

A Martingale Kronecker Lemma and Parameter Estimation for Linear Systems

Robert J. Elliott and John B. Moore

Abstract—A continuous-time martingale Kronecker lemma is proved and used to discuss parameter estimates for linear systems.

I. INTRODUCTION

Kronecker's lemma is well known in discrete time; see for example Loève [2] or Neveu [3]. In Section II of this paper, a continuous-time martingale version is established. This is then applied to discuss parameter estimation for partially observed linear systems. The filtered estimates for the matrices involved have previously been derived in the paper by Elliott and Krishnamurthy [1]. This previous paper applies maximum likelihood arguments in contrast to the direct estimates of this work. Further, rates of convergence are discussed in Section IV. We do not discuss the recursive, adaptive estimation of the parameters of the system.

II. A MARTINGALE KRONECKER LEMMA

For simplicity we consider a scalar martingale; convergence results for the vector case are immediate.

Suppose $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, $t \geq 0$ is a stochastic basis and M is a continuous locally square integrable martingale. Further, u_t is a positive nondecreasing predictable process such that

$$u_t > c > 0 \quad \text{a.s.}$$

Write $z_t := \int_{t_0}^t u_r^{-1} dM_r$ for $0 \leq t_0 \leq t$.

Theorem 2.1: Suppose $\lim_{t \rightarrow \infty} z_t(\omega) = \xi(\omega) < \infty$ a.s. Then $\lim_{t \rightarrow \infty} (1/u_t)(M_t - M_{t_0})$ exists a.s.

If $\lim_{t \rightarrow \infty} u_t(\omega) = +\infty$, this limit is zero.

Proof: For any $s, t_0 < s < t$, because u is nondecreasing

$$\begin{aligned} M_t - M_s &= \int_s^t u_r dz_r = \int_s^t u_r d(z_r - z_s) \\ &= u_t(z_t - z_s) - \int_s^t (z_r - z_s) du_r \quad \text{a.s.} \end{aligned}$$

Consequently

$$|M_t - M_s| \leq 2u_t \sup_{r \geq s} |z_r - z_s|. \tag{1}$$

Suppose first that $\lim_{t \rightarrow \infty} u_t(\omega) = u(\omega) < \infty$. Then $|M_t - M_s| \leq 2u \sup_{r \geq s} |z_r - z_s|$.

From the hypothesis that $\lim_{t \rightarrow \infty} z_t(\omega) = \xi(\omega) < \infty$ a.s., for any $\varepsilon > 0$ there is an s'_ε such that, if $r \geq s \geq s'_\varepsilon$, $|z_r - z_s| < \varepsilon/2u$. Consequently, if $r \geq s \geq s'_\varepsilon$

$$|M_t - M_s| \leq \varepsilon.$$

Manuscript received November 12, 1996. This work was supported by the NSERC under Grant A7964.

R. J. Elliott is with the Department of Mathematical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1 and also the Department of Applied Mathematics, University of Adelaide, Adelaide, SA 5005 Australia (e-mail: relliott@gpu.srv.ualberta.ca, relliott@maths.adelaide.edu.au).

J. B. Moore is with the Department of Systems Engineering, Australian National University, Canberra A.C.T. 0200, Australia.

Publisher Item Identifier S 0018-9286(98)06603-3.

That is, $M_t(\omega)$ satisfies a Cauchy condition and converges to a limit $\mu(\omega)$. Then, $\lim_{t \rightarrow \infty} (1/u_t)(M_t - M_{t_0}) = 1/u(\mu - M_{t_0})$.

