Outline of the course

4. Linear Filtering in Continuous Time
   - The Kalman-Bucy filter: scalar case
   - The Kalman-Bucy filter: general case
   - Linear filtering beyond linear diffusions
5. Nonlinear Filtering in Continuous Time
   - General framework
   - The change of reference probability measure method
   - Unnormalised conditional distribution
   - The Zakai equation
   - The Kushner-Stratonovich equation
   - The innovation process approach
6. Particle Methods

Introduction I

- Particle methods are algorithms which approximate the stochastic process $\pi = \{\pi_t\}_{t\geq 0}$ with discrete random measures of the form

$$\sum_{i=1}^{n} a_j(t)\,\delta_{v_j(t)}.$$

- That is, π is approximated with empirical distributions associated with sets of randomly located particles of stochastic masses $\{a_j\}_{j=1,\ldots,n}$, which have stochastic positions $\{v_j\}_{j=1,\ldots,n} \subseteq S$.
- The basis of this class of numerical method is the representation of π given by the Kallianpur-Striebel formula.

Introduction II

- That is, for any $\varphi \in B_b(S)$ we have

$$\pi_t(\varphi) = \frac{\rho_t(\varphi)}{\rho_t(\mathbf{1})},$$

where

$$\rho_t(\varphi) = \tilde{\mathbb{E}}\left[\varphi(X_t)\tilde{Z}_t \middle| \mathcal{Y}_t\right],$$

and

$$\tilde{Z}_t = \exp\left\{\int_0^t h^T(X_s)\,dY_s - \frac{1}{2}\int_0^t \|h(X_s)\|^2\,ds\right\}.$$

- The expectation $\tilde{\mathbb{E}}$ is taken under the probability measure $\tilde{\mathbb{P}}$ under which the process Y is a Brownian motion independent of X.
- Then one can use a Monte Carlo approximation for

$$\tilde{\mathbb{E}}\left[\varphi(X_t)\tilde{Z}_t \middle| \mathcal{Y}_t\right].$$

Basic convergence results I

- Let $\{v_j\}_{j=1,..,n}$ be n mutually independent stochastic processes and independent of Y, each of them being a solution of the martingale problem for (A, π_0).
- This means that $(v_j, Y)_{j=1,...,n}$ are identically distributed and have the same distribution as the pair (X, Y) (under $\tilde{P}$).
- Let $\{a_j\}_{j=1,...,n}$ be the following exponential martingales

$$a_j(t) = 1 + \int_0^t a_j(s)\, h^T(v_j(s))\, dY_s, \qquad t \geq 0,$$

that is,

$$a_j(t) = \exp\left\{ \int_0^t h^T(v_j(s))\, dY_s - \frac{1}{2}\int_0^t \|h(v_j(s))\|^2\, ds \right\}, \qquad t \geq 0.$$

- Hence, the triples $(v_j, a_j, Y)_{j=1,...,n}$ are identically distributed and have the same distribution as the triple $(X, \tilde{Z}, Y)$ (under $\tilde{P}$).

Basic convergence results II

- Let $\rho^n = \{\rho_t^n\}_{t\geq 0}$ and $\pi^n = \{\pi_t^n\}_{t\geq 0}$ be the following sequences of measure-valued processes

$$\rho_t^n := \frac{1}{n}\sum_{j=1}^{n} a_j(t)\,\delta_{v_j(t)}, \qquad t \geq 0,$$

$$\pi_t^n := \frac{\rho_t^n}{\rho_t^n(\mathbf{1})} = \sum_{j=1}^{n} \bar{a}_j(t)\,\delta_{v_j(t)}, \qquad t \geq 0,$$

where the normalised weights $\bar{a}_j^n$ have the form

$$\bar{a}_j^n(t) = \frac{a_j(t)}{\sum_{k=1}^{n} a_k(t)}, \qquad j = 1, ..., n, \quad t \geq 0.$$

- That is ρ_t^n is the empirical measure of n (random) particles with positions $\{v_j(t)\}_{j=1,...,n}$ and weights $\{a_j(t)/n\}_{j=1,...,n}$ and π_t^n is its normalised version.

