RM-76-41

ON TWO SELECTED TOPICS CONNECTED WITH STOCHASTIC SYSTEMS THEORY

Yu. A. Rozanov

May 1976

Abstract

The existence and the explicit form of the minimal Markov process which contains as a component a given stationary process are established. It is shown in particular that the future/past splitting subspace of the multivariate stationary process is finite-dimensional if and only if the process has a rational spectral densities matrix.

The property of being stochastically continuous is obtained as the condition for continuation of the σ-algebras associated with a process with independent increments which usually represents a stochastic disturbance of the system considered. This property gives us the left-side continuation of the σ-algebras in the case of the arbitrary process in a metric space.

On Two Selected Topics Connected with Stochastic Systems Theory*

Yu. A. Rozanov

I. Markov State Representation

For control of a stochastic system it is desirable to choose a phase-state such that the system evolution in time forms a Markov process. Of course this is not always possible if one is restricted to using a state-space only of a certain type (say, a finite-dimensional vector space); when it is possible, then there are many ways of doing so. In [1] to [4] some relevant problems were considered from the viewpoint of systems theory.

Let us consider the situation in a case of Gaussian stationary processes. Let $x(t) = \{x_\alpha(t)\}$ be an arbitrary family of univariate Gaussian stationary processes $x_\alpha(t)$; $-\infty < t < \infty$, $H^t(x)$ a linear closure of all variables $x_\alpha(t)$; H^{-t} and H^{t+} a linear closure of all subspaces $H^s(x)$, respectively; $s \leq t$, and $H^s(x)$, $s \geq t$, in a Hilbert space of random variables with the usual inner product $(h_1, h_2) = Eh_1h_2$.

Let $p^{-t}(x)$ be the orthogonal projector onto $H^{-t}(x)$. The process $x(t)$ is Markovian if

$$p^{-t}(x)H^{t+}(x) = H^t(x) \quad . \tag{1}$$

Let us say that the Markov process $x(t)$ gives a Markov

---

*This paper was initiated by personal discussion with G. Picci and S.R. Mitter at Padova in December 1975.

state representation for a stationary process y(t) if

$$H^{t}(y) \subseteq H^{t}(x) \quad . \tag{2}$$

In a case of finite dimensional vector processes this means that

$$y(t) = Cx(t) \tag{2'}$$

where C is a constant matrix.

There is a feeling that for any given y(t) there must be an in some sense minimal process x(t) that provides the Markov state representation. How can we describe this minimal Markov process if it exists?

Let us consider a class of all Markov processes x(t) satisfying the condition (2) with the same "innovation" as y(t) namely such that

$$H^{-t}(x) = H^{-t}(y) = H^{-t} \quad . \tag{3}$$

We have

$$H^{t}(y) \subseteq H^{t}(x) \quad , \qquad H^{t+}(y) \subseteq H^{t+}(x)$$

and

$$P^{-t}H^{t+}(y) \subseteq P^{-t}H^{t+}(x) = H^{t}(x)$$

where $P^{-t} = P^{-t}(x) = P^{-t}(y)$. Thus for the Markov process x(t),

$$P^{-t}H^{t+}(y) \subseteq H^{t}(x) \quad , \tag{4}$$

where $P^{-t}H^{t+}(y)$ is the subspace "splitting" the future $H^{t+}(y)$ and the past $H^{-t}(y)$ of the process y(t), $-\infty < t < \infty$.

Let us form a stationary process Y(t), $-\infty < t < \infty$, with

$$H^{t}(Y) = P^{-t}H^{t+}(y) \quad ;$$

say

$$Y(t) = \{V_t h,\ h \in P^{-o}H^{o+}(y)\} \quad , \tag{5}$$

where V_t, $-\infty < t < \infty$, means a family of unitary operators in our Hilbert space, generated by equations

$$x_\alpha(t+s) = V_t x_\alpha(s)\ ;\ -\infty < s,\ t < \infty \quad .$$

Obviously

$$H^t(y) \subseteq H^{-t} \cap H^{t+}(y) \subseteq H^t(Y) \quad , \tag{6}$$

and the stationary process $Y(t)$ has the same innovation as $y(t)$:

$$H^{-t}(Y) = H^{-t} \quad .$$

Let us show that $Y(t)$ is a Markovian process. Indeed, for any $h \in H^o(Y)$ there is $h^t \in H^{o+}(y)$ such that

$$P^{-o}h^+ = h \quad , \qquad h^+ - h \perp H^{-o}$$

and so

$$V_t(h^+ - h) = V_t h^+ - V_t h \perp H^{-t} \quad ,$$

$$H^{-o} \subseteq H^{-t} \quad , \qquad V_t h^+ - V_t h \perp H^{-o} \quad , \qquad t \geq 0 \quad .$$

We have

$$P^{-o}(V_t h^+ - V_t h) = 0 \quad ,$$

$$P^{-o}(V_t h) = P^{-o}(V_t h^+) \in H^o(Y)$$

because for $t \geq 0$

$$V_t h^+ \in H^{o+}(y) \qquad \text{if} \qquad h^+ \in H^{o+}(y) \quad .$$

Thus

$$P^{-o}H^{o+}(Y) = H^o(Y) \quad ,$$

$$P^{-t}H^{t+}(Y) = H^t(Y) \quad ,$$

which we needed to prove.

