



**HAL**  
open science

# Convergence of the gradient algorithm for linear regression models in the continuous and discrete time cases

Laurent Praly

► **To cite this version:**

Laurent Praly. Convergence of the gradient algorithm for linear regression models in the continuous and discrete time cases. [Research Report] PSL Research University; Mines ParisTech. 2017. hal-01423048v4

**HAL Id: hal-01423048**

**<https://hal.science/hal-01423048v4>**

Submitted on 18 May 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Convergence of the gradient algorithm for linear regression models in the continuous and discrete time cases

Laurent Praly\*

Started December 26, 2016,      Revised may 18, 2018

version : 4

## Abstract

We establish convergence to zero of the solutions of

$$\dot{\tilde{\theta}}(t) = -\phi(t)\phi(t)^\top \tilde{\theta}(t) \quad \text{or} \quad \tilde{\theta}(t) = \tilde{\theta}(t-1) - \phi(t)\phi(t)^\top \tilde{\theta}(t-1)$$

under a possibly “vanishing persistent” excitation condition.

## 1 Continuous time case

Given a continuous vector function<sup>1</sup>  $t \in \mathbb{R} \mapsto \phi(t) \in \mathbb{R}^n$ , we let  $\mathfrak{T}(t, s)$  be the transition matrix associated with the non autonomous differential equation<sup>2</sup>

$$\dot{\tilde{\theta}}(t) = -\phi(t)\phi(t)^\top \tilde{\theta}(t) \tag{1}$$

i.e. satisfying

$$\partial_1 \mathfrak{T}(t, s) = -\phi(t)\phi(t)^\top \mathfrak{T}(t, s) \quad , \quad \mathfrak{T}(s, s) = I \tag{2}$$

where  $\partial_1$  denotes the partial derivative with respect to the first argument,  $t$  here. We are interested in sufficient conditions implying

$$\lim_{t \rightarrow +\infty} |\mathfrak{T}(t, s)| = \lim_{t \rightarrow +\infty} \sup_x \frac{|\mathfrak{T}(t, s)x|}{|x|} = 0 .$$

---

\*MINES ParisTech, PSL Research University, CAS - Centre automatique et systèmes, 35 rue St Honoré 77300 Fontainebleau, France

<sup>1</sup>About the smoothness of  $\phi$  we need only the differential equation to have solutions; so, since we have boundedness, measurability is sufficient. We need also to be able to change the order of integration.

<sup>2</sup> The function  $\phi$  is often obtained from a pre-processing from a “raw” function say  $\Phi$ . It may be

$$\phi(t) = \gamma(t) \frac{\Phi(t)}{\sqrt{1 + |\Phi(t)|^2}} \quad \text{or} \quad \phi(t) = \gamma(t) \frac{\Phi(t)}{\sqrt{r(t)}} \quad , \quad \dot{r}(t) = -\lambda r(t) + |\Phi(t)|^2 .$$

Let  $x$  be an arbitrary (constant) vector. We have

$$\partial_1 |\mathfrak{T}(t, s)x|^2 = -2|\phi(t)^\top \mathfrak{T}(t, s)x|^2$$

By defining the function  $\hat{f}_s$  as

$$\hat{f}_s(t) = \phi(t)^\top \mathfrak{T}(t, s)x \quad , \quad \hat{f}_s(s) = \phi(s)^\top x \quad (3)$$

and integrating, we get

$$|\mathfrak{T}(t, s)x|^2 = |x|^2 - 2 \int_s^t |\hat{f}_s(r)|^2 dr \quad (4)$$

On another hand, by integrating (2), we get

$$\mathfrak{T}(t, s) = I - \int_s^t \phi(r)\phi(r)^\top \mathfrak{T}(r, s) dr \quad (5)$$

Incorporating this in the definition of  $\hat{f}_s$  yields

$$\begin{aligned} \hat{f}_s(t) &= \phi(t)^\top \mathfrak{T}(t, s)x \\ &= \phi(t)^\top \left( I - \int_s^t \phi(r)\phi(r)^\top \mathfrak{T}(r, s) dr \right) x \\ &= \phi(t)^\top x - \int_s^t (\phi(t)^\top \phi(r)) (\phi(r)^\top \mathfrak{T}(r, s)x) dr \\ &= f(t) - \int_s^t (\phi(t)^\top \phi(r)) \hat{f}_s(r) dr \end{aligned} \quad (6)$$

