Discrete time filtering equations

Salem SAID

11 August 2010

The discrete time filtering equations are very general! We here show that they can be derived with signal in a general measure space and observation with values in a locally compact group.

1 The filtering problem

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a complete probability space, $(S, \mathcal{S})$ a Polish space (separable, complete metric) and $(G, \mathcal{G})$ a locally compact group with left invariant measure μ. Here $\mathcal{S}, \mathcal{G}$ are the respective Borel σ-algebras of S, G. The discrete time filtering problem involves two processes. The signal $X = \{X_n\}_{n\in\mathbb{N}}$ has values in S, it is assumed to be a Markov chain. The observation $Y = \{Y_n\}_{n\in\mathbb{N}}$ is given by[1]

$$Y_n = h_n(X_n)W_n \qquad n \in \mathbb{N} \tag{1}$$

where $h_n : S \to G$ is a measurable function and $W = \{W_n\}_{n\in\mathbb{N}}$ are independent from each other and W is independent from X. It is assumed that W_n has strictly positive probability density L_n with respect to μ. In the following $\mathcal{X}, \mathcal{Y}, \mathcal{W}$ will denote the natural filtrations of X, Y, W.

We may define for X the transition kernel $K_n : S \times \mathcal{S} \to [0, 1]$ for $n \in \mathbb{N}$

$$\mathbb{P}(X_{n+1} \in A|\mathcal{X}_n) = \mathbb{P}(X_{n+1} \in A|X_n) = K_n(X_n, A) \qquad A \in \mathcal{S} \tag{2}$$

The goal of the filtering problem is to compute $\pi_n : G^{n+1} \times \mathcal{S} \to [0, 1]$, the conditional distribution of X_n given $\mathcal{Y}_n$ using knowledge of K_n, L_n and h_n for $n \in \mathbb{N}$.

Let us introduce some terminology from Bayesian statistics. The distribution of X_0 is denoted q_0 and assumed known. We have, for $n \in \mathbb{N}$

$$q_{n+1}(A) = \int_S K_n(x, A)q_n(dx) = (K_n q_n)(A) \qquad A \in \mathcal{S} \tag{3}$$

We put $\pi_0 = q_0$; for $n \in \mathbb{N}$, q_n is called the prior and π_n the posterior. The function $L^y_n : S \to \mathbb{R}_+$ where $L^y_n(x) = L_n(h^{-1}(x)y)$, $y \in G$ and $h^{-1}(x)$ denoting inverse in G, is called the likelihood function. In addition to π_n, we consider the prediction posterior $p_n : G^n \times \mathcal{S} \to [0, 1]$. This is the conditional distribution of X_n given $\mathcal{Y}_{n-1}$.

[1] In a usual way, G is denoted as a multiplicative group.

2 The recursion formula

The filtering problem may be solved by a recursive formula. The prior $\pi_0 = q_0$ is assumed known. For $n \in \mathbb{N}$, π_{n+1} follows from π_n by two operations, known as the predicting step and the updating step. The predicting step gives $\pi_n \Rightarrow p_{n+1}$ and only implies K_n. The updating step gives $p_{n+1} \Rightarrow \pi_{n+1}$ and only implies the likelihood function. Proposition 1 generalizes this latter operation, proposition 2 describes the predicting step and proposition 3 finally gives the recursion formula.

Proposition 1 *For $n \in \mathbb{N}$, the conditional distribution of Y_n given $\mathcal{X}_n$ has density g_n with respect to μ, where*

$$g_n(y) = L_n(h_n^{-1}(X_n)y) \tag{4}$$

Proof: Let $H_n = h_n(X_n)$. For $B \in \mathcal{G}$,

$$\mathbb{P}(Y_n \in B|\mathcal{X}_n) = \mathbb{E}[\mathbf{1}_B(H_n W_n)|\mathcal{X}_n]$$

By hypothesis, W_n is independent of $\mathcal{X}_n$. On the other hand H_n is $\mathcal{X}_n$-measurable. It follows, using the left invariance of μ,

$$\mathbb{P}(Y_n \in B|\mathcal{X}_n) = \int_G \mathbf{1}_B(H_n w) L_n(w)\mu(dw) = \int_B L_n(H_n^{-1}w)\mu(dw)$$

This gives g_n as the desired conditional density.▲

Proposition 2 *For $n \in \mathbb{N}$ we have $p_{n+1} = K_n \pi_n$. In other words,*

$$p_{n+1}(A) = \int_S K_n(x, A)\pi_n(dx) \qquad A \in \mathcal{S} \tag{5}$$

Proof: Using the definition of p_{n+1} and the fact that $\mathcal{Y}_n \subset \mathcal{X}_n \vee \mathcal{W}_n$, we have

$$p_{n+1}(A) = \mathbb{E}[\mathbb{P}(X_{n+1}|\mathcal{X}_n, \mathcal{W}_n)|\mathcal{Y}_n]$$