Suppose now that $\lim_{t \rightarrow \infty} u_t(\omega) = +\infty$. Give $\varepsilon > 0$. Again using the Cauchy condition for z there is an s_ε such that, if $r \geq s_\varepsilon \vee t_0$

$$|z_{s_\varepsilon} - z_r| < \frac{\varepsilon}{3}.$$

Consequently, $\sup_{r \geq s_\varepsilon \vee t_0} |z_r - z_{s_\varepsilon}| \leq \varepsilon/3$. From (1), if $t \geq s_\varepsilon \vee t_0$

$$\frac{1}{u_t} |M_t - M_{s_\varepsilon \vee t_0}| \leq \frac{2\varepsilon}{3}.$$

Suppose $t_0 \leq s_\varepsilon \vee t_0 < t$. Now $\lim_{t \rightarrow \infty} u_t = +\infty$, so there is a t_ε such that, if $t > t_\varepsilon$

$$u_t \geq 3|M_{s_\varepsilon \vee t_0} - M_{t_0}|/\varepsilon.$$

That is, $1/u_t |M_{s_\varepsilon \vee t_0} - M_{t_0}| \leq \varepsilon/3$. Now

$$\frac{1}{u_t} |M_t - M_{t_0}| \leq \frac{1}{u_t} |M_{s_\varepsilon \vee t_0} - M_{t_0}| + \frac{1}{u_t} |M_t - M_{s_\varepsilon \vee t_0}|.$$

So if $t > s_\varepsilon \vee t_\varepsilon \vee t_0$

$$\frac{1}{u_t} |M_t - M_{t_0}| \leq \varepsilon$$

and the result is proved.

III. DYNAMICS

On a stochastic basis $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ suppose there are m - and n -dimensional Wiener processes w and v , respectively.

A signal process $\{x_t\}$, $t \geq 0$ is described by the equation

$$dx_t = Ax_t dt + dw_t, \quad x_0 \in R^m. \tag{2}$$

An observation process is described by the equation

$$dy_t = Cx_t dt + dv_t, \quad y_0 = 0 \in R^n. \tag{3}$$

We suppose w and v have covariance matrices B and D , respectively. D is assumed positive definite. $\{Y_t\}$, $t \geq 0$, will denote the filtration generated by y .

In the filtering and estimation literature it is often supposed that, under some reference measure \bar{P} , y itself is a standard n -dimensional Brownian motion. These problems are discussed in, for example, Elliott and Krishnamurthy [1].

Prime ($'$) will denote transpose, so for vectors z_1, z_2 , $z_1 z_2'$ is the Kronecker product.

From (2)

$$\int_0^t dx_s x'_s = A \int_0^t x_s x'_s ds + \int_0^t dw_s x'_s. \tag{4}$$

Write

$$J_t = \int_0^t dx_s x'_s, \quad O_t = \int_0^t x_s x'_s ds, \quad M_t = \int_0^t dw_s x'_s$$

and

$$\hat{J}_t = E[J_t|Y_t], \quad \hat{O}_t = E[O_t|Y_t], \quad \hat{M}_t = E[M_t|Y_t].$$

Recursive formulas which evaluate \hat{J} and \hat{O} are given in [1]. From (4) $\hat{J}_t = A\hat{O}_t + \hat{M}_t$. An estimate for A is, therefore

$$\hat{A}_t = \hat{J}_t \hat{O}_t^{-1}$$

and the error $\hat{A}_t - A = \hat{M}_t \hat{O}_t^{-1}$.

In this paper we investigate the convergence of this error to zero.

IV. CONVERGENCE

Consider a function $\rho(t)$, $t \geq 0$, which is positive nondecreasing and such that $\lim_{t \rightarrow \infty} \int_0^t \rho(s)^{-1} ds = \lambda < \infty$. Note from Theorem 2.1, this last condition implies that $\lim_{t \rightarrow \infty} t\rho(t)^{-1} = 0$. An example of such a function is

$$\rho(t) = \max(1, t(\log t)(\log \log t)^\alpha), \quad \alpha > 1.$$

Clearly any function which grows faster than t^α , $\alpha > 1$, at infinity satisfies the condition. The strongest results are those for functions which have the slowest growth at infinity.

