Laboratory of Adaptive Control
Institute of Automation and Robotics
Poznań University of Technology
IAR-PUT
Maciej M. Michałek, Rafał Madoński (transl.)

E3 Parametric identification with Recursive-Least-Squares

This exercise is devoted to a recursive version of the LS method of parametric identification. In comparison to the batch-type LS estimator (see exercise E3), the use of *recursive* algorithm allows one to decrease the overall computational complexity (e.g. lack of a matrix inversion) and to use the persistently updated model of a plant in real-time. The latter feature is especially important in the case of MIAC-type adaptive control techniques.

1 Identification of a dynamic plant with constant parameters

The recursive identification algorithm RLS (*Recursive* LS) results from a transformation of the batch-type LS estimator introduced in exercise E3. Therefore RLS can be considered as a recursive/recurrent version of the LS estimator, which asymptotically has statistical properties equivalent to the LS method.

Let us assume a structure of the identified dynamic plant represented by the following difference equation:

$$y(n) = \boldsymbol{\varphi}^{\top}(n)\,\boldsymbol{p}_0 + v(n), \tag{1}$$

where $v(n)$ is a random noise (white or colored), and $\boldsymbol{p}_0$ is the vector of *true* plant parameters.

Computations in the RLS method are performed using only a single set of input-output data in each iteration of the algorithm. Therefore, the current estimate $\hat{\boldsymbol{p}}(n)$ of parameters is periodically updated using the following general scheme

$$\hat{\boldsymbol{p}}(n) = \hat{\boldsymbol{p}}(n-1) + \boldsymbol{k}(n)\,\varepsilon(n) \tag{2}$$

where $\boldsymbol{k}(n)\,\varepsilon(n)$ is a correction term evaluated based on a new dataset. The term $\varepsilon(n) = y(n) - \boldsymbol{\varphi}^{\top}(n)\,\hat{\boldsymbol{p}}(n-1)$ is a current one-step-ahead prediction error (computed based on the previous estimate $\hat{\boldsymbol{p}}(n-1)$), and $\boldsymbol{k}(n) \in \mathbb{R}^d$ is a vector of time-varying gains which depends on a current covariance matrix of the estimated parameters. The full computational scheme of the RLS method is as follows:

$$\hat{\boldsymbol{p}}^{\mathrm{LS}}(n) = \hat{\boldsymbol{p}}^{\mathrm{LS}}(n-1) \;+\; \boldsymbol{k}(n)\,\varepsilon(n), \tag{3}$$

$$\varepsilon(n) = y(n) \;-\; \boldsymbol{\varphi}^{\top}(n)\,\hat{\boldsymbol{p}}^{\mathrm{LS}}(n-1), \tag{4}$$

$$\boldsymbol{k}(n) = \boldsymbol{P}^{\mathrm{LS}}(n)\,\boldsymbol{\varphi}(n), \tag{5}$$

$$\boldsymbol{P}^{\mathrm{LS}}(n) = \boldsymbol{P}^{\mathrm{LS}}(n-1) - \frac{\boldsymbol{P}^{\mathrm{LS}}(n-1)\boldsymbol{\varphi}(n)\boldsymbol{\varphi}^{\top}(n)\boldsymbol{P}^{\mathrm{LS}}(n-1)}{1+\boldsymbol{\varphi}^{\top}(n)\boldsymbol{P}^{\mathrm{LS}}(n-1)\boldsymbol{\varphi}(n)}, \tag{6}$$

where the above sequence should be computed in the following order: (6)→(5)→(4)→(3). Worth mentioning that equation (3) defines in fact a particular (computational) dynamical system, which starting from initial conditions $\hat{\boldsymbol{p}}(0)$ and $\boldsymbol{P}(0)$ will evolve in the discrete time through the transient phase until reaching the steady state (theoretically reached at infinity). After vanishing of the transient phase (in practice assumed to vanish after considerably large value of N), the estimate $\hat{\boldsymbol{p}}(n)$ in time instant $n = N-1$ should correspond to the value of the batched LS estimator computed upon the dataset of N measurements. Initial conditions $\hat{\boldsymbol{p}}(0)$ and $\boldsymbol{P}(0)$ for the recurrence (3)-(6) can be selected using various approaches. Two of them are:

W1. choice based on *a priori* knowledge of plant dynamics,

W2. arbitrary choice, e.g. $\hat{\boldsymbol{p}}(0) := \mathbf{0}$, $\boldsymbol{P}(0) := \rho \cdot \boldsymbol{I}$, where $\rho \gg 0$, and $\boldsymbol{I} \in \mathbb{R}^{d \times d}$ is the identity matrix.

