Home Exercises for the course

"Identification of Parameters, Filtering, Prediction and Smoothing of Dynamic Models"

Exercise 1 *Let us consider the following ARMA model*

$$\boxed{\left.\begin{array}{c} x_{n+1} = ax_n + bu_n + \zeta_n, \\ \zeta_n = \xi_n + d_1\xi_{n-1} + d_2\xi_{n-2}, \\ x_n, u_n, \zeta_n \in R^1, \ n = 0, 1, ..., \end{array}\right\}} \tag{1}$$

with

$$\left.\begin{array}{c} a = 0.5, \ b = 1, \ d_1 = 0.3, \ d_2 = -0.33, \\ x_0 = 5, \ u_n = \sin(0.3n), \end{array}\right\} \tag{2}$$

and $\{\xi_n\}_{n=0,1...}$ *is a stationary sequence of **independent Gaussian** random variables satisfying*

$$\left.\begin{array}{c} E\{\xi_n\} = 0, \ E\{\xi_n^2\} = \sigma^2 = 1, \\ E\{\xi_n\xi_k\} \overset{n\neq k}{=} 0, \ E\{\xi_n x_n\} = 0, \\ E\{\xi_n u_n\} = E\{\xi_{n-1}u_n\} = E\{\xi_{n-2}u_n\} = 0. \end{array}\right\} \tag{3}$$

The model (1) can be represented in the generalized regression format as

$$\boxed{\left.\begin{array}{c} x_{n+1} = c^{\intercal} z_n + \zeta_n, \\ c := \begin{pmatrix} a \\ b \end{pmatrix}, \ z_n := \begin{pmatrix} x_n \\ u_n \end{pmatrix} - \textit{ generalized regression vector.} \end{array}\right\}} \tag{4}$$

Problem: estimate the vector c ***using in each time*** n ***the observations*** (x_{n+1}, z_n).

To do that let us apply the Instrumental Variable (IV) estimation algorithm

$$\boxed{\left.\begin{array}{c} c_{n+1} = c_n + \Gamma_n v_n (x_{n+1} - z_n^{\intercal} c_n), \ c_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ \Gamma_n = \Gamma_{n-1} - \dfrac{\Gamma_{n-1} v_n z_n^{\intercal} \Gamma_{n-1}}{1 + z_n^{\intercal}\Gamma_{n-1} v_n}, \ z_n^{\intercal}\Gamma_{n-1} v_n \neq -1, \\ \Gamma_0 = \rho^{-1} I_{2\times 2}, \ \rho = 10^{-5}. \end{array}\right\}} \tag{5}$$

Notice that

- *for* $v_n = z_n$ *this is* ***Least Squares Method*** *(LSM);*
- *for* $v_n = z_{n-2}$ *this is* ***Instrumental Variables Method*** *(IVM);*

Show (by numerical simulations) that LSM method does not work in this example, but IVM correctly estimates unknown parameter c, namely,

$$c_n \underset{n\to\infty}{\to} c = \begin{pmatrix} 0.5 \\ 1 \end{pmatrix}.$$

Exercise 2 *Let us consider the following RAR (Regression-Auto-Regression) model with scalar output* x_{n+1} *and independent noise sequence* $\{\xi_n\}_{n=0,1,...}$*:*

$$\boxed{\left.\begin{array}{c} x_{n+1} = a_0 x_n + b u_n + \xi_n, \\ x_n, u_n, \xi_n \in R^1, \;\; n = 0, 1, ... \end{array}\right\}} \tag{6}$$

where

$$a_0 = 0.5, \;\; b = 1, x_0 = 5 \;\} \tag{7}$$

and $\{\xi_n\}_{n=0,1...}$ *is a stationary sequence of* ***independent*** random variables

$$\left.\begin{array}{c} E\{\xi_n\} = 0, \\ E\{\xi_n \xi_k\} \overset{n \neq k}{=} 0, \;\; E\{\xi_n x_n\} = 0, \\ E\{\xi_n u_n\} = E\{\xi_{n-1} u_n\} = E\{\xi_{n-2} u_n\} = 0. \end{array}\right\} \tag{8}$$

which in generalized regression format can be represented as

$$\boxed{\begin{array}{c} x_{n+1} = c^{\intercal} z_n + \xi_n, \\ c := \begin{pmatrix} a_0 \\ b \end{pmatrix}, \;\; z_n := \begin{pmatrix} x_n \\ u_n \end{pmatrix} - \text{ generalized regression vector.} \end{array}} \tag{9}$$

