Recursive identification of multivariable systems using matchable–observable linear models

Rodrigo Alvite Romano and Felipe Pait

Preamble

This is a report about recursive parameter estimation using the parameterization proposed in the context of adaptive control theory.

Matchable–observable parameterizations

Any strictly proper linear system of McMillan degree n_x with n_u inputs ($u_k \in \mathbb{R}^{n_u}$) and n_y outputs ($y_k \in \mathbb{R}^{n_y}$) can be represented by the following state–space model structure

$$x_{k+1} = \left(A + D(\theta)\left(I - G(\theta)\right)^{-1} C\right) x_k + B(\theta) u_k \tag{1}$$

$$y_k = \left(I - G(\theta)\right)^{-1} C x_k, \tag{2}$$

where θ denotes the model parameters, I is an identity matrix of suitable dimensions, (C, A) is an arbitrary observable pair, and A is Hurwitz. The parameter matrix $G(\theta) \in \mathbb{R}^{n_y \times n_y}$ is strictly lower triangular. The other parameter matrices $D(\theta)$ and $B(\theta)$ take values in $\mathbb{R}^{n_x \times n_y}$ and $\mathbb{R}^{n_x \times n_u}$, respectively. This parameterization proposed in [1] in the context of adaptive control have the properties of observability, match-–point controllability, and matchability. As a consequence, no undesired pole-–zero cancellations can appear, the number of model parameters is reduced if compared to fully–parameterized models, and linear least—squares parameter estimation methods is applicable in off–line batch identification [2] or in a recursive context. Our aim is to present an algorithm based on the parameterization (1)–(2) suitable for real–time applications, in which the model parameters must be updated recursively.

[1] A. S. Morse and F. M. Pait. MIMO design models and internal regulators for cyclicly switched parameter–adaptive control systems. *IEEE Transactions on Automatic Control*, 39(9):1809–1818, 1994

[2] R. A. Romano, F. Pait, and C. Garcia. Multivariable system identification using an output-injection based parameterization. In *Proceedings of the 9th IEEE International Conference on Control & Automation*, pages 53–58, Santiago, 2011

Now, suppose that the parameter–independent pair (C, A) is given by

$$C = \text{block diagonal}\left\{C_1, C_2, \ldots, C_{n_y}\right\} \tag{3}$$

$$A = \text{block diagonal}\left\{A_1, A_2, \ldots, A_{n_y}\right\}, \tag{4}$$

where the pairs (C_i, A_i) are constructed such that

1. Given a list of observability indices

$$l = \left\{n_1, \ldots, n_{n_y}\right\},$$

so that $n_1 + \ldots + n_{n_y} = n_x$, choose an arbitrary stable monic polynomial $\alpha(q)$ of degree $\underline{n} = \max(l)$, such that, for $i = \{1, \ldots, n_y\}$, $\alpha(q)$

has a real monic factor $\lambda_i(q)$ of degree n_i. The symbol q denotes the shift operator.

2. The n_i–dimensional observable pairs (C_i, A_i) are composed of matrices A_i in right companion form, whose characteristic polynomial is $\lambda_i(q)$, and $C_i = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}$.

As detailed in a previous work [3], such choice for (C, A) is particularly compelling for system identification as it enables us to write a predictor for the outputs of (1)–(2) in regression form, such that the entries of the parameter matrices $B(\theta)$, $D(\theta)$ and $G(\theta)$ appear linearly, and the information vector φ is essentially composed of delayed versions of filtered input–output samples. This result is formally presented as follows.

[3] R. A. Romano and F. Pait. Matchable—observable linear models and direct filter tuning: An approach to multivariable identification. Submitted to the IEEE Transactions on Automatic Control, 2015

Proposition 1. *Let*

$$\begin{aligned}\xi_k = \Big[& y^f_{1,k-\underline{n}} \ \cdots \ y^f_{1,k-1} \ \cdots \ y^f_{n_y,k-\underline{n}} \ \cdots \ y^f_{n_y,k-1} \\ & u^f_{1,k-\underline{n}} \ \cdots \ u^f_{1,k-1} \ \cdots \ u^f_{n_u,k-\underline{n}} \ \cdots \ u^f_{n_u,k-1} \Big]^T, \end{aligned} \tag{5}$$

where

$$\begin{aligned} y^f_{i,k} &= \frac{q^{\underline{n}}}{\alpha(q)} y_{i,k} &&, \text{for } i = \{1, \ldots, n_y\} \\ u^f_{j,k} &= \frac{q^{\underline{n}}}{\alpha(q)} u_{j,k} &&, \text{for } j = \{1, \ldots, n_u\}, \end{aligned}$$

are filtered versions of the ith output and input sequences at instant k, respectively. Furthermore, for each $i \in \{1, \ldots, n_y\}$, introduce

$$\mathcal{M}_i = \text{block diagonal}\{\underbrace{M_i, \ldots, M_i}_{n_u+n_y \text{ times}}\},$$

where $M_i \in \mathbb{R}^{\underline{n} \times n_i}$ is a Toeplitz matrix of the form

$$M_i = \begin{bmatrix} \kappa_{i,1} & 0 & \cdots & 0 \\ \vdots & \kappa_{i,1} & & \vdots \\ \kappa_{i,\underline{n}-n_i} & \vdots & \ddots & 0 \\ 1 & \kappa_{i,\underline{n}-n_i} & & \kappa_{i,1} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \kappa_{i,\underline{n}-n_i} \\ 0 & \cdots & 0 & 1 \end{bmatrix},$$

whose entries are given by the coefficients of the polynomial

$$\kappa_i(q) = \frac{\alpha(q)}{\lambda_i(q)} = q^{\underline{n}-n_i} + \kappa_{i,\underline{n}-n_i} q^{\underline{n}-n_i-1} + \ldots + \kappa_{i,2} q + \kappa_{i,1}.$$