Consider the (matrix) martingale M_t . M is locally square integrable; $\langle M \rangle$ will denote the predictable nonnegative martingale process such that $M_t M'_t - \langle M \rangle_t$ is a local martingale.

In fact $\langle M \rangle_t = BB' \int_0^t x'_s x_s ds$ and $\text{Tr} \langle M \rangle_t = \text{Tr} BB' \int_0^t x'_s x_s ds$. Consider the martingale

$$R_t := \int_0^t \rho(\text{Tr} \langle M \rangle_s)^{-1/2} dM_s.$$

Lemma 4.1: R_t is a square integrable martingale so $\lim_{t \rightarrow \infty} R_t = \xi(w) < \infty$ exists a.s.

Proof: $E[\text{Tr} R_t R'_t] = E[\int_0^t \rho(\text{Tr} \langle M \rangle_s)^{-1} d(\text{Tr} \langle M \rangle_s)]$. Now $\int_0^t \rho(\text{Tr} \langle M \rangle_s)^{-1} d(\text{Tr} \langle M \rangle_s) < \lambda$ a.s. So $\lim_{t \rightarrow \infty} E[\text{Tr} R_t R'_t] \leq \lambda < \infty$ and R_t is a square integrable martingale for $0 \leq t \leq \infty$. \square

Corollary 4.2: From Theorem 2.1, if ρ is continuous

$$\lim_{t \rightarrow \infty} \rho(\text{Tr} \langle M \rangle_t)^{-1/2} M_t$$

exists.

If $\lim_{t \rightarrow \infty} \text{Tr} \langle M \rangle_t = +\infty$, this limit is zero. ($\text{Tr} \langle M \rangle_t$ is an increasing process so $\lim_{t \rightarrow \infty} \text{Tr} \langle M \rangle_t$ exists and is either finite or $+\infty$.)

Corollary 4.3: $\lim_{t \rightarrow \infty} \rho(\int_0^t x'_s x_s ds)^{-1/2} M_t$ exists a.s.

Proof: Note that, apart from the positive constant $B^* = \text{Tr} BB'$

$$\text{Tr} \langle M \rangle_t \text{ is } \int_0^t x'_s x_s ds.$$

Therefore, as ρ is nondecreasing

$$\begin{aligned} \rho((B^* + 1)^{-1} \text{Tr} \langle M \rangle_t) &= \rho\left((B^* + 1)^{-1} B^* \int_0^t x'_s x_s ds\right) \\ &\leq \rho\left(\int_0^t x'_s x_s ds\right). \end{aligned}$$

So

$$\rho\left(\int_0^t x'_s x_s ds\right)^{-1} \leq \rho((B^* + 1)^{-1} \text{Tr} \langle M \rangle_t)^{-1}.$$

With

$$\bar{R}_t := \int_0^t \rho\left(\int_0^s x'_u x_u du\right)^{-1/2} dM_s$$

we have

$$E[\text{Tr} \bar{R}_t \bar{R}'_t] \leq (B^* + 1)\lambda < \infty.$$

Therefore, $\lim_{t \rightarrow \infty} \bar{R}_t$ exists and is finite a.s. so from Theorem 2.1

$$\lim_{t \rightarrow \infty} \rho\left(\int_0^t x'_s x_s ds\right)^{-1/2} M_t \text{ exists a.s.}$$

Corollary 4.4: Suppose x satisfies the stability property

$$L = \sup_t \frac{1}{t} \int_0^t x'_s x_s ds < \infty$$

and

$$\lim_{t \rightarrow \infty} \rho(t) = \infty.$$

Then

$$\lim_{t \rightarrow \infty} \rho(t)^{-1/2} M_t = 0 \text{ a.s.}$$

Proof:

$$\begin{aligned} &\rho((B^* + 1)^{-1}(L + 1)^{-1} \text{Tr} \langle M \rangle_t) \\ &= \rho\left((B^* + 1)^{-1}(L + 1)^{-1} B^* \int_0^t x'_s x_s ds\right) \\ &= \rho\left((B^* + 1)^{-1}(L + 1)^{-1} B t \frac{1}{t} \int_0^t x'_s x_s ds\right) \\ &\leq \rho((B^* + 1)^{-1}(L + 1)^{-1} B L t) \leq \rho(t). \end{aligned}$$

