stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE BESLISKUNDE BW 141/81 MEI
(DEPARTMENT OF OPERATIONS RESEARCH)

J.H. VAN SCHUPPEN

THE PROBABILISTIC REALIZATION PROBLEM FOR FINITE DIMENSIONAL GAUSSIAN RANDOM VARIABLES

Preprint

kruislaan 413 1098 SJ amsterdam

Printed at the Mathematical Centre, 413 Kruislaan, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

1980 Mathematics subject classification: 93E03, 60G05

The probabilistic realization problem for finite dimensional Gaussian random variables*)

by

J.H. van Schuppen

ABSTRACT

A classification is given of all σ-algebra's that make two given σ-algebra's conditional independent in the case that the σ-algebra's are generated by finite dimensional Gaussian random variables.

KEY WORDS & PHRASES: *Gaussian random variables, realization problem*

*) This report will be submitted for publication elsewhere.

To appear in Proc. European Conference on Circuit Theory and Design 1981, The Hague, The Netherlands.

1. INTRODUCTION

The purpose of this note is to present the solution of the probabilistic realization problem for σ-algebra's generated by finite dimensional Gaussian random variables.

The stochastic realization problem in stochastic system theory is to construct stochastic dynamical system representations for stochastic processes. There is a growing literature on this subject [2,6, 7,10], mainly for Gaussian processes. The problem is still not satisfactorily solved. One open question in the Gaussian case is the explicit classification of all minimal stochastic realizations. In a static setting the stochastic realization problem reduces to the probabilistic realization problem to be formulated below. In this note this problem will be resolved. The solution given may provide insight in the classification of minimal Gaussian stochastic realizations.

The main concept in stochastic realization theory, as shown in [7,10], is the conditional independence relation for σ-algebra's. This relation is a key property in many area's of probability theory and stochastic processes. Examples of such area's are sufficient statistics, Markov processes, information theory, random fields, and stochastic system theory.

What is the problem? Assume given two jointly Gaussian random variables and consider the σ-algebra's that they generate. One may ask for all the σ-algebra's that make the given two σ-algebra's conditional independent. To exclude some trivial answers the concept of a minimal σ-algebra must be introduced. The probabilistic realization problem is then to show existence of σ-algebra's that make two given σ-algebra's minimal conditional independent, to classify all such σ-algebra's, and to develop an algorithm that constructs these σ-algebra's. The contribution of this paper is the solution of this problem.

One may also define a weak probabilistic realization problem, where the underlying probability space may be constructed. This problem is different from the probabilistic realization problem, although they coincide under certain conditions. A still open problem is the probabilistic realization in the case the σ-algebra's are arbitrary, or not necessarily generated by Gaussian random variables.

The approach of the paper is a mixture of probabilistic and geometric analysis. The main objects of the paper are σ-algebra's generated by finite dimensional Gaussian random variables. From Neveu [8] it is clear that a Hilbert space framework may be used in this case. This approach has been followed in [7]. However this line of work is insufficient for the problem to be considered here. Because of the restriction to σ-algebra's generated by finite dimensional Gaussian random variables a more explicit classification may be obtained. Yet our approach will be very much in a geometric spirit emphasizing the spaces and working basis free as much as possible. Not all proofs will be given here; they are deferred to a future publication.

A brief summary of the paper follows. The problem formulation is given in the next section, while some preliminaries are presented in Section 3. The probabilistic realization problem is resolved in section 4.

2. PROBLEM FORMULATION

In this section some notation is introduced and the problem defined.

In this note (Ω,F,P) denotes a complete probability space consisting of a set Ω, a σ-algebra F, and a probability measure P. Let

$$\underline{F} = \{G \subset F \mid G \text{ a } \sigma\text{-algebra completed with all null sets of } F\},$$

and for $G \in \underline{F}$.

$$L^+(G) = \{x: \Omega \to R_+ \mid x \text{ is } G \text{ measurable}\}.$$

If $y: \Omega \to R^k$ is a random variable then $F^y = \sigma(\{y\}) \in \underline{F}$ is the σ-algebra generated by y. If $F_1, F_2 \in \underline{F}$ then $F_1 \vee F_2$ denotes the smallest σ-algebra that contains both F_1 and F_2. The notation $(F_1,F_2) \in I$ is used to indicate that F_1, F_2 are independent σ-algebra's.