Problem: Find the Cramer-Rao information low-bound, that is, calculate the right-hand side of the matrix inequality

$$\boxed{\lim_{n \to \infty} n \, \mathsf{E}\{(c_n - c)(c_n - c)^{\intercal}\} \geq \mathcal{I} = \lim_{n \to \infty} n \mathbb{I}_F^{-1}(c, n)} \tag{10}$$

valid for any estimates c_n *as a function of all* ***the observations*** $(x_1, z_0; ..., x_{n+1}, z_n)$ *available at* ***time*** n *for two cases:*

Case a) $\{u_n\}_{n=0,1,...}$ *is a stationary sequence of independent random variables with* ***Laplace distribution***

$$p_u(x) = \frac{1}{a} \exp\left\{-\frac{|x|}{a}\right\}, \;\; a = 2;$$

and $\{\xi_n\}_{n=0,1,...}$ *is a stationary sequence of independent random variables with* ***Gaussian distribution***

$$p_\xi(x) = \frac{1}{\sqrt{2}\sigma} \exp\left\{-\frac{x^2}{2\sigma}\right\}, \;\; \sigma = 1;$$

Case b) $\{u_n\}_{n=0,1,...}$ *is a stationary sequence of independent random variables with* ***Gaussian distribution***

$$p_\xi(x) = \frac{1}{\sqrt{2}\sigma} \exp\left\{-\frac{x^2}{2\sigma}\right\}, \;\; \sigma = 1;$$

and $\{\xi_n\}_{n=0,1,...}$ *is a stationary sequence of independent random variables with* ***Laplace distribution***

$$p_u(x) = \frac{1}{2a} \exp\left\{-\frac{|x|}{a}\right\}, \;\; a = 2;$$

Remark: *Simulations are not required, only numerical calculation of the information low-bounds.*

Hint (help): *use the following result.*

Theorem 3 *If in the model (9) with i.i.d. (independent identically distributed) centered "noise" sequence $\{\xi_n\}$ the generalized inputs $\{z_n\}$ satisfy the following conditions*

1) *"strong law of large number " (SLNL) for $\{z_n\}$*

$$\boxed{\left\| \frac{1}{n} \sum_{t=0}^{n} \{z_t z_t^\intercal\} - \frac{1}{n} \sum_{t=0}^{n} \mathsf{E}\{z_t z_t^\intercal\} \right\| \underset{n\to\infty}{\overset{a.s.}{\to}} 0} \tag{11}$$

2) *the convergence of "averaged" inputs covariation*

$$\boxed{\frac{1}{n} \sum_{t=0}^{n} \mathsf{E}\{z_t z_t^\intercal\} \underset{n\to\infty}{\to} \mathcal{R} > 0} \tag{12}$$

3) *z_n (under a fixed prehistory $z^{n-1} := (z_1, ..., z_{n-1})$ and fixed x_n) does not depend on c, that is,*

$$\nabla_c \ln p_{z_n}\left(v_n^z \mid x_n, z^{t-1}, c\right) = 0 \ (n = 1, ...)$$

