MARKOVIANITY OF A SUBSET OF COMPONENTS OF A MARKOV PROCESS

P. I. Kitsul*
Department of Mathematics and Statistics
Minnesota State University, Mankato, MN, USA
R. S. Liptser
Department of Electrical Engeneering
Tel Aviv University, Tel Aviv, Israel
A. P. Serebrovski
Laboratory of Stochastic Dynamical Systems
Institute for Information Transmission Problems
Russian Academy of Sciences, Moscow, Russia

Abstract

We consider the necessary and sufficient conditions for a group of the components of a stationary vector Gaussian Markov process to possess a Markov property. The representation by a linear Itô stochastic differential equation is also given.

Keywords : Markov process, stochastic realization, forward and backward representations.

Mailing address: Dr. Pavel I. Kitsul, Department of Mathematics and Statistics, Minnesota State University, Mankato, MN 56001, USA

---

[*] on leave from the Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, Russia

1 Introduction

A Markov property of a random process (even in wide sense) is effectively used in filtering and stochastic control. If a filtered or controlled process is a collection of the components of a Markov process, then the usage of the other components for creating the optimal filter or controller can be required.

That observation makes reasonable the following problem statement. Let X_t be a vector Markov process (vector-column) with components $(X_t^1, ..., X_t^n)$ and

$$Y_t = \widetilde{C} X_t, \tag{1}$$

where $\widetilde{C}$ is an $r \times n$-matrix of rank k, $k \le r < n$. A natural question is: when is Y_t a Markov process as well? To find a constructive answer we restrict ourselves by the consideration of a zero mean stationary Gaussian Markov process described by the linear Itô differential equation

$$dX_t = AX_t dt + BdW_t, \tag{2}$$

where $X_t = (X_t^1, \dots, X_t^n)$ is a column-vector, $(W_t)_{t\ge 0}$ is an n -dimensional standard Wiener process independent of X_0, $A = A_{n\times n}$ and $B = B_{n\times n}$ are constant matrices, and X_0 is a Gaussian random vector.

It is clear that $(X_t, Y_t)_{t\ge 0}$ is a stationary Gaussian Markov process. In general, since $k < n$, the process $(Y_t)_{t\ge 0}$ is not necessarily a Markov one. However, as is known from the stochastic realization theory [7] and [2], under some proper conditions on matrices $A, \widetilde{C}$, and

$$Q = Cov(X_0, X_0),$$

the process $(Y_t)_{t\ge 0}$ is nevertheless a Markov itself.

Without losing generality, one can replace $\widetilde{C}$ in (1) by the matrix

$$C = (\widetilde{C} Q \widetilde{C}^T)(\widetilde{C} Q \widetilde{C}^T)^+ \widetilde{C}, \tag{3}$$

where $\widetilde{C} Q \widetilde{C}^T)^+$ is the Moore-Penrose pseudoinverse of the matrix $(\widetilde{C} Q \widetilde{C}^T)$, see e.g. [1], that is

$$Y_t = CX_t. \tag{4}$$

In fact, since $\mathbf{E}(X_t X_t^T) \equiv Q$, we get

$$\begin{aligned}
&\mathbf{E}[(\widetilde{C} - C)X_t X_t^T (\widetilde{C} - C)^T] \\
&= \Big(I - (\widetilde{C} Q \widetilde{C}^T)(\widetilde{C} Q \widetilde{C}^T)^+\Big)(\widetilde{C} Q \widetilde{C}^T)\Big(I - (\widetilde{C} Q \widetilde{C}^T)(\widetilde{C} Q \widetilde{C}^T)^+\Big) \\
&= 0.
\end{aligned}$$

In this paper necessary and sufficient conditions for the $(Y_t)_{t\geq 0}$ to possess a Markov property in terms of matrices A, C and Q are formulated (not explicitly using [7] and [2]). The Itô linear differential equation for Y_t is also given.