2.1. DEFINITION. The *conditional independence relation* for a triple of σ-algebra's $F_1,F_2,G \in \underline{F}$ is defined by the condition that for all $y_1 \in L^+(F_1)$, $y_2 \in L^+(F_2)$

$$E[y_1 y_2 | G] = E[y_1|G]\, E[y_2|G].$$

Equivalently, if for all $y_1 \in L^+(F_1)$

$$E[y_1|F_2 \vee G] = E[y_1|G].$$

Then one says that F_1, F_2 are conditional independent given G, or that G splits F_1, F_2. Notation: $(F_1,G,F_2) \in CI$. □

The equivalence follows from [1, II.45].

Some notation is introduced. Let

$\mathbb{Z}_+ = \{1,2,3,\ldots\}$, $\quad N = \{0,1,2,\ldots\}$,

and for $n \in Z_+$ let

$Z_n = \{1,2,\ldots,n\}$, $\quad N_n = \{0,1,2,\ldots,n\}$.

If $Q \in R^{n\times n}$ then Q^T denotes the transposed of Q, $Q \geq 0$ that Q is positive definite, and $Q > 0$ that Q is strictly positive definite.

A finite dimensional Gaussian random variable with parameters $n \in Z_+$, $\mu \in R^n$, $Q \in R^{n\times n}$, satisfying $Q = Q^T \geq 0$, is a random variable $x: \Omega \to R^n$ such that for all $u \in R^n$

$$E[\exp(iu^Tx)] = \exp(iu^T\mu - \frac{1}{2} u^TQu).$$

Notation: $x \in G(\mu,Q)$; $(x_1,\ldots,x_m) \in G(\mu,Q)$ denotes that with $x^T = (x_1^T,\ldots,x_m^T)$, $x \in G(\mu,Q)$. If $x \in G$, then Q_{xx} may denote its covariance matrix.

2.2. DEFINITION. The *Gaussian conditional independence relation* for a triple of σ-algebra's $F^{y_1},F^x,F^{y_2} \in \underline{F}$, generated by

$$y_1: \Omega \to R^{k_1}, \quad y_2: \Omega \to R^{k_2}, \quad x: \Omega \to R^n$$

is defined by the conditions

1. $(F^{y_1},F^x,F^{y_2}) \in CI$;
2. $(y_1,x,y_2) \in G$.

Notation: $(F^{y_1},F^x,F^{y_2}) \in CIG$. □

Given $(y_1,y_2) \in G$ there exists random variables x such that $(F^{y_1},F^x,F^{y_2}) \in CIG$. For example $x = y_1$, or $x = y_2$, are such random variables. From many viewpoints it is of interest to ask for a minimal σ-algebra.

2.3. DEFINITION. The *minimal Gaussian conditional independence relation* for a triple of σ-algebra's $F^{y_1},F^{y_2},F^x \in \underline{F}$ generated by

$$y_1: \Omega \to R^{k_1}, \quad y_2: \Omega \to R^{k_2}, \quad x: \Omega \to R^n$$

is defined by the conditions

1. $(F^{y_1},F^x,F^{y_2}) \in CIG$;
2. if $F^{x_1} \in \underline{F}$, $F^{x_1} \subset F^x$, $(y_1,y_2,x,x_1) \in G$, $(F^{y_1},F^{x_1},F^{y_2}) \in CIG$, then $F^{x_1} = F^x$.

Then one says that F^x makes F^{y_1}, F^{y_2} minimal conditional independent, or that F^x is a minimal splitter of F^{y_1},F^{y_2}. Notation: $(F^{y_1},F^x,F^{y_2}) \in CIG_{min}$. □

2.4. PROBLEM. The *Gaussian probabilistic realization problem* for a triple of Gaussian random variables (y_1,y_2,v) is:

(a) *to show existence* of triples (R^n,B_n,F^x), where $x: \Omega \to R^n$, such that

 1. $(F^{y_1},F^x,F^{y_2}) \in CIG_{min}$;

 2. $F^x \subset F^{y_1} \vee F^{y_2} \vee F^v$ and $(y_1,y_2,v,x) \in G$;

 such a triple will then be called a minimal probabilistic realization;

(b) *to classify* all minimal probabilistic realizations;

(c) *to develop an algorithm* that constructs all minimal probabilistic realizations. □

