97-F-35

A Note on the Prediction Process

Hideatsu Tsukahara
University of Tokyo

October 1997

A Note on the Prediction Process

Hideatsu Tsukahara*

Abstract

In this note, we give an example to show that the prediction process may lost Markov property if the future of the process which generates the known past is not included in the future to be predicted.

Keywords: Prediction process; Markov property.

* Author's affiliation: Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan

1. Introduction

Let $X = (X_t)_{t\in\mathbb{R}_+}$ be a measurable process with a Lusin state space $\mathbb{E}$. The prediction process $Z^X = (Z_t^X)_{t\in\mathbb{R}_+}$ of X is a measure-valued process defined by a suitable regularization of $P(\, X_{t+\bullet} \in \bullet \mid \mathcal{F}_t\,)$, where $\mathcal{F}_t$ is a natural pseudo-path filtration of X (for the precise definition, see Knight (1992)). The prediction process originally developed in Knight (1975) was, as we defined above, only relative to the pseudo-path filtration generated by the given process X. However, it is possible to construct the prediction process relative to a more or less arbitrary filtration. This is explained in Knight (1992): see also Knight (1981) and Yor (1977). We have only to introduce an auxiliary process, so that the given and auxiliary processes generate a prescribed filtration. Then define the prediction process $Z_t^{X,Y}$ of the bivariate process (X_t, Y_t) and restrict $Z_t^{X,Y}(A)$ to A that depends only on the first coordinate. Now recall that the prediction process is a homogeneous strong Markov process (see Knight (1975)). In Appendix to Chapter 6 of Knight (1992), it is noted that we have to retain all sets A in the combined σ-field in order to have the Markov properties of $Z_t^{X,Y}$. In other words, the prediction process may not be a Markov process for a general filtration $\mathcal{F}_t$. In this note, we illustrate this situation by giving an explicit example.

2. Example

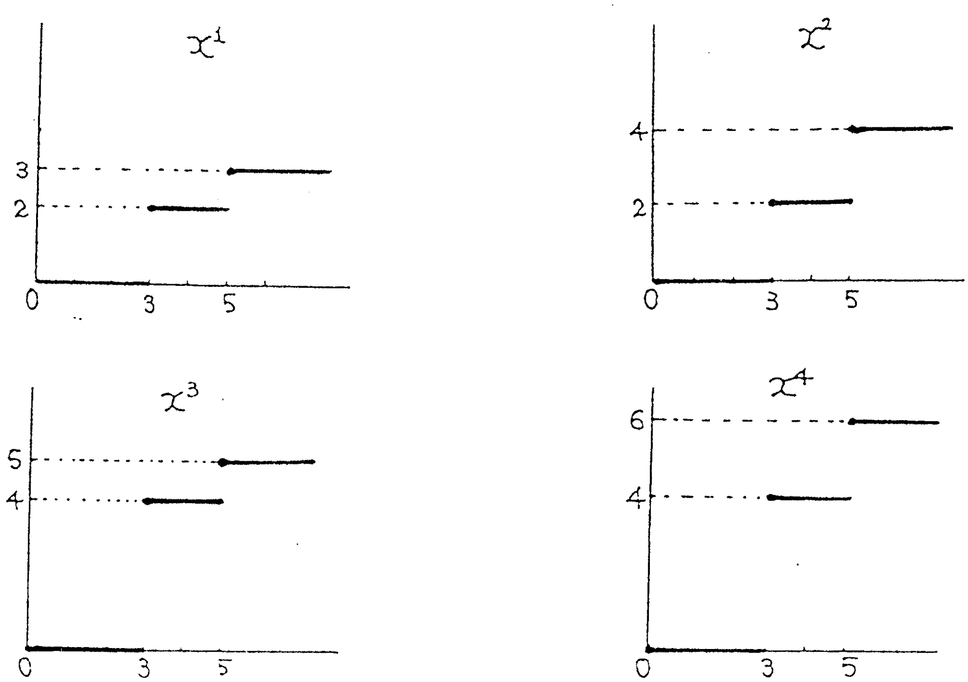
Let $\Omega = \{(\omega_1, \omega_2, \omega_3, \omega_4, \omega_5) \colon \omega_i = 1 \text{ or } 2,\ i = 1,2,3,4,5\} = \{1,2\}^5$ and define

