AN INTRODUCTION TO MARKOV PROCESSES AND THEIR APPLICATIONS IN MATHEMATICAL ECONOMICS

UMUT ÇETIN

1. Markov Property

2. Brief review of martingale theory

3. Feller Processes

4. Infinitesimal generators

5. Martingale Problems and Stochastic Differential Equations

6. Linear continuous Markov processes

7. Optimal filtering

Suppose that we are given on a filtered probability space an adapted process of interest, $X = (X_t)_{0 \le t \le T}$, called the *signal process*, for a deterministic T. The problem is that the signal cannot be observed directly and all we can see is an adapted *observation process* $Y = (Y_t)_{0 \le t \le T}$. The filtering is concerned with finding $\mathbb{E}[f(X_t)|\mathcal{F}_t^Y]$, where $\mathcal{F}^Y$ is the minimal filtration generated by Y and satisfying the usual hypotheses, and f is a measurable function.

Remark 1. *There is a problem with the definition of the process* $(\mathbb{E}[f(X_t)|\mathcal{F}_t^Y])_{0 \le t \le T}$ *as the conditional expectation* $\mathbb{E}[f(X_t)|\mathcal{F}_t^Y]$ *is only defined a.s. and there are uncountably many* t *between* 0 *and* T*! However, there exists a process, let's denote it with* f^o*, called the* $\mathcal{F}^Y$-optional projection *of* $f(X)$*, which satisfies* $f_t^o = \mathbb{E}[f(X_t)|\mathcal{F}_t^Y]$*, for every* t *(and some more). Moreover,* f^o *is uniquely defined. Thus, whenever we define a process by* $(\mathbb{E}[H_t|\mathcal{F}_t^Y])_{0 \le t \le T}$*, we shall always mean, and use, the* $\mathcal{F}^Y$*-optional projection of* H*. The* F^Y*-optional projection of* H *will be denoted with* $\widehat{H}$*. See the second volume of Rogers and Williams for more details. We also suppose the filtration supports two Brownian motions,* W *and* B*, such that* $d[W,B]_t = \rho_t dt$*, for some predictable process* ρ*.*

8. The innovations approach to nonlinear filtering

Let's suppose the observation process is of the form

$$Y_t = \int_0^t h_s ds + W_t, \tag{8.1}$$

where W is a standard Brownian motion and h is an adapted process such that

$$\mathbb{E}\left(\int_0^T h_s^2 ds\right) < \infty. \tag{8.2}$$

The main result of nonlinear filtering theory is the following:

Theorem 8.1. (Fujisaki, Kallianpur and Kunita)

(1) *The process N defined by*

$$N_t = Y_t - \int_0^t \widehat{h}_s ds, \tag{8.3}$$

for each $t \in [0,T]$, is an $\mathcal{F}^Y$-Brownian motion.

(2) *If M is an L^2-bounded $\mathcal{F}^Y$-martingale with $M_0 = 0$, then there exists an $\mathcal{F}^Y$-predictable process C such that*

$$\mathbb{E}\left(\int_0^T C_s^2 ds\right) < \infty,$$

and that

$$Z_t = \int_0^t C_s dN_s.$$

The F^Y-Brownian motion N is called the *innovation process* in filtering literature.

Proof. Let S be an $\mathcal{F}^Y$ stopping time. Since we only observe Y over the finite interval $[0,T]$, $S \le T$, hence bounded. Let $N_T^* = \sup_{t \le T} |N_t|$. Note that N_T^* is dominated by the random variable

$$W_T^* + \int_0^T \left\{|h_s| + |\widehat{h}_s|\right\} ds,$$

which is square integrable due to (8.2). Thus, N_S is also integrable and

$$\begin{aligned} \mathbb{E}(N_S) &= \mathbb{E}\left\{W_S + \int_0^S (h_s - \widehat{h}_s) ds\right\} \\ &= \int_0^T \mathbb{E}\left[(h_s - \widehat{h}_s) 1_{[s \le S]}\right] ds = 0, \end{aligned}$$