Basic convergence results III

Lemma 6.1

For any $\varphi \in B_b(S)$ *we have*

$$\tilde{\mathbb{E}}\left[\left|\rho_t^n(\varphi) - \rho_t(\varphi)\right|^2 \middle| \mathcal{Y}_t\right] = \frac{c_{1,\varphi}(t)}{n},$$

where

$$c_{1,\varphi}(t) := \tilde{\mathbb{E}}\left[\left|\varphi(X_t)\tilde{Z}_t - \rho_t(\varphi)\right|^2 \middle| \mathcal{Y}_t\right].$$

Moreover

$$\tilde{\mathbb{E}}\left[\left|\rho_t^n(\varphi) - \rho_t(\varphi)\right|^4 \middle| \mathcal{Y}_t\right] \leq \frac{c_{2,\varphi}(t)}{n^2},$$

where

$$c_{2,\varphi}(t) := 6\tilde{\mathbb{E}}\left[\left|\varphi(X_t)\tilde{Z}_t - \rho_t(\varphi)\right|^4 \middle| \mathcal{Y}_t\right].$$

Basic convergence results IV

Remark 6.2

More generally one can prove that for any integer p and any $\varphi \in B_b(S)$,

$$\tilde{\mathbb{E}}\left[\left|\rho_t^n(\varphi) - \rho_t(\varphi)\right|^{2p} \middle| \mathcal{Y}_t\right] \leq \frac{c_{p,\varphi}(t)}{n^p},$$

where

$$c_{p,\varphi}(t) := k_p \tilde{\mathbb{E}}\left[\left|\varphi(X_t)\tilde{Z}_t - \rho_t(\varphi)\right|^{2p} \middle| \mathcal{Y}_t\right]$$

and k_p is some universal constant.

Basic convergence results V

Remark 6.3

- The previous lemma and remark are of little use if the random variables $c_{p,\varphi}$ are not finite almost surely.
- One can show, see Exercise 8.15 in **Bain & Crisan** (2009), that if $\tilde{\mathbb{E}}\left[\tilde{Z}_t^{2p}\right] < \infty$ then $\tilde{\mathbb{E}}\left[c_{p,\varphi}(t)\right] < \infty$ and in particular $c_{p,\varphi}(t)$ is finite almost surely.
- If h is bounded then $\tilde{\mathbb{E}}\left[\tilde{Z}_t^{2p}\right] < \infty$.

Remark 6.4

- The convergence of $\rho_t^n(\varphi)$ to $\rho_t(\varphi)$ holds true for larger classes of functions provided that $\varphi(X_t)\tilde{Z}_t$ is $\tilde{P}$-integrable.
- The existence of higher moments of $\varphi(X_t)\tilde{Z}_t$ ensures a control on the rate of convergence.

Basic convergence results VI

Proposition 6.5

If $\tilde{\mathbb{E}}\left[\tilde{Z}_t^{2p}\right] < \infty$, *then for any* $\varphi \in B_b(S)$,*there exists a finite* $\mathcal{Y}_t$*-measurable random variable* $\bar{c}_{p,\varphi}(t)$ *such that*

$$\tilde{\mathbb{E}}\left[\left|\pi_t^n(\varphi) - \pi_t(\varphi)\right|^{2p} \middle| \mathcal{Y}_t\right] \leq \frac{\bar{c}_{p,\varphi}(t)}{n^p}.$$

- Lemma 6.1 show the convergence of $\rho_t^n(\varphi)$ to $\rho_t(\varphi)$ when conditioned to $\mathcal{Y}_t$.
- The following result shows the convergence in expectation and almost surely.

Basic convergence results VII

Theorem 6.6

If $\tilde{\mathbb{E}}\left[\tilde{Z}_t^2\right] < \infty$, *then for any* $\varphi \in B_b(S)$ *we have*

$$\tilde{\mathbb{E}}\left[\left|\rho_t^n(\varphi) - \rho_t(\varphi)\right|\right] \leq \frac{\tilde{c}_1(t)}{\sqrt{n}} \|\varphi\|_\infty,$$

where $\tilde{c}_1(t) := \tilde{\mathbb{E}}\left[\tilde{Z}_t^2\right]^{1/2}$. *In particular,* ρ_t^n *converges to* ρ_t *in expectation.*
If $\tilde{\mathbb{E}}\left[\tilde{Z}_t^{2p}\right] < \infty$, *for* $p \geq 2$ *then for any* $\varepsilon \in \left(0, \frac{1}{2}\left(1 - \frac{1}{p}\right)\right)$ *and* $\varphi \in B_b(S)$ *there exists a positive random variable* $\tilde{c}_{\varepsilon,p,\varphi}$ *which is almost surely finite and such that*

$$\left|\rho_t^n(\varphi) - \rho_t(\varphi)\right| \leq \frac{\tilde{c}_{\varepsilon,p,\varphi}}{n^\varepsilon}.$$

In particular, ρ_t^n *converges to* ρ_t, $\tilde{P}$*-almost surely.*

Basic convergence results VIII

- The almost sure convergence of π_t^n to π_t holds under the same conditions as the convergence of ρ_t^n to ρ_t.
- However, the convergence of π_t^n to π_t in expectation requires additional integrability condition on $\rho_t^{-1}(\mathbf{1})$.