2 Auxiliary Results

The necessary and sufficient condition for a Gaussian zero mean vector process to possess a Markov property was given by J. L. Doob in [4] (see Ch. V, Theorem 8.1 in terms of its correlation matrix $K(t,s)$, for the positive definite matrices $K(u,u)$; see also [6] and [5] for the singular $K(u,u)$).
For any $s \leq u \leq t$

$$K(t,s)K^+(s,s) = K(t,u)K^+(u,u)K(u,s)K^+(s,s).$$

For a stationary Gaussian process that condition has the following form:

$$K(t-s)K^+(0) = K(t-u)K^+(0)K(u-s)K^+(0). \tag{5}$$

In our particular case (recall that X_t is the stationary process with $\mathbf{E}X_0 \equiv 0$ and hence $\mathbf{E}Y_0 \equiv 0$) from (2) it follows that $\mathbf{E}(X_t X_0^T) = e^{tA}Q$, so that for any $t \geq 0$

$$K(t) = \mathbf{E}(Y_t Y_0^T) = Ce^{tA}QC^T. \tag{6}$$

Let us denote

$$M = QC^T(CQC^T)^+. \tag{7}$$

According to (5) and (6) the necessary and sufficient condition for the process $(Y_t)_{t\geq 0}$ to possess a Markov property is formulated as

$$Ce^{(t-s)A}M = Ce^{(t-u)A}MCe^{(u-s)A}M. \tag{8}$$

In the next section, we will need the following

Lemma 1 *Let according to (3)*

$$C = (\widetilde{C}Q\widetilde{C}^T)(\widetilde{C}Q\widetilde{C}^T)^+\widetilde{C}.$$

Then the matrix C satisfies

$$C = (CQC^T)(CQC^T)^+C. \tag{9}$$

Proof. Since

$$CQC^T = (\widetilde{C}Q\widetilde{C}^T)(\widetilde{C}Q\widetilde{C}^T)^+\widetilde{C}Q\widetilde{C}^T(\widetilde{C}Q\widetilde{C}^T)^+(\widetilde{C}Q\widetilde{C}^T) = (\widetilde{C}Q\widetilde{C}^T)$$

we obtain

$$\begin{aligned}(CQC^T)(CQC^T)^+C &= (CQC^T)(CQC^T)^+(\widetilde{C}Q\widetilde{C}^T)(\widetilde{C}Q\widetilde{C}^T)^+\widetilde{C} \\ &= (\widetilde{C}Q\widetilde{C}^T)(\widetilde{C}Q\widetilde{C}^T)^+(\widetilde{C}Q\widetilde{C}^T)(\widetilde{C}Q\widetilde{C}^T)^+\widetilde{C} \\ &= (\widetilde{C}Q\widetilde{C}^T)(\widetilde{C}Q\widetilde{C}^T)^+\widetilde{C} \\ &= C.\end{aligned}$$

3 Markov Property Condition in Terms of A, C, Q

Let us denote (for fixed u, $s \le u \le t$)

$$G(s,u,t) \;=\; Ce^{(t-s)A}M \;-\; Ce^{(t-u)A}MCe^{(u-s)A}M \;.$$

The condition (8) is obviously equivalent to $G(s,u,t) \;\equiv\; 0$. But the latter means that for any $p, q \;=\; 0, 1, 2, \ldots$

$$0 \;\equiv\; \frac{\partial^{p+q}G(s,u,t)}{\partial t^p \partial s^q}\,|_{s=u=t} \;=\; (-1)^q\; C[A^{p+q} \;-\; A^pMCA^q]M \;,$$

which gives

$$CA^p[I \;-\; MC]A^qM \;=\; 0 \qquad (p, q \;=\; 0, 1, 2, ...)\;. \tag{10}$$

So, the necessary and sufficient condition (8) is equivalent to (10).

It is easy to show that (10) can be transformed into an equivalent form that can also be used as the necessary and sufficient condition.

Lemma 2 *Let* $M = QC^T(CQC^T)^+$. *Then*

$$CA^p[I \;-\; MC]A^qM \;=\; 0 \qquad (p, q \;=\; 0, 1, 2, ...)$$

if and only if

$$CA^mM \;=\; (CAM)^m \;\; (m = 1, 2, \ldots). \tag{11}$$

Proof.
A. Let

$$CA^{p+q}M = CA^pMCA^qM.$$

for any whole numbers p and q. Then by the principle of mathematical induction

$$\begin{aligned} (CAM)^m &= (CAM)^{m-1}(CAM) \\ &= (CA^{m-1}M)(CAM) \\ &= CA^{m-1}MCAM \\ &= CA^mM. \end{aligned}$$

B. Let

$$CA^mM = (CAM)^m.$$

for any natural number m. Then

$$\begin{aligned} CA^{p+q}M &= (CAM)^{p+q} \\ &= (CAM)^p(CAM)^q \\ &= (CA^pM)(CA^qM) \\ &= CA^pMCA^qM. \end{aligned}$$

Lemma 3 *Let $M = QC^T(CQC^T)^+$. Then*

$$MCM = M \quad \text{and} \quad CMC = C. \tag{12}$$

Proof.