where we have let

$$f(t) = \phi(t)^\top x . \quad (7)$$

It follows that we have<sup>3</sup>

$$\begin{aligned} |f(u) - \hat{f}_s(u)|^2 &= \left( \phi(u)^\top \int_s^u \phi(r) \hat{f}_s(r) dr \right)^2 \\ &\leq |\phi(u)|^2 \left| \int_s^u \phi(r) \hat{f}_s(r) dr \right|^2 \\ &\leq |\phi(u)|^2 \left( \int_s^u |\phi(r)|^2 dr \right) \left( \int_s^u \hat{f}_s(r)^2 dr \right) \\ \int_s^t |f(u) - \hat{f}_s(u)|^2 du &\leq \int_s^t \left( |\phi(u)|^2 \left( \int_s^u |\phi(r)|^2 dr \right) \left( \int_s^u \hat{f}_s(r)^2 dr \right) \right) du \\ &\leq \left( \int_s^t |\phi(u)|^2 du \right)^2 \left( \int_s^t \hat{f}_s(r)^2 dr \right) \\ \int_s^t f(u)^2 du &\leq 2 \int_s^t \hat{f}_s(u)^2 du + 2 \int_s^t |f(u) - \hat{f}_s(u)|^2 du \end{aligned}$$

---

<sup>3</sup> These inequalities would be in part equalities if there exist functions  $\mu : \mathbb{R} \rightarrow \mathbb{R}$  and  $\nu : \mathbb{R} \rightarrow \mathbb{R}^n$  satisfying,

$$\phi(u) = \mu(u) \int_s^u \phi(r) \hat{f}_s(r) dr = \nu(u) \hat{f}_s(u) \quad \forall u \in [s, t] .$$

$$\begin{aligned}
&\leq 2 \int_s^t \hat{f}_s(u)^2 du + 2 \left( \int_s^t |\phi(u)|^2 du \right)^2 \left( \int_s^t \hat{f}_s(r)^2 dr \right) \\
&\leq 2 \left[ 1 + \left( \int_s^t |\phi(u)|^2 du \right)^2 \right] \left( \int_s^t \hat{f}_s(r)^2 dr \right)
\end{aligned} \tag{8}$$

So we get finally

$$\begin{aligned}
|\mathfrak{I}(t, s)x|^2 &= |x|^2 - 2 \int_s^t \left| \hat{f}_s(r) \right|^2 dr \\
&\leq |x|^2 - \frac{1}{1 + \left( \int_s^t |\phi(u)|^2 du \right)^2} \int_s^t f(u)^2 du \\
&\leq |x|^2 - \frac{1}{1 + \left( \int_s^t |\phi(u)|^2 du \right)^2} \int_s^t [\phi(u)^\top x]^2 du \\
&\leq x^\top \left( I - \frac{1}{1 + \left( \int_s^t |\phi(u)|^2 du \right)^2} \int_s^t \phi(u)\phi(u)^\top du \right) x
\end{aligned}$$

In other words we have simply

$$\mathfrak{I}(t, s)^\top \mathfrak{I}(t, s) \leq I - \frac{\int_s^t \phi(u)\phi(u)^\top du}{1 + \left( \int_s^t |\phi(u)|^2 du \right)^2} \tag{9}$$

### Remark 1

1. Up to (7), we have identities. The conservativeness we may have in this last inequality is only in the majoration obtained in (8). See footnote 3.
2. The steps used up to (??) for the differential equation (1) can also be used for :

$$\dot{\tilde{\theta}}(t) = -\psi(t)\phi(t)^\top \tilde{\theta}(t)$$

as we have for example in the least square algorithm.

3. With adapting the arguments used in the proof of the claim p.369 of Appendix B.2 of [3], it should be possible to extend the above result to the system

$$\dot{\eta}(t) = A\eta(t) + B\phi(t)^\top \tilde{\theta}(t) \quad , \quad \dot{\tilde{\theta}}(t) = -\phi(t)C\eta(t)$$

where the triple  $(B, A, C)$  is strictly positive real.

### Other relations

We have also

$$\partial_2 \mathfrak{I}(t, s) = \mathfrak{I}(t, s)\phi(s)\phi(s)^\top$$

By direct integration, we obtain the following other expression for  $\mathfrak{I}(t, s)$

$$\begin{aligned}\mathfrak{I}(t, s) &= I + \int_t^s \mathfrak{I}(t, u) \phi(u) \phi(u)^\top du \\ &= I - \int_s^t \mathfrak{I}(t, u) \phi(u) \phi(u)^\top du\end{aligned}\quad (10)$$

We “merge” the two expression (5) and (10) by substituting one inside the integral of the other. This gives

$$\begin{aligned}\mathfrak{I}(t, s) &= I - \int_s^t \phi(r) \phi(r)^\top \left( I - \int_s^r \mathfrak{I}(r, u) \phi(u) \phi(u)^\top du \right) dr \\ &= I - \int_s^t \phi(r) \phi(r)^\top dr + \int_s^t \phi(r) \phi(r)^\top \left( \int_s^r \mathfrak{I}(r, u) \phi(u) \phi(u)^\top du \right) dr \\ &= I - \int_s^t \phi(r) \phi(r)^\top dr + \int_s^t \phi(r) \left( \int_s^r k(r, u) \phi(u)^\top du \right) dr\end{aligned}\quad (11)$$

