LINEAR COMPLETENESS
IN A CONTINUOUS TIME GAUSS-MARKOV MODEL

P. Ibarrola

University Complutense of Madrid

email: ibarrola@mat.ucm.es

and

A. Pérez-Palomares*

University of Zaragoza

email: anapp@unizar.es

Short title: Linear completeness in continuous time.

Abstract

In this paper, we study linear completeness in a continuous time linear model. We give a characterization of this property and we show its equivalence with ordinary completeness when a Gaussian process is considered. Furthermore, a characterization of a sufficient and complete estimator in a continuous time Gaussian model is given.

AMS Subject Classifications: 62M09; 62J05.

Key words and phrases: Gauss-Markov model; continuous time; linear completeness; linear sufficiency.

---

* Research supported by the BFM2001-2499 and DGA-P1192001 Grants. This author is a member of the research group: Grupo consolidado "Modelos Estocásticos" de la Diputación General de Aragón.

Mailing address for manuscript correspondence: Ana Pérez Palomares. Departamento de Métodos Estadísticos. Facultad de Ciencias. Edificio de Matemáticas. Pedro Cerbuna 12. Universidad de Zaragoza. 50009 Zaragoza (Spain). email:anapp@unizar.es

1. Introduction

In the classical linear model $\{\mathbf{Y}, \mathbf{X}\beta, \mathbf{V}\}$, $\mathbf{Y}$ is a random vector with expectation $\mathbf{X}\beta$, β is a vector of unknown parameters and $\mathbf{V}$ is the dispersion matrix which is partially or totally known. For this model, Baksalary and Kala (1981) and Drygas (1983) introduced the linear sufficiency and linear completeness properties of a linear statistic $\mathbf{FY}$, characterized in terms of the matrices $\mathbf{X}$ and $\mathbf{V}$. Later, Markiewicz (1996) proposed a definition of general ridge estimators in terms of linear sufficiency and linear admissibility.

The authors of this paper extended, in a previous paper (Ibarrola and Pérez-Palomares (2003)), the definitions of linear sufficiency and general ridge estimators in a continuous time linear model. The aim of this paper is to study the linear completeness concept in this context. We focus on the following model. Let $T > 0$, we consider a stochastic process $(X_t, t \in [0, T])$ with values in $\mathbb{R}$ and mean function $\mu_\theta(t) = A(t)\theta$, $t \in [0, T]$, where $A(t)$ is a $1 \times p$-dimensional vector and $\theta \in \mathbb{R}^p$ is a p-dimensional unknown parameter. The process $(X_t, t \in [0, T])$ is defined in a measurable space $(\mathcal{X}, \mathcal{F})$. The distribution on $(\mathcal{X}, \mathcal{F})$ associated with the parameter value θ will be denoted by P_θ and we assume that the covariance function exists and is known. Thus,

$$E_\theta[X_t] = A(t)\theta, \quad B(t, u) := E_\theta(X_t - A(t)\theta)(X_u - A(u)\theta) = Cov(X_t, X_u).$$

In addition, we suppose that $A(t)$ and $B(t, u)$ are continuous functions in $[0, T]$ and $[0, T] \times [0, T]$, respectively. This assumption allows us to suppose that the process $(X_t, t \in [0, T])$ has measurable paths which are all in $L^2[0, T]$.

In this paper we study linear completeness for estimators of type $\int_0^T F(dt)X_t$, with a family of appropriate functions F. In Section 2, we obtain a characterization of this property, analogous to that given by Drygas (1983) for the classical linear model, and we also show the equivalence between linear completeness and ordinary completeness when the process $(X_t, t \in [0, T])$ is a Gaussian process. Finally, in Section 3 we consider the linear sufficiency and linear completeness properties which lead us to the concept of linear minimal sufficiency. Furthermore, a characterization of this property is also given.

Let us start with some definitions. Let K be a compact subset of $\mathbb{R}$, we say that $F_r(t)$ is a measurable kernel from K to $[0, T]$ if for each $r \in K$, $F_r(\cdot)$ is a bounded variation function from $[0,T]$ to $\mathbb{R}$ and for each $t \in [0,T]$, $F_{\cdot}(t)$ is a measurable function in K.

$CM(K)$ is denoted as the set of measurable kernels $F_r(t)$, $t \in [0,T]$, $r \in K$, satisfying that

$$\int_0^T \int_0^T F_r(dt)B(t,u)F_s(du), \quad (r,s) \in K \times K \quad \text{and} \quad \int_0^T F_r(dt)A(t), \quad r \in K,$$

are continuous functions.