Therefore

$$\rho(t)^{-1} \leq \rho((B^* + 1)^{-1}(L + 1)^{-1} \text{Tr} \langle M \rangle_t).$$

With

$$\bar{\bar{R}}_t := \int_0^t \rho(s)^{-1/2} dM_s$$

we have

$$\begin{aligned} E[\text{Tr} \bar{\bar{R}}_t \bar{\bar{R}}'_t] &= E\left[\int_0^t \rho(s)^{-1} d(\text{Tr} \langle M \rangle_s)\right] \\ &\leq (B^* + 1)(L + 1)\lambda < \infty. \end{aligned}$$

Therefore, $\lim_{t \rightarrow \infty} \bar{\bar{R}}_t$ exists and is finite a.s. Thus from Theorem 2.1

$$\lim_{t \rightarrow \infty} \rho(t)^{-1/2} M_t \text{ is zero a.s.}$$

Theorem 4.5: Suppose x satisfies the stability property of Corollary 4.4 and $\lim_{t \rightarrow \infty} \rho(t) = \infty$. Further, suppose x satisfies the excitation condition

$$\rho(t)^{-1} \hat{O}_t > K > 0$$

where $O_t = \int_0^t x_s x'_s ds$ and $\hat{O}_t = E[O_t|Y_t]$. Then

$$\lim_{t \rightarrow \infty} \hat{M}_t \hat{O}_t^{-1} = 0 \text{ a.s.}$$

with convergence at a rate $\rho(t)^{1/2}$.

Proof: The stability property states that $\sup_t (1/t) \int_0^t x'_s x_s ds \leq L < \infty$ a.s. Therefore

$$\sup_t \frac{1}{t} E \left[\int_0^t x'_s x_s ds \right] \leq L < \infty$$

and, because $\lim_{t \rightarrow \infty} t \rho(t)^{-1} = 0$

$$\sup_t \frac{1}{\rho(t)} E[\text{Tr}\langle M \rangle_t] < \infty$$

and the set of random variables $\{\rho(t)^{-1/2} M_t\}$ is bounded in L^2 . We can, therefore, condition the convergence of Corollary 4.4 and deduce

$$\lim_{t \rightarrow \infty} \rho(t)^{-1/2} \hat{M}_t = 0 \quad \text{a.s.}$$

Now

$$\begin{aligned} \hat{M}_t \hat{O}_t^{-1} &= \rho(t)^{-1/2} \hat{M}_t (\rho(t)^{-1/2} \hat{O}_t)^{-1} \\ &< \rho(t)^{-1/2} \hat{M}_t \rho(t)^{-1/2} K^{-1}. \end{aligned}$$

Therefore, $\lim_{t \rightarrow \infty} \rho(t)^{1/2} \hat{M}_t \hat{O}_t^{-1} = 0$ a.s. and the result follows.

V. OBSERVATION COEFFICIENT

Recall $dy_s = C x_s ds + dv_s$. Consequently

$$\int_0^t dy_s x'_s = C \int_0^t x_s x'_s ds + \int_0^t dv_s x'_s.$$

Write

$$\begin{aligned} T_t &= \int_0^t dy_s x'_s, & N_t &= \int_0^t dv_s x'_s \\ \hat{T}_t &= E[T_t | Y_t], & \hat{N}_t &= E[N_t | Y_t]. \end{aligned}$$

An estimate for C is

$$\hat{C}_t = \hat{T}_t \hat{O}_t^{-1}$$

and the error is

$$\hat{C}_t - C = \hat{N}_t \hat{O}_t^{-1}.$$

Similar discussions to those in Section IV allow us to conclude that, under the stability and excitation conditions, the error $\hat{N}_t \hat{O}_t^{-1}$ converges to zero almost surely at a rate $\rho(t)^{1/2}$.