As a result we obtain the following.

Theorem 1. There is the Markov process Y(t) whose space $H^t(Y)$ coincides with the minimal subspace splitting the future $H^{t+}(y)$ and the past $H^{-t}(y)$ of y(t). This process y(t) gives the minimal Markov state representation for y(t) in the sense that

$$H^t(Y) \subseteq H^t(x)$$

for any other Markov process x(t) if relation (2) holds.

Let us now find a condition for existence of *finite-dimensional* Markov state representation. We consider a finite-dimensional Markov vector process x(t). In a case of discrete time $t = 0, \pm 1, \ldots$ we have

$$x(t) - Ax(t-1) = \sigma u(t) \tag{7}$$

where A, σ are constant matrices of a proper size and u(t) means the corresponding innovation process. The spectral transfer matrix function can be found as

$$\phi_{xu} = (e^{i\lambda} I - A)^{-1}\sigma \quad , \tag{8}$$

and if y(t) = Cx(t) then $\phi_{yu} = C\phi_{xu}$, so the spectral densities matrix $\phi_{yy} = \phi_{yu} \cdot \phi^*_{yu}$ of process y(t) must be a rational function of $z = e^{i\lambda}$.

Similarly in a case of continuous time t we have

$$dx(t) - ax(t) = \sigma du(t) \tag{7'}$$

and

$$\phi_{xu} = (i\lambda I - a)^{-1}\sigma \quad , \tag{8'}$$

so $\phi_{yu} = C\phi_{xu}$ and the spectral density $f_{yy} = \phi_{yu}\phi_{yu}^*$ is a rational function of $z = i\lambda$.

As is well known, for this type of stationary process y(t) there is the explicit finite-dimensional Markov state representation. For example the corresponding Markov process x(t) can be taken as follows. Let

$$t = 0, \pm 1, \ldots$$

and let

$$y(t) = \int_{-\pi}^{\pi} e^{i\lambda t}\phi_{yu}(e^{-i\lambda})d\Phi_u(\lambda)$$

be a spectral representation of y(t) upon its multi-dimensional innovation process u(t) with ϕ_{yu} as the maximal (rational) factor in the factorization

$$\phi_{yu}\phi_{yu}^* = f_{yy}$$

see [5]. Let

$$Q(z) = \sum_{k=0}^{n} q_k z^k$$

be a polynomial with a non-degenerate coefficient q_0(det $q_0 \neq 0$), such that

$$\phi_{yu}Q = \sum_{k=0}^{m} c_k z^k$$

is also a polynomial matrix. Say if ϕ_{yu} is polynomial itself then one can take $Q = I$. Let us set

$$x_o(t) = \int_{-\pi}^{\pi} e^{i\lambda t} Q(e^{-i\lambda})^{-1} d\Phi_u(\lambda)$$

and

$$x(t) = \{x_k(t)\}$$

where

$$x_k(t) = x_0(t - k) \quad ; \quad k = 0,1,\ldots, r-1 \ (r = \max m,n) .$$

Obviously the process x(t) is Markovian because the equation of type (7) holds:

$$x_0(t) - q_0^{-1} \sum_{k=0}^{n-1} q_{k+1} x_k(t-1) = q_0^{-1} u(t)$$

$$x_k(t) - x_{k-1}(t-1) = 0 \quad ; \quad k = 1,\ldots, r-1$$

$$\sum_{k=0}^{m} c_k x_k(t) = \int_{-\pi}^{\pi} e^{i\lambda t} \left[\sum_{k=0}^{m} c_k e^{-i\lambda k}\right] Q(e^{-i\lambda})^{-1} d\Phi_u(\lambda)$$

$$= \int_{-\pi}^{\pi} e^{i\lambda t} \phi_{yu}(e^{-i\lambda}) d\Phi_u(\lambda) = y(t)$$

which gives us t representation (2').

One can proceed similarly in a case of continuous time t.

As an additional result we obtained the following.