*then **the information bound** under the regular data $y_n := \left(x_{n+1}^\intercal, z_n^\intercal\right)^\intercal \in R^{1+K}$ is*

$$\boxed{\mathcal{I} = \lim_{n\to\infty} n\mathbb{I}_F^{-1}(c, n) = \mathcal{R}^{-1} I_F^{-1}(p_\xi)} \tag{13}$$

where

$$\boxed{I_F(p_\xi) := \mathsf{E}\left\{\left[(\ln p_\xi(\xi))'\right]^2\right\} = \int_{v\in R} \frac{\left[\frac{d}{dv} p_\xi(v)\right]^2}{p_\xi(v)} dv} \tag{14}$$

Fisher information is

a) for Gaussian noise $\xi_n : I_F(p_\xi) = \sigma^{-2}$, $\mathsf{E}\{\xi_t^2\} = \sigma^2$,

b) for Laplace noise $\xi_n : I_F(p_\xi) = a^{-2}$, $\mathsf{E}\{\xi_t^2\} = 2a^2$.

To solve the problem it is sufficient to calculate

$$\mathcal{R} = \lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n} \mathsf{E}\{z_t z_t^\intercal\} = \lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n} \begin{bmatrix} \mathsf{E}\{x_t^2\} & \mathsf{E}\{x_t u_t\} \\ \mathsf{E}\{u_t x_t\} & \mathsf{E}\{u_t^2\} \end{bmatrix}.$$

Exercise 3
Recurrent version of the *Maximum Likelihood Estimating Procedure:*

$$\boxed{\begin{gathered} c_n = c_{n-1} - \Gamma_n I_F^{-1}\left(p_\xi\right) z_n \frac{d}{dv} \ln p_\xi\left(x_{n+1} - c_{n-1}^\intercal z_n\right) \\ \Gamma_{n+1} = \Gamma_n - \frac{\Gamma_n z_{n+1} z_{n+1}^\intercal \Gamma_n}{1 + z_{n+1}^\intercal \Gamma_n z_{n+1}}, \ n \geq n_0 \\ c_{n_0} = Z_{n_0}^{-1} V_{n_0}, \ \Gamma_{n_0} = \left(\sum_{t=0}^{n_0} z_t z_t^\intercal\right)^{-1} = Z_{n_0}^{-1} \end{gathered}} \tag{15}$$

If for the scalar model (9) the specific conditions hold, then the procedure (15) is **asymptotically efficient** (**the best one**) under any regular (not only Gaussian) **i.i.d. noise** in the dynamics of the system.

Calculate and prepare the corresponding the best nonlinear transformations

$$\boxed{\varphi\left(v\right) = \varphi^*\left(v\right) := -\mathbb{I}_F^{-1}\left(p_\xi\right) \frac{d}{dv} \ln p_\xi\left(v\right)} \tag{16}$$

(in this case $\mathbb{I}_F^{-1}\left(p_\xi\right) = I_F^{-1}\left(p_\xi\right)$ is a scalar) for

a) Laplace noise density

$$p_\xi\left(v\right) = \frac{1}{2a} \exp\left\{-\frac{|v|}{a}\right\}, \ a = 1; \ I_F\left(p_\xi\right) = a^{-2}$$

b) Cos^2 noise density

$$p_\xi\left(v\right) = \left\{ \begin{array}{ccc} \dfrac{\pi}{a} \cos^2\left(\dfrac{\pi}{2a}\left(v - c\right)\right) & \text{for} & |v - c| \leq a \\ 0 & \text{for} & |v - c| > a \end{array} \right], \ a = 1, \ I_F\left(p_\xi\right) = \left(\frac{\pi}{a}\right)^2$$

Take $c = 0$.