where the kernel  $k$  is defined as

$$k(r, u) = \phi(r)^\top \mathfrak{I}(r, u) \phi(u)$$

By inserting the expression (5) of  $\mathfrak{I}$  in this definition of  $k$  we get

$$\begin{aligned}k(t, s) &= \phi(t)^\top \mathfrak{I}(t, s) \phi(s) \\ &= \phi(t)^\top \left( I - \int_s^t \phi(r) \phi(r)^\top \mathfrak{I}(r, s) dr \right) \phi(s) \\ &= \phi(t)^\top \phi(s) - \int_s^t \phi(t)^\top \phi(r) \phi(r)^\top \mathfrak{I}(r, s) \phi(s) dr \\ &= \phi(t)^\top \phi(s) - \int_s^t \phi(t)^\top \phi(r) k(r, s) dr\end{aligned}\quad (12)$$

With using (10), we obtain :

$$k(t, s) = \phi(t)^\top \phi(s) - \int_s^t k(t, r) \phi(r)^\top \phi(s) dr$$

A way to view the identity (12) is that, letting  $\varphi(t, s) = \phi(t)^\top \phi(s)$ , the kernel  $\delta + \varphi$  is the inverse of the kernel  $\delta - k$  where  $\delta$  is the Dirac distribution. Precisely, for any  $C^1$  test function  $f$  we obtain formally (a rigorous computation is given below)

$$\begin{aligned}&\int_s^t (\delta(t - u) + \varphi(t, u)) \left( \int_s^u (\delta(u - r) - k(u, r)) f(r) dr \right) du \\ &= \int_s^t (\delta(t - u) + \varphi(t, u)) \left( f(u) - \int_s^u k(u, r) f(r) dr \right) du \\ &= \left( f(t) - \int_s^t k(r, s) f(r) dr \right) + \int_s^t \varphi(t, u) \left( f(u) - \int_s^u k(u, r) f(r) dr \right) du \\ &= f(t) - \int_s^t k(r, s) f(r) dr + \int_s^t \varphi(t, u) f(u) du - \int_s^t \left( \int_r^t \varphi(t, u) k(u, r) du \right) f(r) dr \\ &= f(t) - \int_s^t k(r, s) f(r) dr + \int_s^t \varphi(t, u) f(u) du - \int_s^t (\varphi(t, r) - k(t, r)) f(r) dr \\ &= f(t)\end{aligned}\quad (13)$$

We exploit this remark to reestablish the relation (6) between  $\hat{f}_s$  defined in (3) and  $f$  defined in (7). With (11) and (12), we obtain :

$$\begin{aligned}
\hat{f}_s(t) &= \phi(t)^\top x - \int_s^t \phi(t)^\top \phi(r) \phi(r)^\top x dr + \int_s^t \phi(t)^\top \phi(r) \left( \int_s^r k(r,u) \phi(u)^\top x du \right) dr \\
&= [\phi(t)^\top x] - \int_s^t \phi(t)^\top \phi(u) [\phi(u)^\top x] du + \int_s^t \left( \int_u^t \phi(t)^\top \phi(r) k(r,u) dr \right) [\phi(u)^\top x] du \\
&= [\phi(t)^\top x] - \int_s^t \left( \phi(t)^\top \phi(u) - \int_u^t \phi(t)^\top \phi(r) k(r,u) dr \right) [\phi(u)^\top x] du \\
&= [\phi(t)^\top x] - \int_s^t k(t,u) [\phi(u)^\top x] du
\end{aligned}$$

With the definition of  $f$  this yields :

$$\hat{f}_s(r) = f(r) - \int_s^r k(r,u) f(u) du$$

or equivalently

$$\hat{f}_s(u) = \int_s^u (\delta(u-r) - k(u,r)) f(r) dr$$

So, in view of (13), we suspect that we have

$$f(t) = \int_s^t (\delta(t-u) + \varphi(t,u)) \hat{f}_s(u) du .$$

This is indeed the case since, with (12), we obtain

$$\begin{aligned}
\hat{f}_s(t) + \int_s^t \phi(t)^\top \phi(r) \hat{f}_s(r) dr &= f(t) - \int_s^t k(t,r) f(r) dr + \int_s^t \phi(t)^\top \phi(r) \left( f(r) - \int_s^r k(r,u) f(u) du \right) dr \\
&= f(t) + \int_s^t (\phi(t)^\top \phi(r) - k(t,r)) f(r) dr - \int_s^t \phi(t)^\top \phi(r) \left( \int_s^r k(r,u) f(u) du \right) dr \\
&= f(t) + \int_s^t \left( \int_r^t \phi(t)^\top \phi(u) k(u,r) du \right) f(r) dr - \int_s^t \left( \int_u^t \phi(t)^\top \phi(r) k(r,u) dr \right) f(u) du \\
&= f(t)
\end{aligned}$$

This proves that the transformation  $f \mapsto \hat{f}_s$  is invertible and reestablishes (6).