The inversion of matrix $\boldsymbol{P}(0)$ can be interpreted as a level of confidence for initial estimate $\hat{\boldsymbol{p}}(0)$; moreover $\mathrm{cov}[\hat{\boldsymbol{p}}(n)] = \sigma^2 \boldsymbol{P}(n)$. One should also notice the crucial feature of the covariance matrix computed based on equation (6), namely

$$\boldsymbol{P}(n) \overset{n \to \infty}{\longrightarrow} \mathbf{0}, \tag{7}$$

which means that estimate $\hat{\boldsymbol{p}}(n)$ for $n \to \infty$ will converge to some constant terminal value $\hat{\boldsymbol{p}}_{\mathrm{lim}}$ (with probability equal to 1). The terminal estimate $\hat{\boldsymbol{p}}$ will be close to the true parameters $\boldsymbol{p}_0$ if the following assumptions are satisfied:

A1. noise $v(n)$ in structure (1) is white,

A2. structures of the model and the plant are the same,

A3. input signal $u(n)$ is *persistently exciting*.

1.1 Dynamic plant identification using the RLS method.

- File `PlantARMAX.mdl` contains a discrete-time dynamic plant described by the following structure:

$$y(n) = \frac{b_{20} q^{-2}}{1 + a_{10} q^{-1} + a_{20} q^{-2}} u(n) + \frac{1 + c_{10} q^{-1}}{1 + a_{10} q^{-1} + a_{20} q^{-2}} e(n), \tag{8}$$

  where a_{10}, a_{20}, b_{20} and c_{10} represent the true plant parameters, while $e(n)$ is a white noise. Note that (8) belongs to the ARMAX model family of the following general form: $Ay = Bu + Ce \;\Rightarrow\; y = Gu + He, \quad G = B/A,\ H = C/A$. If we assume that $c_{10} := 0$, then we obtain the particular ARX form of the model with white noise signal on the right-hand side. For $c_{10} \neq 0$ the noise $(1 + c_{10} q^{-1}) e(n)$ is colored with all the consequences of this fact.

- Initialize the following global variables: `Tp=0.1, Tend=1000, Td=1500` which represent, respectively, the sampling time, the simulation time, and the time when parameter b_{20} will intentionally change (here `Tend<Td`, so the change will never occur – the plant with constant parameters).

- Initialize (as a global variable) the following true plant parameter $c_{10} = 0$ (that is, we assume that $v(n)$ in (1) is a white noise). Note that c_{10} is unknown in practice! Perform the identification procedure of plant (8) using the RLS method and applying a symmetrical square-wave input signal $u(n)$ with amplitude $A = 1$ and frequency $f_u = 0.2\,\mathrm{Hz}$. Analyze estimates $\hat{\boldsymbol{p}}(n)$ for various values of parameter ρ in a choice of initial matrix $\boldsymbol{P}(0)$ (see W2). Investigate influence of sampling time T_p to the quality of identification – the suggested set of values is (in [s]):

$$T_p \in \{0.01; 0.1; 0.5; 1.0; 2.0\}.$$

- Implement the *simulated model* computed upon a current estimate of parameters. Compare an output of the plant $y(n)$ (as well as $y_0(n)$) with an output of the simulated model in response to the same input signal $u(n)$. Assess quality of identification.
- Analyze a time-plot of a *trace* of covariance matrix $\boldsymbol{P}(n)$ during the identification process (use command `trace(P)`).
- Initialize the true parameter $c_{10} = 0.7$ (that is, we assume that $v(n)$ in (1) is a colored noise). Perform the RLS identification procedure again and analyze time-evolution of estimate $\hat{\boldsymbol{p}}(n)$.