where we used the optional stopping theorem for $\mathbb{E}[W_S]$ in order to get the first equality, the fact that $\widehat{h}_s = \mathbb{E}[h_s | \mathcal{F}_s^Y]$ and $[s \le S] \in \mathcal{F}_s^Y$ to arrive at the last equality. This shows N is an $\mathcal{F}^Y$-martingale. Since $[N,N]_t = t$, for every $t \in [0,T]$, this shows N is an $\mathcal{F}^Y$-Brownian motion by Lévy's characterisation. See Rogers and Williams Chapter VI.8 for the proof of the second part. ☐

Let the signal process X have the following differential:

$$dX_t = \alpha_t dt + \eta_t dB_t, \tag{8.4}$$

where α is adapted and η is predictable. We further suppose

$$\mathbb{E}\left(\int_0^T \alpha_s^2 ds\right) < \infty,$$

and

$$\sup_{t \in [0,T]} \mathbb{E} X_t^2 < \infty.$$

Theorem 8.2. *Let X and Y be as above. Then we have the following fitering equations:*

$$\widehat{X}_t = \widehat{X}_0 + \int_0^t \widehat{\alpha}_s ds + \int_0^t \left\{\widehat{X_s h_s} - \widehat{X}_s \widehat{h}_s + \widehat{\eta_s \rho_s}\right\} dN_s.$$

Proof. First notice that if C is an adapted process such that

$$\mathbb{E}\int_0^T |C_s|ds < \infty,$$

and if $V_t = \int_0^t C_s ds$, then

$$\widehat{V}_t - \int_0^t \widehat{C}_s ds \text{ is an } \mathcal{F}^Y\text{-martingale.} \tag{8.5}$$

In order to prove this it suffices to prove for any $\mathcal{F}^Y$-stopping time $S \le T$, $\mathbb{E}\widehat{V}_S = \mathbb{E}\int_0^S \widehat{C}_s ds$. Indeed,

$$\begin{aligned} \mathbb{E}\widehat{V}_S = \mathbb{E}V_S &= \mathbb{E}\int_0^S C_s ds \\ &= \int_0^T \mathbb{E}\left[1_{[s\le S]}C_s\right] ds \\ &= \int_0^T \mathbb{E}\left[1_{[s\le S]}\widehat{C}_s\right] ds \\ &= \mathbb{E}\int_0^S \widehat{C}_s ds. \end{aligned}$$

This in turn implies

$$M_t := \widehat{X}_t - \widehat{X}_0 - \int_0^t \widehat{\alpha}_s ds,$$

is an $\mathcal{F}^Y$ martingale with $M_0 = 0$. Moreover, it is an L^2-bounded martingale due to the assumed integrability conditions on α and X. Thus, by Theorem 8.1 there exists a predictable process ϕ such that

$$M_t = \int_0^t \phi_s dN_s.$$

Next, we shall calculate the martingale M explicitly. In order to do this we will calculate the optional projection of XY in two different ways. Using integration by parts

$$\begin{aligned} X_tY_t &= \int_0^t X_s dY_s + \int_0^t Y_s dX_s + [X,Y]_t \\ &= \int_0^t \{X_s h_s + Y_s\alpha_s + \eta_s\rho_s\} ds + \ \mathcal{F}\text{-martingale.} \end{aligned}$$

Therefore, using (8.5)

$$\widehat{X_tY_t} = \widehat{X}_tY_t = \int_0^t \left\{\widehat{X_s h_s} + Y_s\widehat{\alpha}_s + \widehat{\eta_s\rho_s}\right\} ds + \ \mathcal{F}^Y\text{-martingale.} \tag{8.6}$$