$$\mathcal{F}_t \triangleq \begin{cases} \{\varnothing, \Omega\}, & 0 \le t < 1; \\ \sigma(\varnothing, \{(1,\cdot,\cdot,\cdot,\cdot)\}, \{(2,\cdot,\cdot,\cdot,\cdot)\}, \Omega), & 1 \le t < 2; \\ \sigma(\varnothing, \{(1,1,\cdot,\cdot,\cdot)\}, \{(1,2,\cdot,\cdot,\cdot)\}, \{(2,1,\cdot,\cdot,\cdot)\}, \{(2,2,\cdot,\cdot,\cdot)\}, \Omega), & 2 \le t < 3; \\ \sigma(\varnothing, \{(1,1,1,\cdot,\cdot)\}, \{(1,1,2,\cdot,\cdot)\}, \dots, \{(2,2,2,\cdot,\cdot)\}, \Omega), & 3 \le t < 4; \\ \sigma(\varnothing, \{(1,1,1,1,\cdot)\}, \{(1,1,1,2,\cdot)\}, \dots, \{(2,2,2,2,\cdot)\}, \Omega), & 4 \le t < 5; \\ \text{all subsets of } \Omega, & 5 \le t, \end{cases}$$

where, for example, $\{(1,\cdot,\cdot,\cdot,\cdot)\} = \{(\omega_1,\omega_2,\omega_3,\omega_4,\omega_5): \omega_1 = 1\}$ and so on. Namely, $\mathfrak{F}_t$ is the set of all $A \subset \Omega$ such that whether $\omega \in A$ depends only on ω_i for $i \le t$. Define

$$X_t(\omega) \triangleq \begin{cases} 2\omega_2 \mathbf{1}_{\{t\ge 3\}} + \omega_4 \mathbf{1}_{\{t \ge 5\}}, & \text{if } \omega_1 = 1; \\ 2\omega_3 \mathbf{1}_{\{t\ge 3\}} + \omega_5 \mathbf{1}_{\{t \ge 5\}}, & \text{if } \omega_1 = 2. \end{cases}$$

There are four possible paths as shown below:



We call these paths x^1, x^2, x^3,and x^4 as marked in the above. We give a uniform probability on Ω: $P(\{\omega\}) = 1/32$ for all ω. $X = (X_t)$ is adapted to $(\mathfrak{F}_t)$ and has càdlàg paths. The distribution of X on the space of real-valued càdlàg functions $\mathbb{D}$ is given by $P(X_\bullet \in \bullet) = \frac{1}{4}(\delta_{x^1} + \delta_{x^2} + \delta_{x^3} + \delta_{x^4})$.

We define the "prediction" process of the whole path by $K_t(\omega, B) = P(X_\bullet \in B \mid \mathfrak{F}_t)(\omega)$, $B \in \mathcal{B}(\mathbb{D})$, where the right-hand side is a regular conditional distribution and $\mathcal{B}(\mathbb{D})$ denotes the usual σ-field generated by the coordinate projections. This coincides with the Borel σ-field with respect to Skorohod topology, and in general, we denote the Borel σ-field on a space S by $\mathcal{B}(S)$. Since $(\mathfrak{F}_t)$ is right continuous, there is no need to regularize (K_t). Denoting by Π the space of probability measures on $\mathbb{D}$, (K_t) is a Π-valued right continuous step process. In fact, it is easy to see that K_t is given by the following: for $0 \leq t < 2$,

$$K_t = \frac{1}{4}(\delta_{x^1} + \delta_{x^2} + \delta_{x^3} + \delta_{x^4}) \quad \text{for all } \omega.$$

For $2 \leq t < 3$,

$$K_t = \begin{cases} \frac{1}{2}(\delta_{x^1} + \delta_{x^2}), & \omega \in \{(1,1,\cdot,\cdot,\cdot)\}; \\ \frac{1}{2}(\delta_{x^3} + \delta_{x^4}), & \omega \in \{(1,2,\cdot,\cdot,\cdot)\}; \\ \frac{1}{4}(\delta_{x^1} + \delta_{x^2} + \delta_{x^3} + \delta_{x^4}) & \omega \in \{(2,\cdot,\cdot,\cdot,\cdot)\}. \end{cases}$$