Then, a predictor for the ith output at instant k, namely $\hat{y}_{i,k}$, based on (1)–(2) *can be written as*

$$\hat{y}_{i,k} = \varphi^T_{i,k} \theta_i, \tag{6}$$

where the regressor or information vector relative to the ith output is

$$\varphi_{i,k} = \begin{cases} \mathcal{M}_1^T \xi_k, & \textit{for } i = 1 \\ \begin{bmatrix} \xi_k^T \mathcal{M}_i & y_{1,k} & \cdots & y_{i-1,k} \end{bmatrix}^T, & \textit{for } i \neq 1 \end{cases},$$

and the respective parameter vector is

$$\theta_i = \begin{cases} \mathrm{vec}\{\begin{bmatrix} D_i(\theta) & B_i(\theta) \end{bmatrix}\}, & \textit{for } i = 1 \\ \begin{bmatrix} \mathrm{vec}\{\begin{bmatrix} D_i(\theta) & B_i(\theta) \end{bmatrix}\}^T & g_{i1} & \cdots & g_{i(i-1)} \end{bmatrix}^T, & \textit{for } i \neq 1 \end{cases},$$

where $\begin{bmatrix} D_i(\theta) & B_i(\theta) \end{bmatrix}$ *is the ith submatrix of* $\begin{bmatrix} D(\theta) & B(\theta) \end{bmatrix}$, *the latter being partitioned according to the list of observability indices, the operator* $\mathrm{vec}\{\cdot\}$ *stacks the columns of the argument on top of each other, and the symbols* g_{ij} *denote the nonzero elements of the matrix* $G(\theta)$.

Proof. The proof follows directly from Lemmas 1 and 2 in [4]. □

[4] R. A. Romano and F. Pait. Matchable—observable linear models and direct filter tuning: An approach to multivariable identification. Submitted to the IEEE Transactions on Automatic Control, 2015

In fact, given the polynomial $\alpha(q)$, the regression form (6) is linear in θ_i. Therefore, the parameters of (1)–(2) can be efficiently estimated using linear least–squares methods.

Remark: The roots of $\alpha(q)$ determine the poles of the low–pass filter used to generate ξ_k (which, in turn, is part of the regressor vector) from input-–output data. Hence, they are variables that can be designed to remove high frequency disturbances that we do not want to include in the modeling. In other words, fixing the polynomial $\alpha(q)$ is equivalent to specifying a prefilter without taking into account the process dynamics.

Recursive parameter estimation

The state–space description (1)–(2) can be straightforwardly extended to the time–varying context by considering the following parameter variation model

$$\theta_{i,k} = \theta_{i,k-1} + w_{i,k} \tag{7}$$
$$y_{i,k} = \varphi_{i,k}^T \theta_{i,k-1} + v_{i,k}, \tag{8}$$

where $\theta_{i,k} \in \mathbb{R}^{n_i(n_u+n_y)}$ is the parameter vector relative to the ith output at instant k, whose change is driven by a zero–mean Gaussian process $w_{i,k}$ with nonnegative definite covariance matrix Q_i. The measurement noise $v_{i,k}$ is assumed to be a white Gaussian process with variance R_i. The random-walk parameter variation model (7)–(8) is the most common parameter–varying description. For other possible choices, the reader is refereed to [5].

[5] L. Ljung and S. Gunnarsson. Adaptation and tracking in system identification — a survey. *Automatica*, 26(1):7–21, 1990

Applying a Kalman filter to (7)–(8) provides the set of recursive relations (see Lemma 2.2 in [6])

[6] L. Ljung and T. Söderström. *Theory and Practice of Recursive Identification*. The MIT Press, Cambridge, Massachusetts, 1983

$$\hat{\theta}_{i,k} = \hat{\theta}_{i,k-1} + K_{i,k}\left(y_{i,k} - \varphi_{i,k}^{T}\hat{\theta}_{i,k-1}\right) \tag{9}$$

$$K_{i,k} = \frac{P_{i,k-1}\varphi_{i,k}}{R_i + \varphi_{i,k}^{T}P_{i,k-1}\varphi_{i,k}} \tag{10}$$

$$P_{i,k} = P_{i,k-1} - \frac{P_{i,k-1}\varphi_{i,k}\varphi_{i,k}^{T}P_{i,k-1}}{R_i + \varphi_{i,k}^{T}P_{i,k-1}\varphi_{i,k}} + Q_i. \tag{11}$$

Therefore, ...