Finally some additional notation for matrices is introduced. Let

$$D_n = \{A \in R^{n\times n} \mid A \text{ a diagonal matrix}\},$$

$$D_n^+ = \{A \in D_n \mid A \geq 0\},$$

$$O_n = \{S \in R^{n\times n} \mid SS^T = I = S^TS\},$$

the set of orthogonal matrices. For $A \in R^{n\times n}$ let

$$C_n(A) = \{(S_1,S_2) \in O_n \times O_n \mid S_2^TS_1A = AS_2^TS_1\}.$$

It is easily verified that $C_n(A)$ is an equivalence relation, with $S_1 \sim S_2$ iff $(S_1,S_2) \in C_n(A)$. The quotient space $O_n/C_n(A)$ is thus well-defined. The class of matrices that commute with a given matrix is described in [2, 1.VIII2].

3. PRELIMINARIES

In this section the canonical variable representation for Gaussian random variables is introduced. Furthermore an equivalent condition for CIG_{min} is derived.

To describe the relationship between two random variables Hotelling [4] has introduced the concept of a canonical variable representation. For Gaussian random variables this representation has a rather explicit structure that is stated below.

3.1. DEFINITION. Given $y_1: \Omega \to R^{k_1}$, $y_2: \Omega \to R^{k_2}$, $(y_1,y_2) \in G(0,K)$. These random variables are said to be in *canonical variable form* if

$$K = \left(\begin{array}{ccc|ccc} I & & & I & & \\ & I & & & \Lambda & \\ & & I & & & 0 \\ \hline I & & & I & & \\ & \Lambda & & & I & \\ & & 0 & & & I \end{array}\right) \in R^{(k_1+k_2)\times(k_1+k_2)},$$

where $\Lambda \in D_{k_{12}}^+$, $\Lambda = \text{diag}(\lambda_1,\lambda_2,\ldots,\lambda_{k_{12}})$ with $1 > \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{k_{12}} > 0$. Compatible with this decomposition let

$$y_1^T = (y_{11}^T,y_{12}^T,y_{13}^T), \quad y_{11}: \Omega \to R^{k_{11}}, \quad y_{12}: \Omega \to R^{k_{12}},$$

$$y_{13}: \Omega \to R^{k_{13}},$$

and similarly

$$y_2^T = (y_{21}^T,y_{22}^T,y_{23}^T).$$

Furthermore let

$$W = \begin{pmatrix} I & & \\ & \Lambda & \\ & & 0 \end{pmatrix} \in R^{k_1\times k_2}, \quad \Sigma = \begin{pmatrix} I & \Lambda \\ \Lambda & I \end{pmatrix} \in R^{2k_{12}\times 2k_{12}}. \ \square$$

Note that $y_{11} = y_{21}$ a.s., and $k_{12} = k_{22}$.

It is a classical result [8] that for any pair $(z_1,z_2) \in G(0,K_1)$ there exists a basis transformation $(z_1,z_2) \mapsto (S_1z_1,S_2z_2)$ such that (S_1z_1,S_2z_2) is in canonical form. Such a transformation is unique up to the equivalence relation $(S_1,S_2) \sim (V_1,V_2)$ defined by $V_1^TS_1W = WS_2^TV_2$. On the basis of the canonical variable representation one may formulate a canonical form for Gaussian measures.

The problem posed in 2.4 is the construction and classification of σ-algebra's that make two given σ-algebra's minimal Gaussian conditional independent. This problem is analogeous to the construction of realizations in linear system theory. There it is known that a dynamical system has a state space of minimal dimension iff the dynamical system is observable and controllable. Furthermore all minimal realizations are equivalent in a well defined sense. What remains of this picture in probabilistic realization? The concept of probabilistic observability will be defined first.

3.2. DEFINITION. Given $(F^{y1},F^{x},F^{y2}) \in CIG$. This triple will be called *probabilistic observable* if the map

$$x \mapsto E[\exp(iu^T y_1) \mid F^x]$$

is injective on the support of x. It will be called *probabilistic reconstructible* iff the map

$$x \mapsto E[\exp(iu^T y_2) \mid F^x]$$

is injective on the support of x. □

Suppose that through multiple experiments one is able to obtain an estimate of the measure of y_1 for a fixed value of x. Then probabilistic observability implies that from this measure one can determine the value of the state x uniquely. This property hopefully motivates the above definition of probabilistic observability.