Exercise 4
Consider the following system

$$\left.\begin{gathered} y(k) = 0.85y(k-1) + 2u(k) + \eta(k), \\ y(0) = 3, \eta(0) = \xi(0) = 0, \\ \\ \eta(k) = -0.3\eta(k-1) + \xi(k) + 0.8\xi(k-1), \\ u(k) = \sin(0.2k), \end{gathered}\right\} \tag{17}$$

with ξ as an independent random sequence having the *Logistic distribution*

$$p_\xi\left(v \mid \mu, \sigma\right) = \frac{\exp\left\{\dfrac{v-\mu}{\sigma}\right\}}{\sigma\left(1 + \exp\left\{\dfrac{v-\mu}{\sigma}\right\}\right)^2}, \quad -\infty < x < \infty \tag{18}$$

where $\mu \in (-\infty, \infty)$ is mean value and σ which is a scalar parameter. For simulation take

$$\mu = 0, \ \sigma = 1.$$

The system (17) can be rewritten as follows

$$\left. \begin{array}{c} y(k) = z(k)^{\top} c + \eta(k), \ k = 1, 2, ... \\ \\ \eta(k) = H(q^{-1})\xi(k), \\ H(q^{-1}) = \dfrac{1 + 0.8q^{-1}}{1 + 0.3q^{-1}}, \ q^{-1} x\,(k) := x\,(k-1) \end{array} \right\} \tag{19}$$

with

$$z(k) = \begin{pmatrix} y(k-1) \\ u(k) \end{pmatrix}, \ c := \begin{pmatrix} 0.85 \\ 2 \end{pmatrix}.$$

The whitening process is then given by

$$\tilde{y}(k) = H^{-1}(q^{-1}) y(k), \ \tilde{z}(k) = H^{-1}(q^{-1}) z(k),$$

or in the extended form,

$$\begin{array}{c} \tilde{y}(k) + 0.8\,\tilde{y}(k-1) = y(k) + 0.3y(k-1), \ \tilde{y}(0) = y(0), \\ \tilde{z}(k) + 0.8\,\tilde{z}(k-1) = z(k) + 0.3z(k-1), \ \tilde{z}(0) = z(0), \end{array}$$

where the "inverse filter" has the transfer function

$$H^{-1}(q^{-1}) = \frac{1 + 0.3q^{-1}}{1 + 0.8q^{-1}}.$$

The recursive WLSM (Whitening-Least-Square-Method) algorithm is

$$\left. \begin{array}{c} c_n = c_{n-1} - I_{F,\xi}^{-1} \Gamma_n \tilde{z}_n \dfrac{p_{\xi}'(v)}{p_{\xi}(v)} \,|_{v = \tilde{y}_n - \tilde{z}_n^{\intercal} c_{n-1}}, \\ \\ \Gamma_n = \Gamma_{n-1} - \dfrac{\Gamma_{n-1} \tilde{z}_n \tilde{z}_n^{\intercal} \Gamma_{n-1}}{1 + \tilde{z}_n^{\intercal} \Gamma_{n-1} \tilde{z}_n}, \ n = 1, 2, ..., \ \Gamma(0) = 10^5. \end{array} \right\} \tag{20}$$

Here we need to calculare the Fisher Information $I_{F,\xi}$. Take $c(0) = 2$.

Task: To simulate this identification WLSM-process with whitening and without, and demonstrate that LSM without whitening does not work, using Matlab for generating the Logistic distribution:

pd = makedist('Logistic','mu',mu,'sigma',sigma),

and

pd = fitdist(x,'Logistic').

Exercise 5 (Recurrent Residual Method (RRM))
Consider the plant

$$\boxed{\left.\begin{array}{c} x_{n+1} = ax_n + bu_n + \xi_n + d\xi_{n-1}, \\ x_n, u_n, \xi_n \in R^1,\ n = 0, 1, ..., \\ n = 0 : x_0 = 0,\ \xi_{-1} = 0. \end{array}\right\}} \tag{21}$$