ACKNOWLEDGMENT

J. B. Moore wishes to thank the Department of Mathematical Sciences at the University of Alberta for its hospitality in July 1996.

REFERENCES

- [1] R. J. Elliott and V. Krishnamurthy, "Exact finite dimensional filters for maximum likelihood parameter estimation of continuous time linear Gaussian systems," *SIAM J. Contr. Optimiz.*, vol. 35, pp. 1908–1923, 1997.
- [2] M. Loève, *Probability Theory I*, 4th ed. New York: Springer-Verlag, 1977.
- [3] J. Neveu, *Discrete Parameter Martingales*. Amsterdam, The Netherlands: North Holland, 1975.

Comments on "Robust, Fragile, or Optimal?"

Perti M. Mäkilä

Abstract—Recently the issue of fragile controllers (high sensitivity of closed-loop stability and/or performance to small changes in controller coefficients) produced by using popular robust and optimal control synthesis methods was raised by the above-mentioned paper.¹ This paper had at least three serious flaws. First, the authors did not provide any references to the wide earlier work in which methods are given to analyze and solve fragility and related robustness issues. Second, Keel and Bhattacharyya used mostly simple textbook examples in which the optimization criteria for controller synthesis were so simple that they do not incorporate realistic design considerations. This also resulted in the third big flaw of the paper by Keel and Bhattacharyya; namely, the possible explanation given to the cause of these problems via holes in parametric stability space for high-order controllers is misplaced. A more direct explanation is in the badly chosen optimization criteria and controller parameterizations which make the controller synthesis and realization problems considered rather unrealistic and (mathematically) ill-posed.

In the present paper we comment on the paper by Keel and Bhattacharyya and discuss fragility and other robustness issues through the well-established tools of coprime factorizations and robustness optimization. The main conclusion is that by adopting sensible optimization criteria, which take into account enough of the important design considerations directly, and numerically robust controller parameterizations, controller fragility should not be a big problem in applications of modern robust and optimal control theory.

Index Terms—Coprime factorizations, fragility, optimality, robustness.

I. INTRODUCTION

In the recent paper by Keel and Bhattacharyya¹ the interesting issue of possible fragility of controllers designed by popular robust and optimal control synthesis tools was discussed mostly via simple textbook examples. It was very disappointing that the authors had completely missed the wide earlier work devoted to methods that can deal not only with fragility but other robustness issues as well. This undermines any possible contribution that the paper by Keel and Bhattacharyya might otherwise have. The present paper provides earlier references and points out well-known analysis and synthesis methods that can address and solve fragility and other robustness problems.

It is widely recognized that typically the most important robustness issue is related to modeling uncertainty and system perturbations. Hence many articles and books concentrate on this case. However, robustness to controller uncertainty also has been addressed by many authors; see e.g., [18], [13], [16], [17], [8], [19], and the references therein. In fact, very detailed controller sensitivity studies have been undertaken in the context of embedded control systems and fixed word-length controller implementations; see, e.g., [12] and the references therein. Furthermore, Keel and Bhattacharyya have missed the fact that controller implementation is in no way restricted to the particular standard transfer function form that they study for coefficient sensitivity. In fact, it is part of the *folklore* in control engineering practice that this form should be avoided especially with

Manuscript received October 20, 1997.

The author is with the Automation and Control Institute, Tampere University of Technology, FIN-33101 Tampere, Finland.

Publisher Item Identifier S 0018-9286(98)07205-5.

¹L. H. Keel and S. P. Bhattacharyya, *IEEE Trans. Automat. Contr.*, vol. 42, pp. 1098–1105, 1997.