Similarly, to any arbitrary  $C^1$  test function  $f$ , its transform  $\hat{f}_t$  defined as

$$\hat{f}_t(s) = f(s) - \int_s^t f(r) k(r,s) dr$$

It satisfies

$$\hat{f}_t(s) + \int_s^t \phi(u)^\top \phi(s) \hat{f}_t(u) du = f(s)$$

And, when  $f$  is given by (10), we have

$$\begin{aligned}\hat{f}_t(s) &= x^\top \phi(s) - \int_s^t x^\top \phi(r) \phi(r)^\top \mathfrak{T}(r, s) \phi(s) dr , \\ &= x^\top \left( I - \int_s^t \phi(r) \phi(r)^\top \mathfrak{T}(r, s) dr \right) \phi(s) , \\ &= x^\top \mathfrak{T}(t, s) \phi(s) ,\end{aligned}$$

where we have used (5) to obtain the last identity.

From this, we can for instance obtain an upperbound for  $\mathfrak{T}(t, s) \mathfrak{T}(t, s)^\top$ .

## 2 Discrete time case

Given a sequence of vectors  $\phi(t) \in \mathbb{R}^n$  bounded<sup>4</sup> in norm by  $\bar{\phi} \leq \sqrt{2}$ , we let  $\mathfrak{T}(t, s)$  be the transition matrix associated with the non autonomous discrete time system

$$\tilde{\theta}(t) = (I - \phi(t) \phi(t)^\top) \tilde{\theta}(t-1)$$

i.e. satisfying

$$\mathfrak{T}(t, s) = (I - \phi(t) \phi(t)^\top) \mathfrak{T}(t-1, s) \quad , \quad \mathfrak{T}(s, s) = I$$

Our problem is to find sufficient conditions implying

$$\lim_{t \rightarrow +\infty} |\mathfrak{T}(t, s)| = \lim_{t \rightarrow +\infty} \sup_x \frac{|\mathfrak{T}(t, s)x|}{|x|} = 0 .$$

Let  $x$  be an arbitrary unit vector. We have

$$\mathfrak{T}(t, s)x = \mathfrak{T}(t-1, s)x - \phi(t) \phi(t)^\top \mathfrak{T}(t-1, s)x$$

With denoting, for  $t \geq s+1$ ,

$$\hat{f}_s(t) = \phi(t)^\top \mathfrak{T}(t-1, s)x \quad , \quad \hat{f}_s(s+1) = \phi(s+1)^\top x$$

we get

$$\begin{aligned}|\mathfrak{T}(t, s)x|^2 &= \left( \mathfrak{T}(t-1, s)x - \phi(t) \hat{f}_s(t) \right)^\top \left( \mathfrak{T}(t-1, s)x - \phi(t) \hat{f}_s(t) \right) \\ &= |\mathfrak{T}(t-1, s)x|^2 - 2x^\top \mathfrak{T}(t-1, s)^\top \phi(t) \hat{f}_s(t) \phi(t)^\top \phi(t) \hat{f}_s(t)^2 \\ &= |\mathfrak{T}(t-1, s)x|^2 - (2 - \phi(t)^\top \phi(t)) \hat{f}_s(t)^2 \\ &\quad \vdots \\ |\mathfrak{T}(s+1, s)x|^2 &= |x|^2 - (2 - \phi(s+1)^\top \phi(s+1)) \hat{f}_s(s+1)^2\end{aligned}$$

So summation gives

$$|\mathfrak{T}(t, s)x|^2 = |x|^2 - \sum_{r=s+1}^t (2 - \phi(r)^\top \phi(r)) \hat{f}_s(r)^2 \tag{14}$$

---

<sup>4</sup>See footnote 2

On another hand, we have

$$\begin{aligned}
\mathfrak{T}(t, s) &= \mathfrak{T}(t-1, s) - \phi(t)\phi(t)^\top \mathfrak{T}(t-1, s) \\
\mathfrak{T}(t-1, s) &= \mathfrak{T}(t-2, s) - \phi(t-1)\phi(t-1)^\top \mathfrak{T}(t-2, s) \\
&\vdots \\
\mathfrak{T}(s+1, s) &= I - \phi(s+1)\phi(s+1)^\top \mathfrak{T}(s, s)
\end{aligned}$$

So again summation gives

$$\mathfrak{T}(t, s) = I - \sum_{r=s+1}^t \phi(r)\phi(r)^\top \mathfrak{T}(r-1, s) \quad (15)$$