For $3 \leq t < 4$,

$$K_t = \begin{cases} \frac{1}{2}(\delta_{x^1} + \delta_{x^2}), & \omega \in \{(1,1,\cdot,\cdot,\cdot)\} \cup \{(2,\cdot,1,\cdot,\cdot)\}; \\ \frac{1}{2}(\delta_{x^3} + \delta_{x^4}), & \omega \in \{(1,2,\cdot,\cdot,\cdot)\} \cup \{(2,\cdot,2,\cdot,\cdot)\}. \end{cases}$$

For $4 \leq t < 5$,

$$K_t = \begin{cases} \delta_{x^1}, & \omega \in \{(1,1,\cdot,1,\cdot)\}; \\ \delta_{x^2}, & \omega \in \{(1,1,\cdot,2,\cdot)\}; \\ \frac{1}{2}(\delta_{x^1} + \delta_{x^2}), & \omega \in \{(2,\cdot,1,\cdot,\cdot)\}; \\ \delta_{x^3}, & \omega \in \{(1,2,\cdot,1,\cdot)\}; \\ \delta_{x^4}, & \omega \in \{(1,2,\cdot,2,\cdot)\}; \\ \frac{1}{2}(\delta_{x^3} + \delta_{x^4}), & \omega \in \{(2,\cdot,2,\cdot,\cdot)\}. \end{cases}$$

For $5 \le t$,

$$K_t = \begin{cases} \delta_{x^1}, & \omega \in \{(1,1,\cdot,1,\cdot)\} \cup \{(2,\cdot,1,\cdot,1)\}; \\ \delta_{x^2}, & \omega \in \{(1,1,\cdot,2,\cdot)\} \cup \{(2,\cdot,1,\cdot,2)\}; \\ \delta_{x^3}, & \omega \in \{(1,2,\cdot,1,\cdot)\} \cup \{(2,\cdot,2,\cdot,1)\}; \\ \delta_{x^4}, & \omega \in \{(1,2,\cdot,2,\cdot)\} \cup \{(2,\cdot,2,\cdot,2)\}. \end{cases}$$

Next, put $\mathcal{G}_t = \sigma(K_s,\ s \le t)$. Clearly $(\mathcal{G}_t)$ is right continuous, and we have

$$\mathcal{G}_t = \begin{cases} \{\varnothing, \Omega\}, & 0 \le t < 2; \\ \sigma(\varnothing, \{(1,1,\cdot,\cdot,\cdot)\}, \{(1,2,\cdot,\cdot,\cdot)\}, \{(2,\cdot,\cdot,\cdot,\cdot)\}, \Omega), & 2 \le t < 3; \\ \sigma(\varnothing, \{(1,1,\cdot,\cdot,\cdot)\}, \{(1,2,\cdot,\cdot,\cdot)\}, \{(2,\cdot,1,\cdot,\cdot)\}, \{(2,\cdot,2,\cdot,\cdot)\}, \Omega), & 3 \le t < 4; \\ \sigma(\varnothing, \{(1,1,\cdot,1,\cdot)\}, \{(1,1,\cdot,2,\cdot)\}, \{(1,2,\cdot,1,\cdot)\}, \{(1,2,\cdot,2,\cdot)\}, \\ \quad \{(2,\cdot,1,\cdot,\cdot)\}, \{(2,\cdot,2,\cdot,\cdot)\}, \Omega), & 4 \le t < 5; \\ \sigma(\varnothing, \{(1,1,\cdot,1,\cdot)\}, \{(1,1,\cdot,2,\cdot)\}, \{(1,2,\cdot,1,\cdot)\}, \{(1,2,\cdot,2,\cdot)\}, \\ \quad \{(2,\cdot,1,\cdot,1)\}, \{(2,\cdot,1,\cdot,2)\}, \{(2,\cdot,2,\cdot,1)\}, \{(2,\cdot,2,\cdot,2)\}, \Omega), & 5 \le t. \end{cases}$$