Problem: estimate numerically parameters $c = (a, b, d)^{\intercal}$, using the algorith

$$\boxed{\begin{array}{c} c_n = c_{n-1} - \Gamma_n z_n \left(x_{n+1} - c_{n-1}^{\intercal} z_n\right) = c_{n-1} - \Gamma_n z_n \varepsilon_{n+1}, \\ \\ \Gamma_{n+1} = \Gamma_n - \dfrac{\Gamma_n z_{n+1} z_{n+1}^{\intercal} \Gamma_n}{1 + z_{n+1}^{\intercal} \Gamma_n z_{n+1}},\ n \geq n_0, \\ \\ c_{n_0} = Z_{n_0}^{-1} V_{n_0},\ \Gamma_{n_0} = \left(\sum\limits_{t=0}^{n_0} z_t z_t^{\intercal}\right)^{-1} = Z_{n_0}^{-1}, \end{array}} \tag{22}$$

where

$$\boxed{z_n = (x_n, u_n, \varepsilon_n)^{\intercal},} \tag{23}$$

with the sequence $\{\varepsilon_n\}$ generated by the recurrency

$$\boxed{\begin{array}{c} \varepsilon_{n+1} = x_{n+1} - c_{n-1}^{\intercal} z_n, \\ n = 1, 2, ...;\ \varepsilon_1 = 0. \end{array}} \tag{24}$$

Here ξ_n is the *Standard Gaussian random value* ($\mathsf{E}\{\xi_n\} = 0$, $\mathsf{E}\{\xi_n^2\} = \sigma^2$) and real values (used for simulations) of the parameters are:

$$\boxed{a = 0.5, b = -1, d = -0.7.}$$

The input u_n is

$$\boxed{u_n = 0.6 \sin(0.1n).}$$

Present the figures with $c_n = (a_n, b_n, d_n)^{\intercal}$ and $\Delta_n = \varepsilon_n - \xi_n$ which should satisfy

$$\Delta_n \underset{n\to\infty}{\overset{a.s.}{\to}} 0.$$

Exercise 6 (Huber's robust identifiers)
In the case of dynamic *autoregression* model (ARX-model) where the generalized inputs are dependent on the state of the system, the matrix $\mathcal{R}$ depends on p_ξ too, and therefore, we deal with the complete problem (**??**), namely, we need to calculate

$$\sup_{p_\xi \in \mathcal{P}} \left[I_{F,\xi}(p_\xi)\, \mathcal{R}(p_\xi)\right]^{-1} \tag{25}$$

and to find the worth distribution p_ξ^* within the considered class $\mathcal{P}$. For the AR-model

$$y_{n+1} = \sum_{s=0}^{L_a} a_s y_{n-s} + \xi_n = \mathbf{c}^\intercal \mathbf{v}_n + \xi_n$$

$$\mathbf{c}^\intercal = (a_0, ..., a_{L_a}), \ \mathbf{v}_n^\intercal = (y_n, y_{n-1}, ..., y_{n-L_a})$$

we have

$$\frac{1}{n} \sum_{t=0}^{n} \mathsf{E}\left\{\mathbf{v}_t \mathbf{v}_t^\intercal\right\} \to \mathcal{R}$$

where $\mathcal{R}$ satisfies

$$\mathcal{R} = A\mathcal{R}A + \sigma^2 \Xi_0$$

with

$$A = \left\| \begin{array}{ccccc} a_0 & a_1 & \cdots & \cdots & a_{L_a} \\ \mathbf{1} & 0 & \cdots & \cdots & 0 \\ 0 & \mathbf{1} & 0 & \cdots & 0 \\ 0 & \cdots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & \mathbf{1} & 0 \end{array} \right\|, \ \Xi_0 := \left\| \begin{array}{ccccc} \mathbf{1} & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 0 & 0 \end{array} \right\|$$

Obviously, $\mathcal{R}$ can be represented as

$$\mathcal{R} = \sigma^2 \mathcal{R}_0,$$

where $\mathcal{R}_0$ is the solution of

$$\mathcal{R} = A\mathcal{R}A + \Xi_0$$

so that the problem (25) is reduced to

$$\sup_{p_\xi \in \mathcal{P}} \left[\sigma^2 (p_\xi) I_F (p_\xi)\right]^{-1}$$

or equivalently, to

$$\boxed{\inf_{p_\xi \in \mathcal{P}} \left[\sigma^2 (p_\xi) I_F (p_\xi)\right]} \tag{26}$$