Theorem 2. The minimal subspace splitting of the future and the past is finite-dimensional if and only if the process y(t) has the spectral densities matrix with rational components.[1]

---

[1]Compare with [6] where a similar result was obtained for univariate processes in a quite complicated analytical way.

As is well known, for univariate process y(t) with spectral density

$$f = \left|\frac{P}{Q}\right|^2$$

(where $f = \frac{P}{Q}$ is the outer factor) the minimal splitting subspace is generated by functions

$$\frac{e^{i\lambda(t-k)}}{Q} \quad ; \qquad k = 0,\dots,r-1 \tag{9}$$

in a case of discrete time, and

$$\frac{(i\lambda)^k e^{i\lambda t}}{Q} \quad ; \qquad k = 0,\dots,r-1 \tag{9'}$$

for continuous time where r is a maximal degree of the polynomial P, Q (see for example [6]).

The explicit description of minimal splitting subspace for multidimensional processes with rational spectrum still seems to be an open problem. Another open problem concerns analysis of the Markov state representations by means of processes x(t) with different innovations u(t), $-\infty < t < \infty$, i.e., such that

$$H^{-t}(x) \supseteq H^{-t}(y) \quad .$$

(Note that the innovation type and the richness of the past $H^{-t}(x)$ can be characterized completely by the inner factor in the corresponding factorization $f_{yy} = \phi_{yu} \cdot \phi_{yu}^*$.)
Note that some results similar to theorem 1 can be obtained for non-stationary processes.

II. Innovation Continuity

Let x(t) be a random process on an interval of the real line, and $\mathcal{A}_t$ a complete σ-algebra generated by the variables x(s), $s \leq t$. The σ-algebras $\mathcal{A}_t$ grow as t increases. We consider the question of continuity of the $\mathcal{A}_t$ growth, which is quite important for different approaches to stochastic optimal control as well as for the general theory of random processes (see for example [7]).

Let $H_t = L^2(\mathcal{A}_t)$ be the subspace of all random variables h, $Eh^2 < \infty$, measurable with respect to the σ-algebra $\mathcal{A}_t$. It is convenient to treat H_t as the subspace in the Hilbert space of all random variables h, $Eh^2 < \infty$, with the inner product $E(h_1 \cdot h_2)$. Because of the obvious correspondence between $\mathcal{A}_t$ and H_t we consider mainly the family H_t treated as a function of t .

Let us set

$$H_{t-o} = \overline{\bigcup_{s<t} H_s} \quad , \qquad H_{t+o} = \bigcap_{u>t} H_u \quad ;$$

the former means the closure of all subspaces H_s, $s < t$.

We have

$$H_{t-o} \subseteq H_t \subseteq H_{t+o}$$

and there can be a gap between H_{t-o} and H_t as well as between H_t and H_{t+o}. Say this occurs if t is a fixed point of discontinuity of the random process x(·). So considering conditions for the family H_t to be continuous we assume that the process x(t) in a metric phase-space R is stochastically continuous:

$$\lim_{s \to t} P\{\rho(x(s), x(t)) \geq \varepsilon\} = 0 \qquad (10)$$

for any $\varepsilon > 0$; here $\rho(x_1,x_2)$ means the distance between points $x_1, x_2 \in R$.

One can verify that under this assumption

$$H_{t-0} = H_t \quad . \tag{11}$$

To clarify this, let us recall that a probability distribution in a metric space is regular; namely for any measurable set B,

$$\inf_{F \subseteq B} P\,(B \setminus F) = 0$$

where inf is over all closed sets F, $F \subseteq B$. For any closed set F we have

$$\inf_{G \supseteq F} P\,(G \setminus F) = 0$$

where inf is over all open sets G, $G \supseteq F$, with boundaries δG of zero probability $P(\delta G) = 0$. Say one takes the proper

$$G = \{x\colon \rho(x,F) < r\} \quad , \qquad r \to 0 \quad ,$$

with the boundaries

$$\delta G \subseteq \{x\colon \rho(x,F) = r\}$$

which are disjoint for different r; then $P(\delta G) = 0$ except for not more than a countable number of r. Thus any event $\{x(t) \in B\}$ can be approximated by a proper event $\{x(t) \in G\}$ with $P\{x(t) \in \delta G\} = 0$. Event $\{x(t) \in G\}$ itself can be approximated by events $\{x(s) \in G\}$, $s < t$. Indeed,

$$P\{x(t) \in G,\ x(s) \notin G\}$$

$$\leq P\{x(t) \in G \setminus F\} + P\{\rho(x(t),x(s)) \geq \varepsilon\} \to 0 \quad ,$$

if we consequently take

$$F = \{x: \rho(x,R\setminus G) \geq \varepsilon\} \quad ,$$

and $s \to t-o$, $\varepsilon \to o$. Applying this to the other open set $G' = R\setminus(G \cup \delta G)$ with the same boundary $\delta G' = \delta G$, we obtain

$$P\{x(t) \notin G,\ x(s) \in G\} \leq P\{x(t) \in G',\ x(s) \notin G'\} \to 0 \quad .$$

Thus the σ-algebra $\mathscr{A}_t$ generated by the events $\{x(s) \in B\}$, $s \leq t$, coincides with the σ-algebra $\mathscr{A}_{t-o}$ generated by the events $\{x(s) \in B\}$, $s < t$.