2. Linear completeness

We are interested in defining linear completeness for estimators of type

$$\theta_r = \int_0^T F_r(dt)X_t, \quad r \in K, \quad F_r(t) \in CM(K), \tag{1}$$

where the integral in (1) is in squared mean sense.

Definition 1. *Let $\theta_r = \int_0^T F_r(dt)X_t$, $r \in K$ with $F_r(t) \in CM(K)$. The process $(\theta_r, r \in K)$ is said to be linearly complete if for each bounded variation function G in K such that $E_\theta[\int_K G(dr)\theta_r] = 0$ for all θ, it follows that $\int_K G(dr)\theta_r = 0$, a.s.*

Theorem 1. *Let $\theta_r = \int_0^T F_r(dt)X_t$, $r \in K$ and $F_r(t) \in CM(K)$. $(\theta_r, r \in K)$ is linearly complete if and only if*

$$\int_0^T F_r(dt)B(t,u) = \left(\int_0^T F_r(dt)A(t)\right) c(u), \; r \in K, \quad u \in [0,T], \tag{2}$$

where $c(u)$ is a vector of continuous functions in $[0,T]$.

Proof. The left implication is immediate since if $E_\theta[\int_K G(dr)\theta_r] = 0$ for all θ, then we obtain from (2) that $\int_K G(dr) \int_0^T F_r(dt)B(t,u) = 0$, $u \in [0,T]$. On the other hand, $Cov(\theta_r, \theta_s) = \int_0^T \int_0^T F_r(dt)B(t,u)F_s(du)$ and applying Fubini's theorem, we have

$$Cov(\int_K G(dr)\theta_r, \theta_s) = \int_0^T F_s(du) \int_K G(dr) \int_0^T F_r(dt)B(t,u) = 0, \quad s \in K.$$

It implies that the estimator $\int_K G(dr)\theta_r$ has a variance equal to 0, so it is 0, a.s.

Conversely, we consider the vector function $g(r) = \int_0^T F_r(dt)A(t)$, $r \in K$ and the symmetric and non-negative definite matrix $M = \int_K g(r)'g(r)dr$. We construct the process $\hat{\theta}_r = g(r)M^- \int_K g(s)'\theta_s ds$, $r \in K$, where M^- is a g-inverse of M. We obtain $E_\theta[\hat{\theta}_r] = g(r)M^-M\theta = g(r)\theta = E_\theta[\theta_r]$, $r \in K$. Both estimators θ_r and $\hat{\theta}_r$ are a $\int_K G(ds)\theta_s$ type, so, due to the linear completeness of $(\theta_r, r \in K)$, for each $r \in K$ we obtain

$$\theta_r = \hat{\theta}_r = g(r)M^- \int_K g(s)'\theta_s ds, \quad a.s. \tag{3}$$

Multiplying by $X_u - E_\theta[X_u]$ in (3) and integrating, we get, for each $r \in K$ and $u \in [0, T]$,

$$\int_0^T F_r(dt)B(t, u) = g(r)M^- \int_K g(s)'E_\theta[\theta_s(X_u - E_\theta[X_u])]ds.$$

Therefore, we obtain (2) with $c(u) = M^- \int_K g(s)'E_\theta[\theta_s(X_u - E_\theta[X_u])]ds$, which is a continuous function due to the continuity of $B(u, u)$, thus completing the proof.

Remark 1. We have just proved that, for a linearly complete process $(\theta_r, r \in K)$ of type (1), we have $\theta_r = g(r)M^- \int_K g(s)'\theta_s ds$, a.s. for each $r \in K$. Therefore, the covariance function of $(\theta_r, r \in K)$ is $Cov(\theta_r, \theta_s) = g(r)M^-RM^-g(s)'$ with

$$R = \int_K \int_K g(r)'Cov(\theta_r, \theta_s)g(s)drds.$$

The matrix R verifies $R = MM^-R = RM^-M = MM^-RM^-M$. The eigenvalues λ_j and the eigenfunctions $e_j(r)$ of $Cov(\theta_r, \theta_s)$ satisfy

$$\lambda_j e_j(r) = \int_K Cov(\theta_r, \theta_s)e_j(s)ds = g(r) \int_K M^-RM^-g(s)'e_j(s)ds, \tag{4}$$

