence rate explicitly using the representation (3.6) □

Remark 3.2 (fixed location and various observations at discrete time instants)**.** *Assume that the solution u is observed at points $\{(t_i,x),\ i=1,\dots,n\}$, where $t_i:=c+(d-c)i/n,\ i=0,1,\dots,n$, and $t_i\in[c,d]\subset(0,+\infty)$ and a fixed x from the interior of G. For the equation* (3.1)*, we may consider the following estimators for θ, and σ^2 respectively,*

$$\widehat{\theta}_{n,x}:=\frac{3(d-c)\sigma^4}{\pi\sum_{i=1}^n(u(t_i,x)-u(t_{i-1},x))^4},\ \textit{(known }\sigma\textit{)}, \tag{3.7}$$

$$\widehat{\sigma}^2_{n,x}:=\sqrt{\frac{\theta\pi}{3(d-c)}\sum_{i=1}^n(u(t_i,x)-u(t_{i-1},x))^4},\ \textit{(known }\theta\textit{)}. \tag{3.8}$$

These estimators are consistent and asymptotically normal, see [Cialenco and Huang, 2020]*, where the schemes for estimating θ and σ concurrently are also discussed.*

For estimations based on Bayesian analysis, interested readers are referred to [Cheng et al., 2018].

Appendix A. Central limit theorems

Theorem A.1. *Suppose that $\sigma_k\in\mathbb{L}^2(\Omega,L^2([0,T]))$ is a sequence of independent real-valued predictable process and the strong law of large numbers holds*

$$\lim_{N\to\infty}\frac{\sum_{k=1}^N\int_0^T\sigma_k^2\,dt}{\sum_{k=1}^N\mathbb{E}[\int_0^T\sigma_k^2\,dt]}=1,\ a.s. \tag{A.1}$$

Then the central limit theorem holds

$$\frac{\sum_{k=1}^N\int_0^T\sigma_k\,dw_k(t)}{\big(\sum_{k=1}^N\mathbb{E}[\int_0^T\sigma_k^2\,dt]\big)^{1/2}} \tag{A.2}$$

converges in distribution to a standard normal random variable when $N \to \infty$.

The first conclusion is a variant of the strong law of large numbers, see e.g. Theorem 6.2.11 of [Lototsky and Rozovsky, 2017] and the second conclusion can be proved by the following theorem.

Theorem A.2 (Martingale central limit theorem, Theorem 5.5.4 of [Liptser and Shiryayev, 1989])**.** *Assume that for* $t \geq 0$ *and* $\varepsilon > 0$, $X_\varepsilon(t)$ *and* $X(t)$ *are real-valued continuous square-integrable martingales. Also,* $X(t)$ *is a Gaussian process with* $X_\epsilon(0) = X(0) = 0$. *If for some* $t_0 > 0$,

$$\lim_{\varepsilon \to 0} \langle X_\varepsilon \rangle (t_0) = \langle X \rangle (t_0), \text{ in probability,} \tag{A.3}$$

Then $\lim_{\varepsilon \to 0} X_\varepsilon(t_0) \overset{d}{=} X(t_0)$.

Appendix B. Some important topics we didn't cover

- Nonlinear problems with non-globally Lipschitiz nonlinear terms (we assume that at most Lipschitz nonlinearity for SPDEs, but we do have non-globally Lipschitiz nonlinear terms for SODEs)
- Weak convergence, including convergence in probability, convergence in law etc.
- Spectral methods, finite/spectral element methods for space and/or time discretization (we focus on finite difference schemes)
- Numerical methods for multiscale stochastic partial differential equations
- Theory and numerics for stochastic partial differential equations with Levy noises, see e.g. [Peszat and Zabczyk, 2007];
- Inference of a stochastic process (or random field)
- Large Deviations for SPDEs and their computation
- $\cdots$

References

[Cheng et al., 2018] Cheng, Z., Cialenco, I., and Gong, R. (2018). Bayesian Estimations for Diagonalizable Bilinear SPDEs. *arXiv*, page arXiv:1805.11747.

[Cialenco, 2018] Cialenco, I. (2018). Statistical inference for SPDEs: an overview. *Stat. Inference Stoch. Process.*, 21(2):309–329.

[Cialenco and Glatt-Holtz, 2011] Cialenco, I. and Glatt-Holtz, N. (2011). Parameter estimation for the stochastically perturbed Navier-Stokes equations. *Stochastic Process. Appl.*, 121(4):701–724.

[Cialenco and Huang, 2020] Cialenco, I. and Huang, Y. (2020). A note on parameter estimation for discretely sampled spdes. *Stochastics and Dynamics*, page 2050016. Online.

[Liptser and Shiryaev, 2001] Liptser, R. S. and Shiryaev, A. N. (2001). *Statistics of random processes. I*, volume 5 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, expanded edition. General theory, Translated from the 1974 Russian original by A. B. Aries, Stochastic Modelling and Applied Probability.

[Liptser and Shiryayev, 1989] Liptser, R. S. and Shiryayev, A. N. (1989). *Theory of martingales*. Kluwer Academic Publishers Group, Dordrecht.

[Lototsky and Rozovsky, 2017] Lototsky, S. V. and Rozovsky, B. L. (2017). *Stochastic partial differential equations*. Universitext. Springer, Cham.

[Øksendal, 2003] Øksendal, B. (2003). *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition. An introduction with applications.

[Peszat and Zabczyk, 2007] Peszat, S. and Zabczyk, J. (2007). *Stochastic partial differential equations with Lévy noise*. Cambridge University Press, Cambridge. An evolution equation approach.