The other way is the following:

$$\begin{aligned}\widehat{X}_tY_t &= \int_0^t \widehat{X}_t dY_s + \int_0^t Y_s d\widehat{X}_s + [\hat{X}, Y]_t \\ &= \int_0^t \widehat{X}_t \left\{ dN_s + \widehat{h}_s ds \right\} + \int_0^t Y_s \left\{ dM_s + \widehat{\alpha}_s ds \right\} + [M, N]_t \\ &= \int_0^t \left\{ \widehat{X}_s \widehat{h}_s + Y_s \widehat{\alpha}_s + \phi_s \right\} ds + \ \mathcal{F}^Y\text{-martingale}\end{aligned} \tag{8.7}$$

(8.6) and (8.7) together imply

$$\int_0^t \left\{ \widehat{X}_s \widehat{h}_s + Y_s \widehat{\alpha}_s + \phi_s \right\} ds - \int_0^t \left\{ \widehat{X_s h_s} + Y_s \widehat{\alpha}_s + \widehat{\eta_s \rho_s} \right\} ds$$

is an $\mathcal{F}^Y$-martingale, thus, must be zero being, of finite variation. This implies

$$\phi_s = \widehat{X_s h_s} - \widehat{X}_s \widehat{h}_s + \widehat{\eta_s \rho_s},$$

for each s. This proves the desired filtering equation. □

Example 8.1. (A change-detection filter.) *Let T be a random variable taking values in $(0, \infty)$ with a probability density function ϕ and tail probability $g(t) := \mathbb{P}(T > t)$. Let B be a Brownian motion independent of T and define*

$$Y_t = \int_0^t A_s ds + B_t,$$

where $A_t = 1_{[T>t]}$. We are interested in $\hat{A}$. Consider the stochastic exponential

$$L_t := \exp\left(- \int_0^t A_s dB_s - \frac{1}{2} \int_0^t A_s^2 ds \right).$$

Assume $E[\exp(T)] < \infty$. Novikov's condition now implies L is a uniformly integrable $\mathbb{P}$-martingale. Define $\mathbb{Q} \sim \mathbb{P}$ by setting $\frac{d\mathbb{Q}}{d\mathbb{P}} = L_\infty$. Then Y is a Brownian motion under $\mathbb{Q}$ by Girsanov's theorem. Moreover, it is independent from T. Note that

$$\frac{d\mathbb{P}}{d\mathbb{Q}} = L_\infty^{-1},$$

and that L^{-1} is a $\mathbb{Q}$-martingale. Now,

$$\begin{aligned}\mathbb{E}[1_{[t<T]} | \mathcal{F}_t^Y] &= \frac{\mathbb{E}^{\mathbb{Q}}[1_{[t<T]} L_\infty^{-1} | \mathcal{F}_t^Y]}{\mathbb{E}^{\mathbb{Q}}[L_\infty^{-1} | \mathcal{F}_t^Y]} \\ &= \frac{\mathbb{E}^{\mathbb{Q}}[1_{[t<T]} L_t^{-1} | \mathcal{F}_t^Y]}{\mathbb{E}^{\mathbb{Q}}[L_t^{-1} | \mathcal{F}_t^Y]} \\ &= C\mathbb{E}^{\mathbb{Q}}[1_{[t<T]} \exp(Y_t - \frac{1}{2}t) | \mathcal{F}_t^Y] \\ &= C \exp(Y_t - \frac{1}{2}t) \mathbb{Q}(T > t),\end{aligned}$$

where $C = (\mathbb{E}^{\mathbb{Q}}[L_t^{-1}|\mathcal{F}_t^Y])^{-1}$ *and in the last line we used that* T *is independent from* Y *under* $\mathbb{Q}$*. Next, we show* $\mathbb{Q}(T > u) = \mathbb{P}(T > u)$ *for all* $u > 0$*. Indeed,*

$$\begin{aligned}
\mathbb{Q}(T > u) &= \mathbb{E}[1_{[T>u]}L_\infty] \\
&= \mathbb{E}[1_{[T>u]}L_u] \\
&= \mathbb{E}[1_{[T>u]}\exp(-B_u - \frac{1}{2}u)] \\
&= \mathbb{P}(T > u)\mathbb{E}[\exp(-B_u - \frac{1}{2}u)] = \mathbb{P}(T > u),
\end{aligned}$$