so $e_j(r) = g(r)c_j$ for some vector c_j. Multiplying in formula (4) by $g(r)'$ and integrating, we get $\lambda_j Mc_j = Rc_j$. In fact $\lambda_j Mc_j = Rc_j$ is equivalent to (4) with $e_j(r) = g(r)c_j$. Therefore, there is a finite number of eigenfunctions, which means that the Karhunen-Loève expansion (see for example Todorovic (1992), pp. 140-141) for θ_r is a finite expansion given by

$$\theta_r = g(r)\theta + \sum_{j=1}^k \frac{1}{\sqrt{\lambda_j}} g(r)c_j \int_K g(s)c_j(\theta_s - E_\theta[\theta_s])ds,$$

where λ_j and c_j are the solutions of $\lambda_j Mc_j = Rc_j$.

To finish this section, we establish the equivalence between linear completeness and ordinary completeness under normality. We denote by $\sigma(\theta_r, r \in K)$ the σ-field generated by the process and we write $f \in \sigma(\theta_r, r \in K)$ when f is measurable with respect to this σ-field. We say that the process $(\theta_r, r \in K)$ is complete, if for each $f \in \sigma(\theta_r, r \in K)$ such that $\int_K f dP_\theta = 0$ for all θ, then $f = 0$, a.s.

Theorem 2. *Assume that the process $(X_t, t \in [0,T])$ is a Gaussian process. Let $(\theta_r, r \in K)$ be a process of type* (1)*. Then, $(\theta_r, r \in K)$ is complete if and only if it is linearly complete.*

Proof. We have only to prove the left implication. From Theorem 1 and Remark 1, it holds that for each $r \in K$, $\theta_r = g(r)M^- Z$ a.s., where $Z = \int_K g(s)'\theta_s ds$ is a normally distributed $p-$dimensional vector with mean $M\theta$ and covariance matrix MM^-RM^-M.

Let $f \in \sigma(\theta_r, r \in K)$, there exists a sequence $r_j \in K$, $j \geq 1$ such that $f \in \sigma(\theta_{r_j}, j \geq 1)$ (see, for example, Doob (1953)). Consider the σ-field $\sigma(\hat{\theta}_{r_j}, j \geq 1)$, where $\hat{\theta}_r = g(r)M^- Z$, $r \in K$. Since $\theta_{r_j} = \hat{\theta}_{r_j}, j \geq 1$, except in a P_θ-null set N, it is easy to see that the subset $\mathcal{H} = \{A \in \sigma(\theta_{r_j}, j \geq 1)$ such that $\exists\, A^* \in \sigma(\hat{\theta}_{r_j}, j \geq 1)$ with $A \cap N^c = A^* \cap N^c\}$ is a λ-system which contains all finite dimensional rectangles. Then, using a $\pi - \lambda$-argument, we obtain that $\mathcal{H} = \sigma(\theta_{r_j}, j \geq 1)$. By classic techniques, we can conclude that for each $f \in \sigma(\theta_r, r \in K)$ there exists a function $f^* \in \sigma(\hat{\theta}_r, r \in K)$ with $f = f^*$ a.s. On the other hand, since Z is complete (see Drygas (1983) Lemma 4.2.), $(\hat{\theta}_r, r \in K)$ is complete because a measurable function of a complete statistic is complete. Therefore, if $\int_K f dP_\theta = 0$ for all θ, then $\int_K f^* dP_\theta = 0$ for all θ. Thus $f^* = 0$, a.s. and $f = 0$, a.s. This shows the result.

3. Linear completeness and linear sufficiency

In this section we give a characterization of a linearly sufficient and linearly complete estimator, using the previous results given in Ibarrola and Pérez-Palomares (2003). The study of these properties simultaneously helps us to characterize the concept of linearly minimal sufficient estimator, which is introduced afterwards. The main result of this section, Theorem 3, is analogous to that given in Drygas (1983) (Theorem 3.7) for discrete time.

First of all, we recall some definitions and we introduce the new notation necessary for this purpose.