Problem: to design the asymptotically robust optimal identification algorithm in the format

$$\boxed{\begin{array}{c} \mathbf{c}_n = \mathbf{c}_{n-1} - \Gamma_n \mathbf{v}_n I_F^{-1}\left(p_\xi^*\right) \dfrac{d}{dv} \ln p_\xi^* (v) \mid_{v = x_{n+1} - \mathbf{c}_{n-1}^\intercal \mathbf{v}_n}, \\ \Gamma_{n+1} = \Gamma_n - \dfrac{\Gamma_n \mathbf{v}_{n+1} \mathbf{v}_{n+1}^\intercal \Gamma_n}{1 + \mathbf{v}_{n+1}^\intercal \Gamma_n \mathbf{v}_{n+1}}, \ n \geq n_0, \\ c_{n_0} = Z_{n_0}^{-1} V_{n_0}, \ \Gamma_{n_0} = \left(\sum_{t=0}^{n_0} z_t z_t^\intercal\right)^{-1} = Z_{n_0}^{-1}, \end{array}} \tag{27}$$

where

$$p_\xi^*(v) = \arg \inf_{p_\xi \in \mathcal{P}} \left[\sigma^2 \left(p_\xi \right) I_F \left(p_\xi \right) \right],$$

for the AR model

$$\boxed{\left.\begin{array}{c} x_{n+1} = a_0 x_n + a_1 x_{n-1} + \xi_n, \\ \\ x_n, \xi_n \in R^1, \;\; n = 0, 1, ..., \\ \\ a_0 = 0.5, a_1 = 0.3, x_0 = -1, x_{-1} = 0 \end{array}\right\}}$$

with noise ξ_n from the class **Class** $\mathcal{P}_2^{AR}$ (*containing all centered distributions with a variance not less than a given value*):

$$\boxed{\mathcal{P}_2^{AR} := \left\{ p_\xi : \int_{\mathbb{R}} x^2 p_\xi \left(x \right) dx \geq \sigma_0^2 \right\}.} \tag{28}$$

and in the same time from the class **Class** $\mathcal{P}_1$ (*of all symmetric distributions nonsingular in the point* $x = 0$):

$$\boxed{\mathcal{P}_1 := \left\{ p_\xi : \;\; p_\xi \left(0 \right) \geq \frac{1}{2a} > 0 \right\},} \tag{29}$$

that is

$$p_\xi \in \mathcal{P}_2^{AR} \cap \mathcal{P}_1.$$

Hint:

Lemma 4 (on the class $\mathcal{P}_2^{AR}$)

$$\boxed{p_\xi^* \left(x \right) = \arg \inf_{p_\xi \in \mathcal{P}_2^{AR} : \sigma^2 \left(p_\xi \right) = \sigma_0^2} I_F \left(p_\xi \right),} \tag{30}$$

*that is, the worth on $\mathcal{P}_2^{AR}$ distribution density $p_\xi^* \left(x \right)$ coincides with the worth distribution density on the classes $\mathcal{P}_i$ characterizing distribution uncertainties for **static** (R-models) provided that*

$$\boxed{\sigma^2 \left(p_\xi^* \left(x \right) \right) = \sigma_0^2.} \tag{31}$$

Proof. It follows directly from the inequality

$$\sigma^2 \left(p_\xi \right) I_F \left(p_\xi \right) \geq \sigma_0^2 I_F \left(p_\xi \right).$$

Lemma 5 (on the class $\mathcal{P}_1$)

$$\boxed{p_\xi^* \left(x \right) = \arg \inf_{p_\xi \in \mathcal{P}_1} I_F \left(p_\xi \right) = \frac{1}{2a} \exp \left\{ -\frac{|x|}{a} \right\},} \tag{32}$$

that is, the worth on $\mathcal{P}_1$ distribution density is the Laplace one given by (32).