since $(\exp(-B_u - \frac{1}{2}u))$ *is a martingale independent of* T *under* $\mathbb{P}$*. On the other hand,*

$$\begin{aligned}
\mathbb{E}[1_{[t\geq T]}|\mathcal{F}_t^Y] &= C\mathbb{E}^{\mathbb{Q}}[1_{[t\geq T]}L_\infty^{-1}|\mathcal{F}_t^Y] \\
&= C\mathbb{E}^{\mathbb{Q}}[1_{[t\geq T]}L_t^{-1}|\mathcal{F}_t^Y] \\
&= C\mathbb{E}^{\mathbb{Q}}[1_{[t\geq T]}\exp(Y_T - \frac{1}{2}T)|\mathcal{F}_t^Y] \\
&= C\int_0^t \exp(Y_u - \frac{1}{2}u)\phi(u)du,
\end{aligned}$$

since T *is independent from* Y *and the* $\mathbb{Q}$*-distribution of* T *is same as its* $\mathbb{P}$*-distribution from above. We must have*

$$1 = C\left(\int_0^t \exp(Y_u - \frac{1}{2}u)\phi(u)du + \exp(Y_t - \frac{1}{2}t)g(t)\right).$$

Let $Z_t = \exp(-Y_t + \frac{1}{2}t)\int_0^t \exp(Y_u - \frac{1}{2}u)\phi(u)du$ *so that*

$$C = \frac{\exp(-Y_t + \frac{1}{2}t)}{Z_t + g(t)}.$$

Recall that

$$\widehat{A}_t = \mathbb{P}(T > t|\mathcal{F}_t^Y] = C\exp(Y_t - \frac{1}{2}t)\mathbb{Q}(T > t) = \frac{g(t)}{Z_t + g(t)}$$

Now let $\widehat{F} = 1 - \widehat{A}$ *and define* $\bar{Y}$ *by* $\bar{Y}_0 = 0$ *and* $d\bar{Y}_t = \widehat{F}_t dt - dN_t$ *so that* $\bar{Y}_t = t - Y_t$*. Under this change of variable*

$$Z_t = \exp(\bar{Y}_t - \frac{1}{2}t)\int_0^t \exp(-\bar{Y}_u + \frac{1}{2}u)\phi(u)du.$$

Using integration by parts on Z *we obtain the* Zakai equation

$$dZ_t = Z_t d\bar{Y}_t + \phi(t)dt.$$

An application of Itô's formula on the function $f(t,x) := \frac{x}{g(t)+x}$ *and the process* Z *yields*

$$d\widehat{F}_t = -\widehat{F}_t(1 - \widehat{F}_t)dN_t + (1 - \widehat{F}_t)\frac{\phi(t)}{g(t)}dt.$$

Since $\widehat{F} = 1 - \widehat{A}$ *we also have*

$$d\widehat{A}_t = \widehat{A}_t(1 - \widehat{A}_t)dN_t - \widehat{A}_t\frac{\phi(t)}{g(t)}dt.$$

9. The Markov case

Observe that we have not made any Markov assumption on X or Y. We will now take a look at this special case and obtain equations that determine the conditional distribution of X.