$$\begin{aligned} MCM &= QC^T(CQC^T)^+CQC^T(CQC^T)^+ \\ &= QC^T(CQC^T)^+ \\ &= M \end{aligned}$$

and by Lemma 1

$$\begin{aligned} CMC &= CQC^T(CQC^T)^+C \\ &= C. \end{aligned}$$

4 Two Sufficient Conditions

Lemma 4 *Let there exist the matrices* Λ_F *or* Λ_B *such that*

$$CA \;=\; \Lambda_F C$$

or

$$AM \;=\; M\Lambda_B \ .$$

Then for any $p, q = 1, 2, \ldots$ *respectively*

$$CA^p \;=\; \Lambda_F^p C \ \textit{or} \qquad A^q M \;=\; M\Lambda_B^q \ . \tag{13}$$

Proof. For the first equality in (13) we obtain

$$\begin{aligned} CA^p &= (CA)A^{p-1} \\ &= (\Lambda_F C)A^{p-1} \\ &= \Lambda_F(CA^{p-1}) \\ &= \ldots \\ &= \Lambda_F^{p-1}(CA) \\ &= \Lambda_F^p C \end{aligned}$$

while for the second one

$$\begin{aligned} A^q M &= A^{q-1}(AM) \\ &= A^{q-1}(M\Lambda_B) \\ &= (A^{q-1}M)\Lambda_B \\ &= \ldots \\ &= (AM)\Lambda_B^{q-1} \\ &= M\Lambda_B^q. \end{aligned}$$

Theorem 1 *In the assumptions of Lemma* 4 *the process* $(Y_t)_{t\geq 0}$ *is Markov.*

Proof. First let $CA \;=\; \Lambda_F C$. Then according to the properties of pseudoinverse matrices (see [1] or [8]), Lemmas 4 and 3, and $C[I \;-\; MC] = C - CMC = 0$ (see (12)) we have

$$CA^p[I \;-\; MC]A^q M \;=\; \Lambda_F^p C[I \;-\; MC]A^q M \;= 0.$$

Now let $AM = M\Lambda_B$. Then, by virtue of $[I - MC]M = M - MCM = 0$ (see (12)), we get

$$CA^p[I - MC]A^qM = CA^p[I - MC]M\Lambda_B^q = 0.$$

Thus each of the equalities from (12) is sufficient for (10) to be valid. This means that both conditions from (12) are sufficient conditions for the process $(Y_t)_{t\geq 0}$ to possess a Markov property.

5 Markov Representation

It follows from (2) and (4) that

$$dY_t = CAX_t\,dt + CB\,dW_t\,.$$

Let us denote by $(\mathcal{F}_t^Y)_{t\geq 0}$ the filtration generated by the process $(Y_t)_{t\geq 0}$ and satisfying the standard conditions [8]. Let $\Pi_t(Y) \equiv \mathbf{E}(X_t|\mathcal{F}_t^Y)$ and introduce an innovation process

$$z_t = Y_t - Y_0 - \int_0^t CA\Pi_s(Y)ds.$$

It is well known that z_t is a Gaussian martingale with respect to the filtration $(\mathcal{F}_t^Y)_{t\geq 0}$ with

$$\mathbf{E}(z_tz_t^T) \equiv CB\mathbf{E}(W_tW_t^T)B^TC^T = CBB^TC^Tt := DD^Tt.$$

Therefore, an innovation Wiener process $\overline{W}_t$ (defined on an extended probability space in case of a singular matrix D) can be found such that

$$dY_t = CA\Pi_t(Y)\,dt + Dd\overline{W}_t.$$

Below we show that under certain conditions

$$CA\,\Pi_t(X) = CA\,\mathbf{E}(X_t|Y_t) \qquad (\mathbf{P}-a.s.). \tag{14}$$

Lemma 5 *For M defined in (7) we have*

$$MCQC^T = QC^T.$$

Proof. According to the Lemma 1 and (7)

$$MCQC^T \;=\; QC^T(CQC^T)^+ \, CQC^T \;= QC^T.$$

Lemma 6 *Under (10)*

$$CA^{1+k}QC^T \;=\; CAM \; CA^kQC^T, \qquad k = 0, 1, 2, \dots \; . \tag{15}$$

Proof. In accordance with (10) for every $p, q \;=\; 0, 1, 2, \dots$ we have

$$CA^{p+q}M \;=\; CA^pM \; CA^qM.$$

In particular for $p = 1$ and $q = k$ the latter becomes

$$CA^{1+k}M \;=\; CAM \; CA^kM \; . \tag{16}$$

Then by Lemma 5 and (16) for every $k = 0, 1, 2, \dots$ we get

$$\begin{aligned} CA^{1+k}QC^T &= CA^{1+k}MCQC^T \\ &= (CA^{1+k}M)CQC^T \\ &= CAM \; CA^kM \; CQC^T \\ &= CAM \; CA^kQC^T. \end{aligned}$$