Incorporating this in the expression of  $\hat{f}_s$  yields, for  $t \geq s+2$ ,

$$\begin{aligned}
\hat{f}_s(t) &= \phi(t)^\top \mathfrak{T}(t-1, s)x \\
&= \phi(t)^\top \left( I - \sum_{r=s+1}^{t-1} \phi(r)\phi(r)^\top \mathfrak{T}(r-1, s) \right) x \\
&= \phi(t)^\top x - \sum_{r=s+1}^{t-1} (\phi(t)^\top \phi(r)) (\phi(r)^\top \mathfrak{T}(r-1, s)x) \\
&= f(t) - \sum_{r=s+1}^{t-1} \phi(t)^\top \phi(r) \hat{f}_s(r) \quad \forall t \geq s+2,
\end{aligned}$$

where we have let

$$f(t) = \phi(t)^\top x \quad (16)$$

We have also

$$\hat{f}_s(s+1) = f(s+1)$$

With the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}
f(u)^2 &\leq \left( 1 + \sum_{r=s+1}^{u-1} [\phi(u)^\top \phi(r)]^2 \right) \left( \sum_{r=s+1}^u \hat{f}_s(r)^2 \right) \\
&\leq \left( 1 + |\phi(u)|^2 \sum_{v=s+1}^{u-1} |\phi(v)|^2 \right) \left( \sum_{r=s+1}^u \hat{f}_s(r)^2 \right) \quad \forall u \geq s+2 \\
f(s+1)^2 &= \hat{f}_s(s+1)^2
\end{aligned}$$

and therefore



$$\begin{aligned}
& \sum_{u=s+1}^t f(u)^2 \\
& \leq \hat{f}_s(s+1)^2 + \sum_{u=s+2}^t \left( 1 + |\phi(u)|^2 \sum_{v=s+1}^{u-1} |\phi(v)|^2 \right) \left( \sum_{r=s+1}^u \hat{f}_s(r)^2 \right) \\
& \leq \left( 1 + \sum_{u=s+2}^t \left( 1 + |\phi(u)|^2 \sum_{v=s+1}^{u-1} |\phi(v)|^2 \right) \right) \hat{f}_s(s+1)^2 \\
& \quad + \sum_{r=s+2}^t \left( \sum_{u=r}^t \left( 1 + |\phi(u)|^2 \sum_{v=s+1}^{u-1} |\phi(v)|^2 \right) \right) \hat{f}_s(r)^2 \\
& \leq \left( (t-s) + \left( \sum_{u=s+2}^t |\phi(u)|^2 \right) \left( \sum_{v=s+1}^{t-1} |\phi(v)|^2 \right) \right) \hat{f}_s(s+1)^2 \\
& \quad + \sum_{r=s+2}^t \left( (t-r+1) + \left( \sum_{u=r}^t |\phi(u)|^2 \right) \left( \sum_{v=s+1}^{t-1} |\phi(v)|^2 \right) \right) \hat{f}_s(r)^2 \\
& \leq \left( (t-s) + \left( \sum_{u=s+2}^t |\phi(u)|^2 \right) \left( \sum_{v=s+1}^{t-1} |\phi(v)|^2 \right) \right) \left( \sum_{r=s+1}^t \hat{f}_s(r)^2 \right) \\
& \leq \left( (t-s) + \left( \sum_{u=s+1}^t |\phi(u)|^2 \right)^2 \right) \left( \sum_{r=s+1}^t \hat{f}_s(r)^2 \right) \tag{17}
\end{aligned}$$

With the definition (16) of  $f(u)$  this inequality is

$$x^\top \left( \sum_{u=s+1}^t \phi(u)\phi(u)^\top \right) x \leq \left( (t-s) + \left( \sum_{u=s+1}^t |\phi(u)|^2 \right)^2 \right) \left( \sum_{r=s+1}^t \hat{f}_s(r)^2 \right)$$

With (14), it allows us to obtain the following upperbound for  $|\mathfrak{F}(t, s)x|^2$

$$\begin{aligned}
|\mathfrak{F}(t, s)x|^2 & \leq |x|^2 - \min_{r \in \{s+1, \dots, t\}} \{2 - \phi(r)^\top \phi(r)\} \left( \sum_{r=s+1}^t \hat{f}_s(r)^2 \right) \\
& \leq |x|^2 - \frac{\min_{r \in \{s+1, \dots, t\}} \{2 - \phi(r)^\top \phi(r)\}}{\left( (t-s) + \left( \sum_{u=s+1}^t |\phi(u)|^2 \right)^2 \right)} x^\top \left( \sum_{u=s+1}^t \phi(u)\phi(u)^\top \right) x
\end{aligned}$$

In other words we have

$$\mathfrak{F}(t, s)^\top \mathfrak{F}(t, s) \leq I - \frac{\min_{r \in \{s+1, \dots, t\}} \{2 - |\phi(r)|^2\}}{\left( (t-s) + \left( \sum_{u=s+1}^t |\phi(u)|^2 \right)^2 \right)} \left( \sum_{u=s+1}^t \phi(u)\phi(u)^\top \right). \tag{18}$$

**Remark 2** The same final remarks as for the continuous time case can be done here. In particular about the majoration (17), a less conservative bound is obtained in the proof of [2, Theorem 4.5] (or of [5, Theorem 2.2, 1st column, p. 2052], for the case where the preprocessing mentioned at the beginning of this section is

$$\phi(t) = \frac{\Phi(t)}{\sqrt{r(t)}} \quad , \quad r(t) = r(t-1) + |\Phi(t)|^2$$

and the above analysis is carried out exploiting the assumption that the sequence  $r$  goes to infinity in some specific way.