With $(y_1,x,y_2) \in G(0,Q)$ and a basis for x such that $Q_{xx} > 0$ one has

$$E[\exp(iu^T y_1) \mid F^x]$$

$$= \exp(iu^T Q_{y_1x}Q_{xx}^{-1}x - \tfrac{1}{2}\, u^T[Q_{y_1y_1} - Q_{y_1x}Q_{xx}^{-1}Q_{xy_1}]u).$$

Thus $(F^{y1},F^{x},F^{y2}) \in CIG$ is probabilistic observable iff $\mathrm{rank}(Q_{y_1x}) = \mathrm{rank}(Q_{xx})$. The following result is now motivated.

3.3. THEOREM. Given

$$y_1: \Omega \to R^{k_1}, \quad y_2: \Omega \to R^{k_2}, \quad x: \Omega \to R^n.$$

The following are equivalent:

(a) $(F^{y1},F^{x},F^{y2}) \in CIG_{min}$;

(b) 1. $(y_1,x,y_2) \in G$;

2. $Q_{y_1y_2} = Q_{y_1x}Q_{xx}^{-1}Q_{xy_2}$;

3. $\mathrm{rank}(Q_{y_1x}) = \mathrm{rank}(Q_{xx}) = \mathrm{rank}(Q_{y_2x})$.

Here it has been assumed that a basis has been chosen such that $Q_{xx} > 0$. □

The proof of 3.3 is based on the following intermediate results.

3.4. PROPOSITION. Given

$$y_1: \Omega \to R^{k_1}, \quad y_2: \Omega \to R^{k_2}, \quad x: \Omega \to R^n$$

$(y_1,y_2,x) \in G$. Suppose that a basis for x has been chosen such that $Q_{xx} > 0$. Then the following are equivalent:

(a) $(F^{y1},F^{x},F^{y2}) \in CIG$;

(b) $Q_{y_1y_2} = Q_{y_1x}Q_{xx}^{-1}Q_{xy_2}$.

PROOF. This is a calculation via the conditional characteristic function. □

3.5. PROPOSITION. Given

$$y_1: \Omega \to R^{k}, \quad y_2: \Omega \to R^{k_2}, \quad x: \Omega \to R^n$$

with $(y_1,y_2,x) \in G$. If $(F^{y1},F^{x},F^{y2}) \in CIG$

$$x_1: \Omega \to R^{k_1} \quad x_1 = E[y_1 \mid F^x],$$

$$x_1: \Omega \to R^{k_2} \quad x_2 = E[y_2 \mid F^{x_1}],$$

then $F^{y1},F^{x2},F^{y2}) \in CIG$,

$$\mathrm{rank}(Q_{y_1x}) = \mathrm{rank}(Q_{xx}) = \mathrm{rank}(Q_{y_2x}).$$

PROOF. Not given here. □

3.6. PROPOSITION. Given

$$y_1: \Omega \to R^{k_1}, \quad y_2: \Omega \to R^{k_2} \quad (y_1,y_2) \in G(0,K)$$

with K as given in 3.1. The notation of 3.1. is adopted. Then $(F^{y1},F^{x},F^{y2}) \in CIG_{min}$ iff there exists a basis for x such that

$$x^T = (x_1^T, x_2^T), \quad x_1: \Omega \to R^{k11}, \quad x_2: \Omega \to R^{k12},$$

$$(F^{y11},F^{x1},F^{y21}) \in CIG_{min} \text{ and } (F^{y12},F^{x2},F^{y22}) \in CIG_{min}.$$

PROOF. Not given here. □

4. THE PROBABILISTIC REALIZATION PROBLEM

In this section the probabilistic realization problem is resolved.

4.1. DEFINITION. Given a complete probability space (Ω,F,P) and three Gaussian random variables defined on it

$$y_1: \Omega \to R^{k_1}, \quad y_2: \Omega \to R^{k_2}, \quad v: \Omega \to R^m$$

with $(y_1,y_2,v) \in G(0,L)$. Let the *set of probabilistic realizations* be

$$PR(R^{k_1+k_2+m}, B_{k_1+k_2+m}, G(0,L))$$

$$= \{(R^n,B_n),\ F^x \in \underline{F} \mid x: \Omega \to R^n,\ F^x = \sigma(\{x\}),$$

$$(F^{y1},F^{x},F^{y2}) \in CIG_{min},\ (y_1,y_2,v,x) \in G,$$

$$F^x \subset F^{y1} \vee F^{y2} \vee F^v\}.$$

In the above definition v represents additional information on which the minimal σ-algebra may be based. It is clear that for an arbitrary Gaussian random variable w, representing the additional information, one may construct a Gaussian random variable v such that $F^v \subset F^w$ and $(F^v,F^{y1} \vee F^{y2}) \in I$.