Let us now consider the right-continuity of the family H_t:

$$H_{t+o} = H_t \quad . \tag{12}$$

Generally this property does not hold even for a very smooth process $x(t)$; moreover it can be an arbitrary type of discontinuity (see for example [8]). But it holds for a case of stochastically continuous processes with independent increments (which are more and more used in the martingale approach to stochastic optimal control). Apart from the Wiener process case, we don't know where this phenomenon is described, though it looks like one of the classical results of probability theory, being the direct generalization of the famous 0-1 low.

Theorem. For a stochastically continuous process with independent increments,

$$H_{t-o} = H_t = H_{t+o} \quad .$$

The proof of the theorem is based on the following

Lemma. The orthogonal complement $H_u \ominus H_t$ in H_u to the subspace H_t, $t < u$, is a linear closure of variables

$$h = h_t h_{t_1}(A_1) \cdots h_{t_n}(A_n) \quad , \tag{13}$$

where $h_t \in H_t$ and

$$h_{t_k}(A_k) = 1_{A_k} - E1_{A_k}$$

(1_{A_k} are indicators of the events

$$A_k = \{x(t_k) - x(t_{k-1}) \in B_k\}; \ t_o = t < t_1 < \cdots < t_n \leq u) \quad .$$

Indeed the subspace H_u coincides with the linear closure of elements

$$h = 1_{A_1} \cdots 1_{A_n} \in H_t$$

(where A_k are the events of the type $A_k = \{x(t_k) \in B_k\}$; $t_1 < \cdots < t_n \leq t$) together with elements of type (13), and the last are obviously orthogonal to the subspace H_t.

As was actually shown in the case of a stochastically continuous process by the proof of the equation (11), any event

$$A = \{x(t_1) - x(t) \in B\}$$

can be approximated by the event

$$A' = \{x(t_1) - x(t+\delta) \in B\} \quad , \qquad \delta \to +0 \quad ;$$

and therefore any element h of type (13) can be approximated by a similar element h' which is obtained from h by the substitution of the second factor:

$$h_{t_1}(A_1) \to h_{t_1}(A') \quad .$$

Thus any element h_{t+o} from the orthogonal complement $H_{t+o} \ominus H_t$

in H_{t+o} to the subspace H_t as an element of the subspace $H_u \ominus H_t$ can be approximated by the linear combination of the proper elements h'. According to the lemma they belong to the subspaces $H_u \ominus H_{t+\delta}$, so they are orthogonal to $H_{t+o} \subseteq H_{t+\delta}$. Because the element h_{t+o} belongs to H_{t+o} we conclude that $h_{t+o} = 0$ and thus $H_{t+o} = H_t$.

References

[1] Kalman, R. E., "Linear Stochastic Filtering Theory - Reappraisal and Outlook," in *Proceedings of Symposium on System Theory*, Polytechnic Institute of Brooklyn (1965), pp. 197-205.

[2] Faurre, P., "Réalisations Markoviennes de processus stationnaires," IRIA Report No. 13, Institut de Recherche d'Informatique et d'Automatique, Le Chesnay, 1973.

[3] Akaike, H., "Markovian Representation of Stochastic Processes by Canonical Variables," *SIAM J. on Control*, 13, 1 (1975), 162.

[4] Akaike, H., "Stochastic Theory of Minimal Realization," *IEEE Transactions on Automatic Control*, 19, 6 (1974), 667.

[5] Rozanov, Yu. A., *Stationary Random Processes* (in Russian), Fismatgiz, Moscow, 1963; English edition Holden-Day, San Francisco, California, 1967.

[6] Ibragimov, I. A., and Yu. A. Rozanov, *Gaussian Random Processes*, Nauka, Moscow, 1970.

[7] Liptser, Y., and Schiryaev, *Statistics of Random Processes* (in Russian), Nauka, Moscow, 1974.

[8] Rozanov, Yu. A., *The Theory of Innovative Processes* (in Russian), Nauka, Moscow, 1973.