Basic convergence results IX

Theorem 6.7

If $\tilde{\mathbb{E}}\left[\tilde{Z}_t^2\right] < \infty$ *and* $\tilde{\mathbb{E}}\left[\rho_t^{-2}(\mathbf{1})\right] < \infty$, *then for any* $\varphi \in B_b(S)$, *we have*

$$\tilde{\mathbb{E}}\left[\left|\pi_t^n(\varphi) - \pi_t(\varphi)\right|\right] \leq \frac{\hat{c}_1(t)}{\sqrt{n}} \|\varphi\|_\infty,$$

where $\hat{c}_1(t) := 2\sqrt{\tilde{\mathbb{E}}\left[\tilde{Z}_t^2\right]\tilde{\mathbb{E}}\left[\rho_t^{-2}(\mathbf{1})\right]}$. *In particular,* π_t^n *converges to* π_t *in expectation.*

If $\tilde{\mathbb{E}}\left[\tilde{Z}_t^{2p}\right] < \infty$, *for* $p \geq 2$ *then for any* $\varepsilon \in \left(0, \frac{1}{2}\left(1 - \frac{1}{p}\right)\right)$ *and* $\varphi \in B_b(S)$ *there exists a positive random variable* $\hat{c}_{\varepsilon,p,\varphi}$ *which is almost surely finite and such that*

$$\left|\pi_t^n(\varphi) - \pi_t(\varphi)\right| \leq \frac{\hat{c}_{\varepsilon,p,\varphi}}{n^\varepsilon}.$$

In particular, π_t^n *converges to* π_t, $\tilde{P}$*-almost surely.*

Basic convergence results X

Remark 6.8

One can show that:

- If $\tilde{\mathbb{E}}\left[\tilde{Z}_t^{2p}\right] < \infty$, then there exists a positive constant $\tilde{c}_p(t)$ such that for any $\varphi \in B_b(S)$ we have

$$\tilde{\mathbb{E}}\left[|\rho_t^n(\varphi) - \rho_t(\varphi)|^p\right] \leq \frac{\tilde{c}_p(t)}{n^{p/2}} \|\varphi\|_\infty^p.$$

  - If $\tilde{\mathbb{E}}\left[\tilde{Z}_t^{2p}\right] < \infty$ and $\tilde{\mathbb{E}}\left[\rho_t^{-2p}(\mathbf{1})\right] < \infty$, then there exists a positive constant $\hat{c}_p(t)$ such that for any $\varphi \in B_b(S)$ we have

$$\tilde{\mathbb{E}}\left[|\pi_t^n(\varphi) - \pi_t(\varphi)|^p\right] \leq \frac{\hat{c}_p(t)}{n^{p/2}} \|\varphi\|_\infty^p.$$

Final remarks I

- We have essentially proved that both ρ_t^n converges to ρ_t, and π_t^n converges to π_t in expectation with the rate $1/\sqrt{n}$. The almost sure convergence is at a lower rate.
- The convergence analysis rely on the existence of higher moments of $\tilde{Z}_t$.
- The previous results show that the Monte Carlo method will produce good approximations of ρ_t (and π_t) provided enough particles are used.
- However, given a error tolerance, the number of particles needed will depend on the constants appearing in the upper bounds of the rates of convergence.
- The magnitude of these constants depend on the higher moments of $\tilde{Z}_t$.
- This is bad news as these higher moments of $\tilde{Z}_t$ increase very rapidly as functions of time.

Final remarks II

- The particle picture makes clearer the reason for the observed deterioration in the accuracy of these approximations with time.
- Each particle has a trajectory that is independent of the signal trajectory and its weight depends on how close its trajectory is to the signal trajectory: the weight is the likelihood of the trajectory given the observation.
- Tipically, most particles' trajectories diverge very quickly from the signal trajectory, with a few remaining close to the signal.
- Therefore the majority of the weights decrease to zero and a small minority become very large.
- This phenomenon makes the Monte Carlo method more computationally demanding since there is the need to use a larger number of particles in order to have a suffcient number of particles in the right area.

Final remarks III

- In order to fix this problem a bunch of methods (under the name of particle filters or sequential Monte Carlo methods) have been developed.
- These methods use a correction mechanism that culls particles with small weights and multiplies particles with large weights.
- The correction procedure depends on the trajectory of the particle and the observation data.
- This is effective as particles with small weight are not carried forward and those surviving to the current time contribute significantly to the estimate of the conditional distribution of the signal given the observation.