4.2. THEOREM. Given a complete probability space (Ω,F,P) with three Gaussian random variables defined on it

$$y_1: \Omega \to R^{k_1}, \quad y_2: \Omega \to R^{k_2}, \quad v: \Omega \to R^m$$

$(y_1,y_2,v) \in G(0,L)$ where

$$L = \begin{pmatrix} K & 0 \\ 0 & I_m \end{pmatrix}$$

with K as given in 3.1; the notation of 3.1 is adopted. Let

$$PA = \{(n_1,n_2,n_3) \in N^3,\ b \in (0,1)^{n_3},\ S \in O_{k_{12}},\ H \in R^{n_3 \times m} \mid$$

if $B = \mathrm{diag}(b_1,\ldots,b_{n_3})$, decreasingly ordered,

$k_{12} = n_1+n_2+n_3$, $U = \mathrm{blockdiag}(I_{n_1},B,0_{n_2}) \in$

$\in R^{k_{12}\times k_{12}}$, then $S \in O_{k_{12}}/C_{k_{12}}(U)$, $HH^T = I\}$.

Define the map

$$r\colon PA \to PR(R^{k_1+k_2+m}, B_{k_1+k_2+m}, G(0,L))$$

as $(n_1,n_2,n_3,b,S,H) \in PA$, $U \in R^{k_{12}\times k_{12}}$ as constructed in the definition of PA, $n = k_{11} + k_{12}$,

$$A = (\Lambda^{-1}-\Lambda)^{-1},\ A^{\frac{1}{2}} \in D^{+}_{k_{12}},$$

$$P_1 = A^{-\frac{1}{2}}S(I-U)S^T\Lambda^{\frac{1}{2}}A^{\frac{1}{2}},$$

$$P_2 = A^{-\frac{1}{2}}SUS^T\Lambda^{-\frac{1}{2}}A^{\frac{1}{2}},$$

$$P_3 = A^{-\frac{1}{2}}S(U-U^2)\begin{pmatrix}0\\H\\0\end{pmatrix}$$

$$x\colon \Omega \to R^n,\ x = \begin{pmatrix} y_{11} \\ P_1y_{12}+P_2y_{22}+P_3v\end{pmatrix},\ F^x = \sigma(\{x\}),$$

$$(R^n, B_n, F^x) \in PR.$$

Then, with respect to the given basis for (y_1,y_2,v), r is well defined and a bijection. □.

The reader is reminded of the fact that the transformation to the canonical variable representation is nonunique. Hence the bijection part of 4.2 is valid only with respect to the given basis.
The result 4.2 resolves the probabilistic realization problem 2.4 in that it classifies all minimal realizations and provides an algorithm to construct these realizations.

The proof of 4.2 is based on the following lemma which treats of special case of 4.2. This special case is motivated by 3.6.

4.3. LEMMA. Given three Gaussian random variables $y_1\colon \Omega \to R^k$, $y_2\colon \Omega \to R^k$, $v\colon \Omega \to R^m$ $(y_1,y_2,v) \in G(0,L)$

$$L = \begin{pmatrix}\Sigma & 0\\ 0 & I_m\end{pmatrix}$$

with Σ as defined in 3.1. Let

$$PA_1 = \{Q \in R^{k\times k}, P_3 \in R^{k\times m} \mid A = (\Lambda^{-1}-\Lambda)^{-1}, Q = Q^T \geq 0,\ Q + Q\Lambda A + A\Lambda Q - QAQ - A = P_3P_3^T\}.$$

Define the map $r_1\colon PA_1 \to PR(R^{2k+m}, B_{2k+m}, G(0,L))$ as $(Q,P_3) \in PA_1$,

$$P_1 = (I-Q\Lambda)\Lambda^{\frac{1}{2}}(I-\Lambda^2)^{-1},$$

$$P_2 = (Q-\Lambda)\Lambda^{\frac{1}{2}}(I-\Lambda^2)^{-1},$$

$$x\colon \Omega \to R^k,\ x = P_1y_1 + P_2y_2 + P_3v,\ F^x = \sigma(\{x\}),$$

$$(R^k, B_k, F^x) \in PR.$$

Then, with respect to the given basis for (y_1,y_2,v), the map r_1 is well defined and a bijection. □

Some calculations used in the proof are summarized below.