Definition 2. (a) *An estimator* $(\theta_r,\ r \in K)$ *of type* (1)*, unbiased for an estimable function* $(g(r)\theta,\ r \in K)$*, is called BLUE (best linear unbiased estimator) if, for each* $r \in K$*,* θ_r *is of minimum variance among all the estimators of type* (1) *unbiased for* $(g(r)\theta,\ r \in K)$*.*
(b) *An estimator* $(\theta_r,\ r \in K)$ *of type* (1) *is linearly sufficient if for each estimable function there exists a BLUE of type* $\int_K G(dr)\theta_r$*, where* G *is a function or a family of functions of bounded variation.*

Then, we can introduce the concept of linear minimal sufficiency in a natural way, as follows.

Definition 3. *An estimator* $(\theta_r,\ r \in K)$ *of type* (1) *is linearly minimal sufficient if it is linearly sufficient and for each estimator* $(\hat{\theta}_s,\ s \in S)$ *of type* (1)*, linearly sufficient, there exists a family of bounded variation functions* G_r*,* $r \in K$*, such that, for each* $r \in K$*,* $\theta_r = \int_S G_r(ds)\hat{\theta}_s$*, a.s.*

Now, we recall the function W defined in Ibarrola and Pérez-Palomares (2003) as

$$W(t,u) = B(t,u) + A(t)MA(u)', \quad t,u \in [0,T],$$

where M is any $p \times p$ symmetric and non-negative matrix such that

$$A(t) = \int_0^T W(t,u)V(du), \quad t \in [0,T],$$

for a vector $V = (V^1, \ldots, V^p)$ of bounded variation functions.

Let R, R' and K be compact subsets of the real line and C, ν real-valued functions defined in $R \times K$ and $K \times R'$, respectively. Let $C \circ \nu$ be the function in $R \times R'$ given by

$$C \circ \nu(t,s) = \int_K C(t,u)\nu(du,s), \quad t \in R, \quad s \in R',$$

whenever it makes sense. If C is defined in $R \times K$, C' corresponds to the function in $K \times R$ given by $C'(u,t) = C(t,u)$. Finally, $im(C)$ is the set of all functions of type $C \circ \nu$, whenever this composition is defined. With this notation, if we consider each element F_r in $CM(K)$ as a function defined in $K \times [0,T]$ and A as a map defined in $[0,T] \times \{1, \ldots, p\}$, where $A(t,i) = A^i(t)$, we have that $im(F \circ A)$ is the set of functions defined in K of the form $\int_0^T F_r(dt)A(t)c$, with $c \in \mathbb{R}^p$. In the same way, $im(A)$ is the set of functions defined in $[0,T]$ of type $A(t)c$, with $c \in \mathbb{R}^p$.

With these elements we can establish the following result.

Theorem 3. *Let* $(\theta_r, r \in K)$ *be an estimator of type* (1)*. Then*

(a) $(\theta_r, r \in K)$ *is linearly sufficient* $\Longleftrightarrow im(A) \subseteq im(W \circ F')$.

(b) $(\theta_r, r \in K)$ *is linearly complete* $\Longleftrightarrow im(F \circ W) \subseteq im(F \circ A)$.

(c) $(\theta_r, r \in K)$ *is linearly sufficient and linearly complete* $\Longleftrightarrow im(A) = im(W \circ F')$.

(d) $(\theta_r, r \in K)$ *is linearly sufficient and linearly complete* $\Longleftrightarrow$ *it is linearly minimal sufficient.*

Proof. (a) This assertion is equivalent to Theorem 3 in Ibarrola and Pérez-Palomares (2003).

(b) By Theorem 1 we can assure that an estimator is linearly complete if and only if $im(F \circ B) \subseteq im(F \circ A)$. Thus, to prove (b) we have to show that $im(F \circ B) \subseteq im(F \circ A)$ is equivalent to $im(F \circ W) \subseteq im(F \circ A)$. This is immediate because we have, for each $r \in K$ and $u \in [0, T]$,

$$\int_0^T F_r(dt) W(t, u) = \int_0^T F_r(dt) B(t, u) + \int_0^T F_r(dt) A(t) M A(u)'.$$

(c) First suppose that $(\theta_r, r \in K)$ is linearly sufficient and linearly complete. By (a) we have that $im(A) \subseteq im(W \circ F')$. On the other hand, since $(\theta_r, r \in K)$ is linearly sufficient, there exists a family G_r, $r \in K$, of bounded variation functions such that $\int_K G_r(ds)\theta_s$ is the BLUE for $(E_\theta[\theta_r], r \in K)$. As $(\theta_r, r \in K)$ is a linearly complete estimator, we have that $\theta_r = \int_K G_r(ds)\theta_s$, a.s., $r \in K$. Therefore, $(\theta_r, r \in K)$ is the BLUE and, using Theorem 2 in Ibarrola y Pérez-Palomares (2003), we obtain $W \circ F' \in im(A)$.