2 Identification of a dynamic plant with time-varying parameters

Until now, we assumed that parameters of the plant were constant during plant operation. Such an assumption, however, is not always valid. Numerous examples can be pointed out in which the dynamics changes over time, either abruptly but occasionally (e.g. once for some period of time) or slowly but persistently[1]. In that case, the ability of the algorithm to track the time-varying parameters in real time in order to possess the up-to-date model of the plant is highly desirable. The method presented in the previous section did not have the ability to adapt to the parameter variation, since the trace of the matrix $\boldsymbol{P}(n)$ converges asymptotically to zero in time and hence the ability to correct the vector of estimated parameters vanishes in time. In order to keep the tracking (adaptation) ability of the estimator one has to assure that matrix $\boldsymbol{P}(n)$ does not converge to zero in time. It can be done using two basic approaches:

- by introducing the so-called *forgetting factor* $\lambda \in (0, 1)$ (usually used in the case of slowly-varying parameters) to matrix $\boldsymbol{P}(n)$; equations of the RLS$_\lambda$ method take the following form:

$$\hat{\boldsymbol{p}}(n) = \hat{\boldsymbol{p}}(n-1) \; + \; \boldsymbol{k}(n)\,\varepsilon(n), \tag{9}$$

$$\varepsilon(n) = y(n) \; - \; \boldsymbol{\varphi}^\top(n)\,\hat{\boldsymbol{p}}(n-1), \tag{10}$$

$$\boldsymbol{k}(n) = \boldsymbol{P}(n)\,\boldsymbol{\varphi}(n), \tag{11}$$

$$\boldsymbol{P}(n) = \frac{1}{\lambda}\left[\boldsymbol{P}(n-1) - \frac{\boldsymbol{P}(n-1)\boldsymbol{\varphi}(n)\boldsymbol{\varphi}^\top(n)\boldsymbol{P}(n-1)}{\lambda + \boldsymbol{\varphi}^\top(n)\boldsymbol{P}(n-1)\boldsymbol{\varphi}(n)}\right], \tag{12}$$

- by *reseting* the covariance matrix $\boldsymbol{P}$, which means *re-initializing* matrix $\boldsymbol{P}$ when a prescribed condition is satisfied (usually used in order to track the parameters that change rarely but abruptly):

  R1. periodic reseting

$$\boldsymbol{P}(n) := \rho\boldsymbol{I} \qquad \text{when} \qquad n = kT,\ k = 1, 2, \ldots \tag{13}$$

  R2. reseting using a condition related to the prediction error or output error

$$\boldsymbol{P}(n) := \rho\boldsymbol{I} \qquad \text{when} \qquad |\varepsilon(n)| > \varepsilon_{\max} \quad \text{or} \quad |\varepsilon_{\mathrm{OE}}(n)| > \varepsilon_{\max} \tag{14}$$

  R3. reseting using a condition related to the trace of the covariance matrix

$$\boldsymbol{P}(n) := \rho\boldsymbol{I} \qquad \text{when} \qquad \mathrm{tr}(\boldsymbol{P}(n)) < P_{\min}, \tag{15}$$

  where $\rho \gg 0$, $T, \varepsilon_{\max} > 0$, and $P_{\min} > 0$ are all the design parameters.

[1] We assume a change of parameters only – a structure of the model remains the same.

The cost one pays for introducing of any from the above modifications corresponds to the increased variance of the parameters estimates during the identification process. Hence, the modifications corresponding to the choice of values for λ, $\rho, T, \varepsilon_{\max}$, and $P_{\min}$ should result from a compromise between the tracking ability of the estimator and a level of fluctuations of the estimated parameters.

2.1 ***Adaptive*** **identification of a dynamic plant using the modified RLS method.**

- Initialize parameters `Tend=1000, Td=500` (a step change of parameter $b_{20}(n)$ of plant (8) will take place at time instant $n = 500$), and $c_{10} = 0$ (that is, we assume that $v(n)$ in (1) is a white noise).
- Using the RLS$_\lambda$ method, perform the identification procedure of plant (8) selecting the forgetting factor λ from the range $[0.98; 0.999]$. Verify the influence of forgetting factor $\lambda \in (0, 1)$ on the identification quality and on fluctuations of the estimated parameters. Pay special attention on the tracking ability of the estimator with respect to parameter $b_{20}(n)$. Analyze a time-plot of a *trace* of covariance matrix $\boldsymbol{P}(n)$ during the identification process.
- Analyze quality of the adaptive identification when applying the covariance matrix resetting (utilize criteria R1 to R3). Verify the influence of coefficient ρ on the quality of identification and on time-variability of the estimated parameters.
- Verify the identification quality of the RLS$_\lambda$ method in the case of the colored noise (that is, for $c_{10} = 0.7$).

□