### Other relations

We have also

$$\mathfrak{T}(t, s) = \mathfrak{T}(t, s+1) (I - \phi(s+1)\phi(s+1)^\top)$$

and therefore

$$\begin{aligned} \mathfrak{T}(t, s) &= \mathfrak{T}(t, s+1) - \mathfrak{T}(t, s+1)\phi(s+1)\phi(s+1)^\top \\ \mathfrak{T}(t, s+1) &= \mathfrak{T}(t, s+2) - \mathfrak{T}(t, s+2)\phi(s+2)\phi(s+2)^\top \\ &\vdots \\ \mathfrak{T}(t, t-1) &= I - \mathfrak{T}(t, t)\phi(t)\phi(t)^\top \end{aligned}$$

So again by summation, we get an identity similar to (15)

$$\mathfrak{T}(t, s) = I - \sum_{u=s+1}^t \mathfrak{T}(t, u)\phi(u)\phi(u)^\top \quad (19)$$

Let us mix the two expressions (15) and (19) of  $\mathfrak{T}$  we have found :

$$\begin{aligned} \mathfrak{T}(t, s) &= I - \sum_{r=s+1}^t \phi(r)\phi(r)^\top \mathfrak{T}(r-1, s) \\ &= I - \phi(s+1)\phi(s+1)^\top - \sum_{r=s+2}^t \phi(r)\phi(r)^\top \left( I - \sum_{u=s+1}^{r-1} \mathfrak{T}(r-1, u)\phi(u)\phi(u)^\top \right) \\ &= I - \sum_{r=s+1}^t \phi(r)\phi(r)^\top + \sum_{r=s+2}^t \phi(r) \left( \sum_{u=s+1}^{r-1} k(r, u)\phi(u)^\top \right) \\ &= I - \sum_{u=s+1}^t \phi(u)\phi(u)^\top + \sum_{u=s+1}^{t-1} \left( \sum_{r=u+1}^t \phi(r)k(r, u) \right) \phi(u)^\top \\ &= I + \sum_{u=s+1}^t \left( \sum_{r=u}^t \phi(r)k(r, u) \right) \phi(u)^\top \end{aligned}$$

with denoting

$$k(r, u) = \phi(r)^\top \mathfrak{T}(r-1, u)\phi(u) \quad \forall r \geq u+1 .$$

Note that we have

$$k(r, r-1) = \phi(r)^\top \phi(r-1). \quad (20)$$

Also, with the expression (15) of  $\mathfrak{T}$ , we see that  $k$  satisfies, for  $t \geq s+1$ ,

$$\begin{aligned} k(t, s) &= \phi(t)^\top \mathfrak{T}(t-1, s) \phi(s) \\ &= \phi(t)^\top \left( I - \sum_{r=s+1}^{t-1} \phi(r) \phi(r)^\top \mathfrak{T}(r-1, s) \right) \phi(s) \\ &= \phi(t)^\top \phi(s) - \sum_{r=s+1}^{t-1} \phi(t)^\top \phi(r) \phi(r)^\top \mathfrak{T}(r-1, s) \phi(s) \\ &= \phi(t)^\top \phi(s) - \sum_{r=s+1}^{t-1} \phi(t)^\top \phi(r) k(r, s) \end{aligned} \quad (21)$$

Note that we have also

$$k(t, s) = \phi(t)^\top \phi(s) - \sum_{r=s+1}^{t-1} k(t, r) \phi(r)^\top \phi(s) \quad (22)$$

The identities (21) and (22) are is very meaningful. A closer look shows that by denoting :

- $K(t, s)$  be the lower triangular matrix with 1 on the diagonal and  $k(u, v)$  as element of row  $u$  and column  $v$ ,
- $L(t, s)$  be the lower triangular matrix with 1 on the diagonal and  $\phi(u)^\top \phi(v)$  as element of row  $u$  and column  $v$ ,