Any numerical simulations are not required! Only you need to give the analitical formula for

$$I_F^{-1}\left(p_\xi^*\right)\frac{d}{dv}\ln p_\xi^*\left(v\right)$$

in the algorithm (27).

Exercise 7 (Kalman's filtering)
Consider the stochastic system

$$\boxed{\begin{array}{c} dx\left(t\right)=\left[Ax(t)+b(t)\right]dt+\Xi dW_x(t),\ x(0)=x_0, \\ \\ dy(t)=c^{\intercal}x(t)dt+rdW_y(t),\ r>0, \end{array}} \tag{33}$$

where $W_x(t)$ and $W_y(t)$ are standard (with unite variance) independent Wiener processes.

For the simulation in Simulink here may be used the, so-called, *engineering presentation* of this system as

$$\boxed{\begin{array}{c} \dot{x}\left(t\right)=Ax(t)+b(t)+\sigma\xi_x\left(t\right),\ x(0)=x_0, \\ \\ \dot{y}(t)=c^{\intercal}x(t)+r\xi_y\left(t\right),\ r>0, \end{array}}$$

where $\xi_x\left(t\right)$ and $\xi_y\left(t\right)$ are associated gaussian white noises (with zero mean and variance equal 1), namely,

$$\xi_x\left(t\right)\sim\frac{d}{dt}W_x(t),\ \xi_y\left(t\right)\sim\frac{d}{dt}W_y(t)$$

(which do not exist in rigorous mathematical sense).

To estimate the current states $\hat{x}\left(t\right)$ let us apply the Kalman filter

$$d\hat{x}\left(t\right)=\left[A\hat{x}\left(t\right)+b(t)\right]dt+L\left(t\right)\left[dy_t\left(t\right)-c^{\intercal}\hat{x}\left(t\right)dt\right],\hat{x}_0\ is\ fixed. \tag{34}$$

that In the engineering interpretation it looks as

$$\boxed{\frac{d}{dt}\hat{x}\left(t\right)=A\hat{x}\left(t\right)+b(t)+L\left(t\right)\left[\frac{d}{dt}y_t\left(t\right)-c^{\intercal}\hat{x}\left(t\right)\right],\ \hat{x}_0\text{ is fixed.}} \tag{35}$$

We assume that $\frac{d}{dt}y_t\left(t\right)$ **is available!**

Here

$$\boxed{L\left(t\right)=L^*\left(t\right):=r^{-2}P\left(t\right)c,} \tag{36}$$

where

$$\boxed{P\left(t\right)=\mathsf{E}\left\{\Delta x\left(t\right)\Delta^{\intercal}x\left(t\right)\right\}} \tag{37}$$

satisfies the following differential Riccati equation

$$\boxed{\begin{array}{c}\dot{P}(t) = AP(t) + P(t)A^{\intercal} + \Xi\Xi^{\intercal} - r^{-2}P(t)cc^{\intercal}P(t), \\ \\ P(0) = \mathsf{E}\{\Delta x(0)\Delta^{\intercal}x(0)\}\end{array}} \tag{38}$$

For simulation take

$$x(t) \in R^2,\ A = \begin{bmatrix} -1 & 1 \\ 0 & -2 \end{bmatrix},\ \Xi = 0.5\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ r = 0.9,$$
$$x(0) = \begin{pmatrix} 1 \\ -1 \end{pmatrix},\ b(t) = 0.6\begin{pmatrix} \sin(0.5t) \\ \cos(0.5t) \end{pmatrix},\ P(0) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

To analyze 3 situations:

$$1)\ c = \begin{pmatrix} 1 \\ -1 \end{pmatrix},\ 2)\ c = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ 3)\ c = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and for each of them to draw the pictures

$$\boxed{\hat{x}(t) = \begin{pmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \end{pmatrix},\ x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}}$$

($\hat{x}_i(t)$ and $x_i(t)$ in the same graphic) and

$$\boxed{\mathrm{tr}\{P(t)\} = \mathsf{E}\left\{\|\Delta x(t)\|^2\right\}.}$$