4.4. PROPOSITION. Given the matrix $\Lambda \in R^{n\times n}$, of the form as presented in 3.1, $Q \in R^{n\times n}$, and

$$L = \begin{pmatrix} I & \Lambda & \Lambda^{\frac{1}{2}} \\ \Lambda & I & \Lambda^{\frac{1}{2}}Q \\ \Lambda^{\frac{1}{2}} & Q\Lambda^{\frac{1}{2}} & Q\end{pmatrix} \in R^{3n\times 3n}.$$

Assume that $Q = Q^T$. The following are equivalent:
a. $L \geq 0$;
b. $Q \in \underline{Q} = \{Q \in R^{n\times n} \mid Q = Q^T \geq 0,\ A := (\Lambda^{-1}-\Lambda)^{-1},\ Q + Q\Lambda A + A\Lambda Q - QAQ - A \geq 0\}$;
c. $\Lambda \leq Q \leq \Lambda^{-1}$.

PROOF. With Σ as defined in 3.1,

$$\Sigma^{-1} = \begin{pmatrix}(I-\Lambda^2)^{-1} & -A \\ -A & (I-\Lambda^2)^{-1}\end{pmatrix}.$$

Elementary row and column operations now yield that $L \geq 0$ iff

$$Q \geq \Lambda \quad\text{and}\quad I - \Lambda^{\frac{1}{2}}Q\Lambda^{\frac{1}{2}} \geq 0,$$

iff

$$Q - (\Lambda^{\frac{1}{2}}\ Q\Lambda^{\frac{1}{2}})\Sigma^{-1}\begin{pmatrix}\Lambda^{\frac{1}{2}}\\ \Lambda^{\frac{1}{2}}Q\end{pmatrix} \geq 0.$$

A calculation then gives the result. □

PROOF OF 4.3. 1. With A and Σ as given above

$$(P_1\ P_2) = (\Lambda^{\frac{1}{2}}\ Q\Lambda^{\frac{1}{2}})\Sigma^{-1},$$

$$(P_1\ P_2)\Sigma\begin{pmatrix}P_1^T\\ P_2^T\end{pmatrix} = QAQ + A - Q\Lambda A - A\Lambda Q.$$

2. To show that r_1 is well defined let $(Q,P_3) \in PA_1$. Then F^x is well defined. It is then a calculation to show that $Q_{y_1x} = \Lambda^{\frac{1}{2}}$, $Q_{y_2x} = \Lambda^{\frac{1}{2}}Q$, and using an equality of 1 above and $(Q,P_3) \in PA_1$,

$$Q_{xx} = (P_1P_2)\Sigma\begin{pmatrix}P_1^T\\ P_2^T\end{pmatrix} + P_3P_3^T = QAQ + A - Q\Lambda A - A\Lambda Q + P_3P_3^T = Q.$$

Now $Q = Q^T \geq 0$ and $Q + Q\Lambda A + A\Lambda Q - QAQ - A = P_3P_3^T \geq 0$ imply by 4.4 that $Q \geq \Lambda > 0$. Then

$$Q_{y_1y_2} = \Lambda = \Lambda^{\frac{1}{2}}Q^{-1}Q\Lambda^{\frac{1}{2}} = Q_{y_1x}Q_{xx}^{-1}Q_{xy_2},$$

$$\operatorname{rank}(Q_{y_1x}) = \operatorname{rank}(Q_{xx}) = \operatorname{rank}(Q_{y_2x}),$$

hence by 3.3 $(F^{y_1}, F^x, F^{y_2}) \in CIG_{min}$.

3. It will be shown that r_1 is surjective. Let $(R^n, B_n, F^x) \in PR$. Because

$$(F^{y_1}, F^x, F^{y_2}) \in CIG_{min}$$

and 3.3 one may choose a basis for x of dimension $k = \operatorname{rank}(Q_{y_1y_2})$, hence $\operatorname{rank}(Q_{xx}) = k$ and by 3.3 $\operatorname{rank}(Q_{y_1x}) = k$. Take the basis transformation

$x_1 = \Lambda^{\frac{1}{2}}Q_{y_1x}^{-1}x$ and set $Q = Q_{x_1x_1}$. Then it follows that $Q_{y_1x_1} = \Lambda^{\frac{1}{2}}$, $Q_{y_2x_1} = \Lambda^{\frac{1}{2}}Q$. Furthermore

$$E[x_1 \mid F^{y_1} \vee F^{y_2}] = (\Lambda^{\frac{1}{2}}\ Q\Lambda^{\frac{1}{2}})\Sigma^{-1}\begin{pmatrix}y_1\\ y_2\end{pmatrix} = P_1y_1 + P_2y_2,$$

with 1 above. Let $z\colon \Omega \to R^k$, $z = x - P_1y_1 - P_2y_2$.