we have

$$L(t, s)K(t, s) = K(t, s)L(t, s) = I$$

Precisely, let  $f(t)$  be arbitrary. We obtain, with (20) and (21),

$$\begin{aligned} &\left( f(t) - \sum_{u=s+1}^{t-1} k(t, u) f(u) \right) + \sum_{r=s+1}^{t-1} \phi(t)^\top \phi(r) \left( f(r) - \sum_{u=s+1}^{r-1} k(r, u) f(u) \right) \\ &= f(t) - \sum_{u=s+2}^{t-1} k(t, u) f(u) - k(t, s+1) f(s+1) \\ &\quad + \sum_{r=s+2}^{t-1} \phi(t)^\top \phi(r) \left( f(r) - \sum_{u=s+1}^{r-1} k(r, u) f(u) \right) + \phi(t)^\top \phi(s+1) f(s+1) \\ &= f(t) + \sum_{u=s+1}^{t-1} (\phi(t)^\top \phi(u) - k(t, u)) f(u) - \sum_{u=s+1}^{t-2} \left( \sum_{r=u+1}^{t-1} \phi(t)^\top \phi(r) k(r, u) \right) f(u) \\ &= f(t) + \sum_{u=s+1}^{t-1} (\phi(t)^\top \phi(u) - k(t, u)) f(u) - \sum_{u=s+1}^{t-2} (k(t, u) - \phi(t)^\top \phi(u)) f(u) \\ &= f(t) + (\phi(t)^\top \phi(t-1) - k(t, t-1)) f(t-1) \\ &= f(t) \end{aligned} \quad (23)$$

A similar result can be obtained from (22).

### 3 Convergence for the discrete and continuous time case

Let  $t_i$  be strictly positive real numbers going to  $+\infty$  with  $t_0 = 0$ . For any  $t$ , there exists  $\tau(t)$  such that  $t$  is between  $t_{\tau(t)}$  and  $t_{1+\tau(t)}$ . We have seen that both in the discrete and continuous time case, there exist a real number  $\pi_i$  in  $[0, 1]$  satisfying

$$|\mathfrak{T}(t_i, t_{i-1})| \leq 1 - \pi_i$$

Specifically, since  $\sqrt{1-a} \leq 1 - \frac{a}{2}$

– **in the continuous time case**, (9) gives

$$\pi_i \geq \frac{1}{2} \frac{\lambda_{\min} \left( \int_{t_{i-1}}^{t_i} \phi(u) \phi(u)^\top du \right)}{1 + \left( \int_{t_{i-1}}^{t_i} |\phi(u)|^2 du \right)^2} \quad (24)$$

With denoting  $\bar{\phi}_i = \text{esssup}_{t \in [t_{i-1}, t_i]} |\phi(t)|$ , a more conservative lowerbound for  $\pi_i$  is

$$\pi_i \geq \frac{1}{2} \frac{\lambda_{\min} \left( \int_{t_{i-1}}^{t_i} \phi(u) \phi(u)^\top du \right)}{1 + (t_i - t_{i-1})^2 \bar{\phi}_i^4} \quad (25)$$

Also, in [1], the following assumption is introduced: *There exist  $\alpha$  and  $\beta$  such that there exists a function  $T$  such that we have, for all  $t \geq 0$*

$$\alpha I \leq \int_t^{T(t)} \phi(r) \phi(r)^\top dr \leq \beta I .$$

Since we have

$$\int_t^{T(t)} |\phi(r)|^2 dr = \text{trace} \left( \int_t^{T(t)} \phi(r) \phi(r)^\top dr \right) \leq n \lambda_{\max} \left( \int_t^{T(t)} \phi(r) \phi(r)^\top du \right)$$

The above assumption implies that, by letting

$$t_i = T^i(0) ,$$

we have, for all  $i$

$$\alpha \leq \lambda_{\min} \left( \int_{t_{i-1}}^{t_i} \phi(r) \phi(r)^\top dr \right) , \quad \int_{t_{i-1}}^{t_i} |\phi(r)|^2 dr \leq n \beta$$

and therefore

$$\pi_i = \frac{\alpha}{1 + n^2 \beta^2}$$

– **in the discrete time case**, with  $\bar{\phi}^2$  smaller than 2, (18) gives

$$\pi_i \geq \frac{\min_{r \in \{1+t_{i-1}, \dots, t_i\}} \{2 - |\phi(r)|^2\}}{2 \left( (t_i - t_{i-1} - 1) + \left( \sum_{u=1+t_{i-1}}^{t_i} |\phi(u)|^2 \right) \right)} \lambda_{\min} \left( \sum_{1+t_{i-1}}^{t_i} \phi(u) \phi(u)^\top \right) \quad (26)$$

or the more conservative lower bound

$$\pi_i = \frac{2 - \bar{\phi}^2}{2(t_i - t_{i-1} - 1)(1 + (t_i - t_{i-1} - 1)\bar{\phi}^4)} \lambda_{\min} \left( \sum_{1+t_{i-1}}^{t_i} \phi(u) \phi(u)^\top \right)$$

So we have

$$\begin{aligned} |\mathfrak{I}(t, 0)| &= \left| \mathfrak{I}(t, t_{\tau(t)}) \prod_{i=1}^{\tau(t)} \mathfrak{I}(t_i, t_{i-1}) \right| \\ &\leq |\mathfrak{I}(t, t_{\tau(t)})| \prod_{i=1}^{\tau(t)} |\mathfrak{I}(t_i, t_{i-1})| \\ &\leq \prod_{i=1}^{\tau(t)} (1 - \pi_i) \\ &\leq \exp \left( - \sum_{i=1}^{\tau(t)} \pi_i \right) \end{aligned}$$

where to obtain the last inequality we have used the property

$$(1 - x) \leq \exp(-x) \quad \forall x \in \mathbb{R} .$$

We conclude that  $|\mathfrak{I}(t, 0)|$  tends to 0 if we can find  $T$  and the  $t_i$ 's such that we get

$$\sum_{i=1}^{\infty} \pi_i = +\infty .$$

**Discussion :** To show the interest of this result, we compare it with the persistent excitation (spanning) condition. We do this here for the continuous time case only, but the same holds for the discrete time case.