5.1 Verification of (14)

First let us show that

$$CA \; \mathbf{E}(\Pi_t(X)Y_s^T) \;=\; CA \; \mathbf{E}(\mathbf{E}(X_t|Y_t) \; Y_s^T) \tag{17}$$

for any $s \leq t$.

By the property of conditional expectation we have

$$\mathbf{E}\Pi_t(Y)Y_s^T = \mathbf{E}(X_tY_s^T) = \mathbf{E}(X_tX_s^T)C^T = e^{A(t-s)}QC^T$$

and, at the same time, by the theorem on normal correlation (see e.g. [8], Vol. II, Ch. 13)

$$\begin{aligned} \mathbf{E}(X_t|Y_t) &= \mathbf{E}(X_tY_t^T)(\mathbf{E}(Y_tY_t^T))^+ \; Y_t \\ &= CQ^T(CQC^T)^+ \; Y_t \;=\; M \; Y_t \; , \end{aligned}$$

we get

$$CA\,\mathbf{E}(X_tY_s^T) \;=\; CAM\;\mathbf{E}(Y_tY_s^T) := CAM\;Ce^{(t-s)A}QC^T.$$

Hence (17) is just

$$CA\;e^{(t-s)A}QC^T \;=\; CAM\;Ce^{(t-s)A}QC^T \tag{18}$$

which is equivalent to (15). The latter is obtained by differentiating both sides of (18) with respect to t and letting $t = s$.

Thus (17) holds. It is obvious that (17) remains valid for Y_s replaced by any linear combination $\sum\limits_k c_kY_{s_k}$, $s_k \leq t$ and for limits (in the mean square sense) of such linear combinations. That is, if $L\{Y_0^t\}$ is such a limit, then

$$CA\;\mathbf{E}\Pi_t(X)L\{Y_0^t\} \;=\; CA\;\mathbf{E}(\mathbf{E}(X_t|Y_t)L\{Y_0^t\})$$

or

$$CA\;\mathbf{E}((\Pi_t(X) - \mathbf{E}(X_t|Y_t))L\{Y_0^t\}) = 0.$$

Now, choosing $L\{Y_0^T\} = (\mathbf{E}((\Pi_t(X) - \mathbf{E}(X_t|Y_t)\;))^TA^TC^T$ we find

$$CA\;\mathbf{E}((\Pi_t(X) - \mathbf{E}(X_t|Y_t))((\Pi_t(X) - \mathbf{E}(X_t|Y_t))^TA^TC^T = 0,$$

so (14) holds.

5.2 Main Theorem

We are now in the position to formulate the main result.

Theorem 2 *Under (10) (or equivalently (11)), the process* $(Y_t)_{t\geq 0}$ *is a stationary Markov diffusion w.r.t. an* $(\mathcal{F}_t^Y)_{t\geq 0}$ *-innovation Wiener process* $(\overline{W}_t)_{t\geq 0}$ *with the drift vector* $CA\mathbf{E}(X_t|Y_t)(= CAMY_t)$ *and diffusion matrix* DD^T*, namely*

$$dY_t \;=\; CAMY_t\;dt \;+\; D\;d\overline{W}_t.$$

Proof. Under (15) the statement of the theorem is already proved. We should also emphasize that (15) is implied by (10) due to Lemma 6.

6 Interpretation of Sufficient Conditions

Theorem 1 states that any of conditions $CA = \Lambda_F C$ or $AM = M\Lambda_B$ is sufficient for the process $(Y_t)_{t\geq 0}$ to be Markov.

Using the condition $CA = \Lambda_F C$ we can derive the Itô differential equation for Y_t in a more simple way. Since $Y_t = CX_t$, we have

$$dY_t = CAX_t dt + CBdW_t$$

so that under the above mentioned condition

$$dY_t = \Lambda_F Y_t dt + CBdW_t,$$

and, therefore, there exists an r-dimensional Wiener process $\overline{W}_t$ adapted to the filtration $(\mathcal{F}_t^Y)_{t\geq 0}$ such that

$$dY_t = \Lambda_F Y_t dt + DdW_t,$$

where $D = (CBB^T C^T)^{1/2}$.