Then

$F^z \subset F^{y_1} \vee F^{y_2} \vee F^v$, z is independent of (y_1,y_2), and $(y_1,y_2,v,z) \in G$, hence there exists a $P_3 \in R^{k\times m}$ such that $z = P_3v$ a.s. Furthermore

$$P_3P_3^T = E[zz^T] = Q - (P_1\ P_2)\Sigma\begin{pmatrix}P_1^T\\ P_2^T\end{pmatrix} = Q + Q\Lambda A + A\Lambda Q - QAQ - A.$$

Thus $(Q,P_3) \in PA_1$.

4. To establish that r_1 is injective let

$$(R^{n_1}, B_{n_1}, F^{x_1}) = (R^{n_2}, B_{n_2}, F^{x_2}) \in PR.$$

Because r is surjective one may associate with these (Q_1,P_{13}), $(Q_2,P_{23}) \in PA_1$. Because

$$F^{x_1} = F^{x_2} \subset (F^{y_1} \vee F^{y_2} \vee F^v),$$

$$(y_1,y_2,v,x_1) \in G,\ (y_1,y_2,v,x_2) \in G,$$

there exists a nonsingular $S \in R^{k\times k}$ such that $x_2 = Sx_1$.

Let

$x_1 = P_{11}y_1 + P_{12}y_2 + P_{13}v,$

$x_2 = P_{21}y_1 + P_{22}y_2 + P_{23}v$. Then

$P_{21}y_1 + P_{22}y_2 + P_{23}v = SP_{11}y_1 + SP_{22}y_2 + SP_{23}v,$

$(SP_{11}-P_{21}) = -(SP_{12}-P_{22})\Lambda,$

$(SP_{11}-P_{21})\Lambda = -(SP_{12}-P_{22}),\ P_{23} = SP_{13}.$

Using the expressions for P_{11}, P_{12}, P_{21}, P_{22} and performing some calculations one obtains

$SP_{11} - P_{21} = [(S-I)-(SQ_1-Q_2)\Lambda]\Lambda^{1/2}(I-\Lambda^2)^{-1},$

$SP_{12} - P_{22} = [(SQ_1-Q_2)-(S-I)\Lambda]\Lambda^{1/2}(I-\Lambda^2)^{-1},$

$(S-I)(1-\Lambda^2) = 0$ or $S = I$,

$(Q_1-Q_2)(I-\Lambda^2) = 0$ or $Q_1 = Q_2$, $P_{23} = P_{13}$. □

PROOF OF 4.2. 1. That r is well defined follows from a calculation and 3.3.

2. To show that r is surjective let $(R^n, B_n, F^x) \in PR$. From 3.6 it follows that there exists a basis for x such that

$x^T = (x_1^T, x_2^T),\ x_1: \Omega \to R^{k11},\ x_2: \Omega \to R^{k12},$

$(F^{y11}, F^{x_1}, F^{y21}) \in CIG_{min},$

$(F^{y12}, F^{x_2}, F^{y22}) \in CIG_{min},\ F^x = F^{x_1} \vee F^{x_2}.$

It is an exercise to show that

$(F^{y11}, F^{x_1}, F^{y21}) \in CIG_{min}$

implies that with respect to some basis $x_1 = y_{11} = y_{21}$ a.s. From $(F^{y12}, F^{x_2}, F^{y22}) \in CIG_{min}$ and 4.3 one obtains that there exists a $(Q, P_3) \in PA_1$ such that

$r_1(Q, P_3) = (R^{k12}, B_{k12}, F^{x_2})$. Then

$Q + Q\Lambda A + A\Lambda Q - QAQ - A = P_3P_3^T$. Using the transformation

$M = A^{1/2}(Q-\Lambda)A^{1/2} \in R^{k12\times k12}$ one obtains that

$A^{-1/2}(M-M^2)A^{-1/2} = P_3P_3^T \geq 0,\ M = M^T \geq 0$, or $M^2 \leq M$
$M = M^T \geq 0$.