The vector function  $\phi$  is said persistently exciting or spanning if there exist two strictly positive real numbers  $\varepsilon$  and  $T$  such that, for any  $t$ , the Gram matrix  $\int_t^{t+T} \phi(s) \phi(s)^\top ds$  on a time window with width  $T$  is above the level  $\varepsilon$ , i.e.

$$\lambda_{\min} \left( \int_t^{t+T} \phi(s) \phi(s)^\top ds \right) \geq \varepsilon \quad \forall t \geq 0 .$$

It is established in [4] that this condition is necessary and sufficient to have the uniform asymptotic stability of the origin for (1)

The condition of non summability of  $\pi_i$  above implies attractiveness but not uniform attractiveness. It can be seen weaker than the persistent excitation or spanning condition in two ways: the level  $\varepsilon$  may decrease with  $t$ , the width  $T$  of the time window may increase with  $t$ .

Specifically, let  $T$  be fixed and let  $\varepsilon_i$  be the level reached by the Gram matrix on the  $i$ th time window  $[(i-1)T, iT]$ , i.e.

$$\varepsilon_i = \lambda_{\min} \left( \int_{(i-1)T}^{iT} \phi(s)\phi(s)^\top ds \right)$$

then, with (25),  $\pi_i$  is not summable if  $\varepsilon_i$  is not, i.e.

$$\sum_{i=1}^{\infty} \lambda_{\min} \left( \int_{(i-1)T}^{iT} \phi(s)\phi(s)^\top ds \right) = +\infty .$$

Now, let  $\varepsilon$  be fixed and, with  $t_0 = 0$ , let  $t_i$  be the smallest time such that the Gram matrix on the time window  $[t_{i-1}, t]$  is larger than the level  $\varepsilon$ , i.e.

$$t_i = \min_{t : \lambda_{\min} \left( \int_{t_{i-1}}^t \phi(s)\phi(s)^\top ds \right) \geq \varepsilon} t$$

then, with (25),  $\pi_i$  is not summable if

$$\sum_{i=1}^{\infty} \frac{1}{1 + (t_i - t_{i-1})^2} = +\infty .$$

## References

- [1] D. Aeyels, R. Sepulchre. On the convergence of a time-variant linear differential equation arising in identification. *Kybernetika* (1994) Vol. 30, N. 6, pp. 715-723
- [2] H.-F. Chen, L. Guo. Identification and stochastic adaptive control. Springer Science+Business, LLC 1991.
- [3] R. Marino, P. Tomei. Nonlinear control design, geometric, adaptive and robust. Prentice Hall 1995.
- [4] A. Morgan, K. Narendra. On the uniform asymptotic stability of certain linear nonautonomous differential equations. *SIAM J. Control and Optimization* Vol. 15, No. 1, January 1977
- [5] W. Ren, P. Kumar. Stochastic adaptive prediction and model reference control. *IEEE Transactions on Automatic Control*, Vol. 39. No. 10. october 1994 2047

## 4 History of the versions

### Modifications on January 20, 2017

Addition of the reference [1] suggested by Romeo Ortega and of the comment on how  $\pi_i$  can be chosen constant when the assumption proposed in that paper holds.

### Modifications on January 15, 2017

Addition of the discussion on the relation between non summability of  $\pi_i$  and the persistent excitation (spanning) condition.

### Modifications on January 13, 2017

The lower bounds (24) and (26) have been changed to follow a suggestion of Romeo Ortega of giving less conservative lower bound for  $\pi_i$ . Before they were, with  $t_{i+1} - t_i \leq T$ ,

– for the continuous time case,

$$\pi_i = \frac{1}{2(1 + T^2 \bar{\phi}^4)} \lambda_{\min} \left( \int_{t_{i-1}}^{t_i} \phi(u) \phi(u)^\top du \right)$$

– for the discrete time case

$$\pi_i = \frac{2 - \bar{\phi}^2}{2T(1 + T\bar{\phi}^4)} \lambda_{\min} \left( \sum_{1+t_{i-1}}^{t_i} \phi(u) \phi(u)^\top \right)$$

### Modifications on may 18, 2018

The paragraphs “Other relations” have been added. They are actually extracted from a previous note.