Conversely, suppose that $im(W \circ F') = im(A)$. Part (a) assures that $(\theta_r, r \in K)$ is linearly sufficient. On the other hand, since $W \circ F' \in im(A)$, we have that, for each $r \in K$,

$$\int_0^T W(t, u) F_r(du) = A(t) c_r, \quad t \in [0, T], \quad c_r \in \mathbb{R}^p. \tag{5}$$

Now, we integrate (5) with respect to $V'(dt)$ and we obtain

$$\int_0^T A(u)' F_r(du) = \Sigma\, c_r, \quad \text{where} \quad \Sigma = \int_0^T V'(dt) A(t),$$

which verifies $A(t) = A(t)\Sigma^{-}\Sigma$. Thus, we have proved that

$$A(t)c_r = A(t)\Sigma^{-}\Sigma\, c_r = A(t)\Sigma^{-}\int_0^T A(u)' F_r(du).$$

Last expression, together with (5), give us $im(F \circ W) \subseteq im(F \circ A)$, and therefore we have that $(\theta_r, r \in K)$ is linearly complete.

(d) For the right implication, suppose that $(\theta_r, r \in K)$ is a linearly sufficient and linearly complete estimator and $(\hat{\theta}_s, s \in S)$ is a linearly sufficient estimator. There exist bounded variation functions G_r, $r \in K$, such that, $\int_S G_r(ds)\hat{\theta}_s$ is the BLUE for $(E_\theta[\theta_r], r \in K)$. On the other hand, since $(\theta_r, r \in K)$ is linearly sufficient and complete, it is the BLUE for $(E_\theta[\theta_r], r \in K)$. Then, necessarily $\theta_r = \int_S G_r(ds)\hat{\theta}_s$, a.s. $r \in K$, and the linear minimal sufficiency is proved.

Conversely, suppose that $(\theta_r, r \in K)$ is linearly minimal sufficient. By definition and part (a), $im(A) \subseteq im(W \circ F')$. On the other hand, we consider $\hat{\theta} = \int_0^T V'(dt)X_t$, which is linearly sufficient. Since $(\theta_r, r \in K)$ is linearly minimal sufficient, for each $r \in K$, there exists a vector c_r such that $\theta_r = c_r\hat{\theta}$, a.s. Then,

$$\int_0^T F_r(dt)B(t,u) = E_\theta[\theta_r(X_u - E_\theta[X_u])] = c_r \int_0^T V'(dt)B(t,u).$$

Due to the properties of V and W, last formula implies that $im(W \circ F') \subseteq im(A)$. This proves that $(\theta_r, r \in K)$ is linearly sufficient and linearly complete. The proof of Theorem 3 is completed.

Theorem 4. *Assume that the process $(X_t, t \in [0,T])$ is a Gaussian process. Let $(\theta_r, r \in K)$ be a process of type* (1)*. Then, $(\theta_r, r \in K)$ is a sufficient and complete estimator if and only $im(A) = im(W \circ F')$.*

Proof. It is an immediate consequence of Theorem 4 in Ibarrola and Pérez-Palomares (2003) together with Theorems 2 and 3 (c) in the present paper.

Acknowledgments

We express our gratitude to the referee for his helpful comments and suggestions.

References

Baksalary, J.K and Kala, R. 1981, Linear transformations preserving best linear unbiased estimators in a general Gauss-Markoff model, *Ann. Statist.* **9**, 913-916.

Doob, J.L. 1953, *Stochastic Processes*, John Wiley. New York.

Drygas, H. 1983, Sufficiency and completeness in the general Gauss-Markov model, *Sankhya Ser. A* **45**, 88-98.

Ibarrola, P. and Pérez-Palomares, A. 2003, Linear sufficiency and linear admissibility in a continuous time Gauss-Markov model. *J. Multivariate Anal.* **87**, 315-327.

Markiewicz, A. 1996, Characterization of general ridge estimators, *Statist. Prob. Letters* **27**, 145-148.

Müeller, J. 1987, Sufficiency and completeness in the linear model, *J. Multivariate Anal.* **21**, 312-323.

Todorovic, P. 1992, *An introduction to Stochastic Processes and Their Applications.* Springer-Verlag, New York.