From the fact that $\mathcal{X}_{n+1}$ and $\mathcal{W}_n$ are independent and the Markov property of X, we may now write

$$p_{n+1}(A) = \mathbb{E}[\mathbb{P}(X_{n+1} \in A|\mathcal{X}_n)|\mathcal{Y}_n] = \mathbb{E}[\mathbb{P}(X_{n+1} \in A|X_n)|\mathcal{Y}_n]$$

From (2) and the definition of π_n, we finally obtain

$$p_{n+1}(A) = \mathbb{E}[K_n(X_n, A)|\mathcal{Y}_n] = \int_S K_n(x, A)\pi_n(dx)$$

Which is the required result.▲

Proposition 3 *For $n \in \mathbb{N}$, we have*

$$\pi_{n+1}(A) = \frac{\int_A L_{n+1}^{Y_{n+1}}(x)p_{n+1}(dx)}{\int_S L_{n+1}^{Y_{n+1}}(x)p_{n+1}(dx)} = (L_{n+1}^{Y_{n+1}} * p_{n+1})(A) \qquad A \in \mathcal{S} \tag{6}$$

Proof: The operation $L_n^{Y_n} * p_n$ is known as the projective product of $L_n^{Y_n}$ and p_n. Note that the projective product of a strictly positive function and a probability distribution is again a probability distribution.

We will denote $Y^n = (Y_0, \dots, Y_n)$ and $Y^{n+1} = (Y_0, \dots, Y_{n+1})$. In the rest of the proof, $p_{n+1}(A) = p_{n+1}^{Y^n}(A)$. We must show, for $B_0, \dots, B_{n+1} \in \mathcal{G}$ and $C_{n+1} = B_0 \times \dots \times B_{n+1}$,

$$\mathbb{P}(X_{n+1} \in A; Y^{n+1} \in C_{n+1}) = \mathbb{E}[(L_{n+1}^{Y_{n+1}} * p_{n+1}^{Y^n})(A); Y^{n+1} \in C_{n+1}]$$

Let $C^n = B_0 \times \dots \times B_n$ and note that,

$$\mathbb{P}(X_{n+1} \in A; Y^{n+1} \in C_{n+1}) = \mathbb{P}[Y_{n+1} \in B_{n+1}; X_{n+1} \in A, Y^n \in C^n] \tag{7}$$

We will now use proposition 1. Note first that,

$$\mathbb{P}[Y_{n+1} \in B_{n+1} | X_{n+1}, \mathcal{Y}_n] = \mathbb{E}[\mathbb{P}(Y_{n+1} \in B_{n+1} | \mathcal{X}_{n+1}, \mathcal{W}_n) | X_{n+1}, \mathcal{Y}_n]$$

It follows from (4), that we have a conditional density for Y_n given $X_{n+1}, \mathcal{Y}_n$,

$$\mathbb{P}[Y_{n+1} \in B_{n+1} | X_{n+1}, \mathcal{Y}_n] = \int_{B_{n+1}} L_{n+1}(h_{n+1}^{-1}(X_{n+1})y)\mu(dy) \tag{8}$$

Replacing in (7), we find after using Fubini's theorem

$$\mathbb{P}(X_{n+1} \in A; Y^{n+1} \in C_{n+1}) = \mathbb{E}\left[\int_{B_{n+1}} \mu(dy) \int_A L_{n+1}^y(x) p_{n+1}^{Y^n}(dx); Y^n \in C^n\right] \tag{9}$$

By an immediate transformation, the right hand side is equal to

$$\mathbb{E}\left[\int_S p_{n+1}^{Y^n}(dx) \int_{B_{n+1}} \left[L_{n+1}(h_{n+1}^{-1}(x)y)(L_{n+1}^y * p_{n+1}^{Y^n})(A)\right] \mu(dy); Y^n \in C^n\right]$$

By definition of p_{n+1}, this is again equal to

$$\mathbb{E}\left[\mathbb{E}\left[\int_{B_{n+1}} L_{n+1}(h_{n+1}^{-1}(X_{n+1})y)(L_{n+1}^y * p_{n+1}^{Y^n})(A)\mu(dy) \middle| \mathcal{Y}_n\right]; Y^n \in C^n\right]$$

From expression (8) of the conditional density of Y_{n+1} given $X_{n+1}, \mathcal{Y}_n$ and the fact that $\mathcal{Y}_n \subset X_{n+1}, \mathcal{Y}_n$,

$$\mathbb{P}(X_{n+1} \in A; Y^{n+1} \in C_{n+1}) = \mathbb{E}\left[\mathbb{E}\left[(L_{n+1}^{Y_{n+1}} * p_{n+1}^{Y^n})(A); Y_{n+1} \in B_{n+1} \middle| \mathcal{Y}_n\right]; Y^n \in C^n\right]$$

We may directly conclude.▲