Now consider $E(\, X_t \mid \mathcal{F}_4 \,)$, which is the expectation of the measure $K_4(\omega, \pi_6^{-1}\bullet)$ on $\mathbb{R}$. Here $\pi_t\colon \mathbb{D} \to \mathbb{R}$ is the coordinate projection: $x \mapsto x(t)$ for $x \in \mathbb{D}$. It is $\mathcal{B}(\mathbb{D})$ measurable since $\mathcal{B}(\mathbb{D}) = \sigma(\pi_t,\ t \ge 0)$. Thus the mapping $\mu \mapsto \pi_t(\mu)$ from Π into $\mathcal{P}(\mathbb{R})$ is $\mathcal{B}(\Pi)/\mathcal{B}(\mathcal{P}(\mathbb{R}))$ measurable, where $\mathcal{P}(\mathbb{R})$ is the space of probability measures on $\mathbb{R}$. Moreover, the mapping $\nu \mapsto \int x\, \nu(dx)$ from $\mathcal{P}(\mathbb{R})$ into $\mathbb{R}$ is measurable. To be precise, the domain should be restricted to those ν's which have a finite expectation: $\{\nu \in \mathcal{P}(\mathbb{R})\colon \int |x|\, \nu(dx) < \infty\} \in \mathcal{B}(\mathcal{P}(\mathbb{R}))$. Consequently, $E(\, X_6 \mid \mathcal{F}_4 \,)$ can be written as $f(K_4)$ for some $f \in b\mathcal{B}(\Pi)$. We have

$$f(K_4) = \begin{cases} 3, & \omega \in \{(1,1,\cdot,1,\cdot)\}; \\ 4, & \omega \in \{(1,1,\cdot,2,\cdot)\}; \\ 3.5, & \omega \in \{(2,\cdot,1,\cdot,\cdot)\}; \\ 5, & \omega \in \{(1,2,\cdot,1,\cdot)\}; \\ 6, & \omega \in \{(1,2,\cdot,2,\cdot)\}; \\ 5.5, & \omega \in \{(2,\cdot,2,\cdot,\cdot)\}. \end{cases}$$

Define $g\colon \Pi \to \mathbb{R}$ by $g(\mu) = [f(\mu)]^2$. Then $g \in b\mathcal{B}(\Pi)$ and

$$g(K_4) = \begin{cases} 9, & \omega \in \{(1,1,\cdot,1,\cdot)\}; \\ 16, & \omega \in \{(1,1,\cdot,2,\cdot)\}; \\ 12.25, & \omega \in \{(2,\cdot,1,\cdot,\cdot)\}; \\ 25, & \omega \in \{(1,2,\cdot,1,\cdot)\}; \\ 36, & \omega \in \{(1,2,\cdot,2,\cdot)\}; \\ 30.25, & \omega \in \{(2,\cdot,2,\cdot,\cdot)\}. \end{cases}$$

It follows that

$$E(\, g(K_4) \mid K_4 \,) = \begin{cases} 12.375, & \omega \in \{(1,1,\cdot,\cdot,\cdot)\} \cup \{(2,\cdot,1,\cdot,\cdot)\}; \\ 30.375, & \omega \in \{(1,2,\cdot,\cdot,\cdot)\} \cup \{(2,\cdot,2,\cdot,\cdot)\}. \end{cases}$$

But on the other hand, we have

$$E(\, g(K_4) \mid \mathcal{G}_3 \,) = \begin{cases} 12.5, & \omega \in \{(1,1,\cdot,\cdot,\cdot)\}; \\ 30.5, & \omega \in \{(1,2,\cdot,\cdot,\cdot)\}; \\ 12.25, & \omega \in \{2,\cdot,1,\cdot,\cdot)\}; \\ 30.25, & \omega \in \{2,\cdot,2,\cdot,\cdot)\}. \end{cases}$$

This shows that (K_t) is not $(\mathcal{G}_t)$-Markov.

The point of this example is now clear: $\sigma(K_3)$ does not have any information about ω_1, while $\sigma(K_2)$ does. And we use it to make prediction at $t = 4$. Thus, given $\sigma(K_3)$, K_2 and K_4 cannot be independent.

It may be proved analogously that the prediction process (Z_t) of X is not Markov with respect to its natural filtration.

Acknowledgements.

The author would like to thank Frank Knight for his helpful comments.

References

Knight, F.B. (1975). A predictive view of continuous time processes, *Ann. Probab.* **3**, 573-596.

Knight, F.B. (1981). *Essays on the Prediction Process*, Lect. Notes Series, 1, Institute of Mathematical Statistics, Hayward, CA.

Knight, F.B. (1992). *Foundations of the Prediction Process*, Oxford University Press, Oxford.

Yor, M. (1977). Sur les théories du filtrage et de la prédiction, *Séminaire de Probabilités XI, Lect. Notes in Math.* **581**, Springer-Verlag, Berlin Heidelberg New York, 257-297.