A similar interpretation under $AM = M\Lambda_B$ uses a backward representation for the process X_t with respect to a backward Wiener process (standard Wiener process in the reverse time) W_t^B (see e.g. [2], [3]):

$$dX_t = QA^T Q^+ X_t dt + BdW_t^B.$$

Hence, Y_t obeys the backward time Itô differential

$$dY_t = dCX_t = CQA^T Q^+ X_t dt + CBdW_t^B$$

with the drift "$CQA^T Q^+ X_t$" that will be described below. Since by Lemma 1 we have $C = (CQC^T)(CQC^T)^+ C$ and $AM = M\Lambda_B$ (recall $M = CQ^T(CQC^T)^+$) we find

$$\begin{aligned}
CQA^T Q^+ X_t &= (CQC^T)(CQC^T)^+ CQA^T Q^+ X_t \\
&= (CQC^T)M^T A^T Q^+ X_t \\
&= (CQC^T)\Lambda_B^T M^T Q^+ X_t \\
&= (CQC^T)\Lambda_B^T (CQC^T)^+ CQQ^+ X_t \\
&= (CQC^T)\Lambda_B^T (CQC^T)^+ CX_t \\
&= (CQC^T)\Lambda_B^T (CQC^T)^+ Y_t.
\end{aligned}$$

Consequently,

$$dY_t = (CQC^T)\Lambda_B^T(CQC^T)^+Y_t dt + CBW_t^B.$$

On the other hand, according to [3] the forward representation for Y_t derived from the backward one has the form

$$dY_t = (CQC^T)(CQC^T)^+\Lambda_B(CQC^T)(CQC^T)^+Y_t dt + CBdW_t,$$

or according to Lemma 1

$$dY_t = (CQC^T)(CQC^T)^+\Lambda_B Y_t dt + CBdW_t,$$

where $CBdW_t$ can be also replaced by $Dd\overline{W}_t$.

Finally, we claim that

$$CA \;=\; \Lambda_F C \implies \Lambda_F Y_t = CAMY_t \tag{19}$$

and

$$AM \;=\; M\Lambda_B \implies (CQC^T)(CQC^T)^+\Lambda_B Y_t = CAMY_t. \tag{20}$$

From formula (9) it can be easily derived that $Y_t = (CQC^T)(CQC^T)^+Y_t$. Hence, due to (7) we get

$$\Lambda_F Y_t = \Lambda_F(CQC^T)(CQC^T)^+Y_t = \Lambda_F CMY_t = CAMY_t$$

and implication (19) is valid. To establish (20), it is sufficient to note that

$$(CQC^T)(CQC^T)^+\Lambda_B Y_t = CM\Lambda_B Y_t = CAMY_t.$$

References

[1] A. Albert, Regression and the Moore-Penrose Pseudoinverse, Academic Press, New York, 1972.

[2] F. Badawi, A. Lindquist, and M. Pavon, A Stochastic Realization Approach to the Smoothing Problem, IEEE Trans. Automat. Contr., AC, 24 (1979) pp. 983-989.

[3] G. B. Di Masi, P. I. Kitsul, and R. S. Liptser, Minimal Dimensional Linear Filters for Stationary Markov Processes with Finite State Space, Stochastics and Stochastic Reports, 36 (1991) pp. 1-19.

[4] J. L. Doob, Stochastic Processes, N.Y., John Wiley and Sons, London-Chapman and Hall, 1953.

[5] F. N. Grigori'ev, N. A. Kuznetsov, and A. P. Serebrovski, Observation Control in Automatic Systems, Nauka, Moscow, 1986 (in Russian).

[6] P. I. Kitsul and R. S. Liptser, On Recurrent Estimation of the Random Sequences, Institute for Control Sciences Press, Moscow, 1974 (in Russian).

[7] A. Lindquist and G. Picci, Realization Theory for Multivariate Stationary Gaussian Processes, SIAM J. Control and Optimization, 23, 6 (1985) pp. 809-857.

[8] R. S. Liptser and A. N. Shiryayev, Statistics of Random processes, Springer-Verlag, Vol. I, 1977, Vol. II, 1978.

[9] G. Marsaglia, Conditional Means and Covariance of Normal Variables with Singular Covariance Matrix, J. Amer. Statist. Assoc., 59, 308 (1964) pp. 1203-1204.