Hence there exists $U \in D^+_{k12}$, $S_1 \in O_{k12}$ such that $M = S_1US_1^T$. By convention the diagonal elements of U are chosen in decreasing order. Let $S \in O_{k12}/C_{k12}(U)$ be the element corresponding to $S_1 \in O_{k12}$. Let $n_1, n_3, n_2 \in N$ be respectively the number of diagonal elements of U in {1},(0,1),{0}. Let $B \in D^+_{n_3}$ be the diagonal matrix corresponding to the elements in (0,1) in U. Let

$V_1 \in R^{n_3\times m},\ V_2 \in R^{n_3\times m},\ V_3 \in R^{n_2\times m},$

$\begin{pmatrix}V_1\\V_2\\V_3\end{pmatrix} = S^TA^{1/2}P_3$. Then

$P_3P_3^T = A^{-1/2}(M-M^2)A^{-1/2} = A^{-1/2}S(U-U^2)S^TA^{-1/2}$ implies that

$V_1 = 0,\ V_3 = 0$, and $V_2V_2^T = (B-B^2)$.

Let $(B-B^2)^{1/2} \in D^+_{n_3}$, $H = (B-B^2)^{-1/2}V_2 \in R^{n_3\times m}$, then $HH^T = I$.

3. Finally it will be shown that r is injective. Let

$(R^{n_1}, B_{n_1}, F^{x_1}) = (R^{n_2}, B_{n_2}, F^{x_2}) \in PR.$

As proven in 2 above r is surjective so there exist $(n_{11},n_{12},n_{13},b_1,S_1,H_1), (n_{21},n_{22},n_{23},b_2,S_2,H_2) \in PA$ corresponding to these elements. As mentioned in 2 above there exist $(Q_1,P_{13}) = (Q_2,P_{23}) \in PA_1$, equal because of 4.3. Note that

$A^{-1/2}S_1U_1S_1^TA^{-1/2}+\Lambda = Q_1 = Q_2 = A^{-1/2}S_2U_2S_2^TA^{-1/2} + \Lambda,$

$S_1U_1S_1^T = S_2U_2S_2^T.$

Because of the ordering of the diagonal elements of U_1, U_2 one obtains that $U_1 = U_2$. Then

$S_2^TS_1U_1 = U_1S_2^TS_1,\ S_1, S_2 \in O_{k12}/C_{k12}(U_1)$

imply that $S_1 = S_2$. Finally

$A^{-1/2}S_1(U_1-U_1^2)\begin{pmatrix}0\\H_1\\0\end{pmatrix} = P_{13} = P_{23} = A^{-1/2}S_1(U_1-U_1^2)\begin{pmatrix}0\\H_2\\0\end{pmatrix}$

imply that $H_1 = H_2$. □

REFERENCES

[1] Dellacherie, C. and P.A. Meyer, "Probabilités et potentiel", Ch. I-IV, Hermann, Paris, 1975.

[2] Faurre, P., M. Clerget and F. Germain, "Opérateurs rationnels positifs", Dunod, Paris, 1979.

[3] Gantmacher, F.R., "The theory of matrices", volume 1 and 2, Chelsea, New York, 1959.

[4] Hotelling, H., "Relations between two sets of variates", Biometrika 28 (1936), pp.321-377.

[5] Kalman, R.E., P.L. Falb, and M.A. Arbib, "Topics in mathematical system theory", McGraw-Hill, New York, 1969.

[6] Lindquist, A. and G. Picci, "State space models for Gaussian stochastic processes", in: "Stochastic Systems: The Mathematics of Filtering and Identification and Applications", M. Hazewinkel, J.C. Willems (eds), Reidel Publ. Co., Dordrecht, to appear.

[7] Lindquist, A., G. Picci and G. Ruckebusch, "On minimal splitting subspaces and Markovian representations", Math. Systems Theory 12 (1979), pp.271-279.

[8] Neveu, J., "Processus aléatoires Gaussiens", Les Presses de l'Université de Montréal, Montréal, 1968.

[9] Rao, C.R., "Linear statistical inference and its applications", Wiley, New York, 1973.

[10] Willems, J.C., "System theoretic models for the analysis of physical systems", Ricerche di Automatica,