Let's suppose X is a diffusion with generator L:

$$L = \frac{1}{2}\sigma^2(x)\frac{d^2}{dx^2} + b(x)\frac{d}{dx}.$$

We will also suppose that B and W are independent for simplicity and $h_s = h(X_s)$ and

$$\mathbb{E}\left[\int_0^T h_s^2\, ds\right] < \infty.$$

Then, using the already obtained formulas we obtain the following

Theorem 9.1. *Let $f \in C_K^2$ and define $\pi_t f = \mathbb{E}[f(X_t)|\mathcal{F}_t^Y]$. Then,*

$$\pi_t f = \pi_0 f + \int_0^t \pi_s Lf\, ds + \int_0^t \{\pi_s hf - \pi_s h \pi_s f\}\, dN_s.$$

The equation in the above theorem is called *Kushner-Stratonovic equations* or simply *filtering equations.*

Exercise 9.1. *Suppose that W and B are not independent but $d[W,B]_t = \rho(X_t, Y_t)dt$ for some measurable function $\rho(x,y)$ which is bounded when x belongs to a bounded interval. Obtain the filtering equations in this setting.*

Exercise 9.2. *Extend the filtering equations to a multidimensional setting.*

10. Kalman-Bucy filter

Kalman-Bucy filter is a celebrated example of filtering which finds widespread use in real-world problems. We assume the signal process satisfies and Ornstein-Uhlenbeck SDE:

$$X_t = X_0 + B_t + \int_0^t aX_s ds,$$

where X_0 is a normal random variable, and the observation process is given by

$$Y_t = W_t + \int_0^t cX_s ds.$$

We assume W and B are independent so that $\rho \equiv 0$. Since the bivariate process (X,Y) is Gaussian, the conditional distribution of X given Y is also Gaussian. The mean is given by $\widehat{X}$ which is given by

$$\widehat{X}_t = \mathbb{E}X_0 + \int_0^t a\widehat{X}_s ds + c\int_0^t \left\{\widehat{X_s^2} - (\widehat{X}_s)^2\right\} dN_s$$

by Theorem 8.2. Let $v_t := \mathbb{E}[(X_t - \widehat{X}_t)^2|\mathcal{F}_t^Y]$ be the conditional variance of X_t given $\mathcal{F}_t^Y$. I.e. $v_t = \widehat{X_t^2} - (\widehat{X}_t)^2$. Next, let's find the filtering equation for $\widehat{X_t^2}$. Again, using Itô's formula and Theorem 8.2

$$\widehat{X_t^2} = \mathbb{E}X_0^2 + \int_0^t (1 + 2a\widehat{X^2}_s)ds + c\int_0^t \left\{\widehat{X_s^3} - \widehat{X_s^2}\widehat{X}_s\right\} dN_s.$$

Recall that for $Z \sim N(\mu, \sigma^2)$, $\mathbb{E}Z^3 = \mu(\mu^2 + 3\sigma^2)$. Thus, since X_t is conditionally Gaussian,

$$\widehat{X^3_s} - \widehat{X^2_s}\widehat{X}_s = \widehat{X}_s\left(\widehat{X^2_s} + 3v_s - \widehat{X}^2_s\right) = 2v_s\widehat{X}_s.$$

Thus,

$$\begin{aligned} dv_t &= d(\widehat{X^2_t} - \widehat{X}^2_t) \\ &= 2cv_t\widehat{X}_t dN_t + (1 + 2a\widehat{X^2_t})dt - 2\widehat{X}_t(cv_t dN_t + a\widehat{X}_t dt) - c^2 v_t^2 dt \\ &= (1 + 2av_t - c^2 v_t^2)dt, \end{aligned}$$

so that v solves an ordinary differential equation. This differential equation has a solution. If $\beta > 0$ and $-\gamma < 0$ are two roots of the quadratic $1 + 2ax - c^2x^2$, and if $\lambda = c^2(\beta + \gamma)$, then

$$\begin{aligned} v_t &= \frac{\delta\beta e^{\lambda t} - \gamma}{\delta e^{\lambda t} + 1}, \qquad \text{where} \\ \delta &= \frac{\sigma^2 + \gamma}{\beta - \sigma^2}, \end{aligned}$$

and $\sigma^2 = \mathrm{var}(X_0)$. Note that $v(\infty) = \beta$.

Department of Statistics, London School of Economics and Political Science, 10 Houghton st, London, WC2A 2AE, UK
E-mail address: u.cetin@lse.ac.uk