Kalman Filter for Observer-ARMA Model with Parameter Estimation

Juan Sebastián Cárdenas-Rodríguez
Mathematical Engineering
Universidad EAFIT
Medellín, Colombia
jscardenar@eafit.edu.co

David Plazas Escudero
Mathematical Engineering
Universidad EAFIT
Medellín, Colombia
dplazas@eafit.edu.co

***Abstract*—In this work we suggest a novel extension of the standard Kalman-Filter (KF), applied to parameter estimation and output-ARMA models. The procedure includes an extended KF for the bilinear form whilst estimating parameters and applying instrumental variables to eliminate state correlation, which can be reduced to a re-sample of the available data and apply the known KF. We include a short discussion regarding this method and some future work suggestions.**

***Index Terms*—State and parameter estimation, Kalman-Filter, instrumental variables, ARMA, extended Kalman-Filter.**

I. Introduction

Autoregressive Moving-Average (ARMA) models are widely used in areas such as economics and finance, with applications in modelling time series. The main problems with ARMA models is the estimation of its parameters, using only state data. One important remark is that ARMA models do not consider an observer, which cannot be always assumed in real conditions.

Classical parameter estimation of ARMA models is often addressed with standard statistical methods, for instance, instrumental least-square (ILS) method. The main problem with the ILS method is the requirement of the complete state of the real system, which is not always achievable since the system is perceived only through the observer.

This work is focused on the design of a Kalman-based filter for the estimation of ARMA parameters through the output of the observer, taking into consideration that the disturbances in the ARMA models can be correlated and this violates one of the main hypothesis of the standard Kalman-Filter (KF). We propose in this work a filter based on instrumental variables and extended KF.

II. Preliminaries

A. ARMA Model

The general scheme for the observer-ARMA model is given by

$$\begin{aligned} z_{t+1} &= \sum_{s=0}^{l_1} a_s z_{t-s} + \sum_{s=0}^{l_2} b_s \xi_{t-s} \\ y_t &= h_t z_t + \zeta_t \\ z_0 &\in \mathbb{R},\ t = 0, 1, \ldots \end{aligned} \tag{1}$$

where ξ_t and ζ_t are random perturbations.

B. The Kalman-Filter State-Observer Model

The KF state-observer model, given by

$$\begin{aligned} x_{t+1} &= F_t x_t + \xi_t \\ y_t &= H_t x_t + \zeta_t \\ x_0 &\in \mathbb{R}^n,\ t = 0, 1, \ldots \end{aligned} \tag{2}$$

where ξ_t and ζ_t are random noises. Every ARMA model can be translated into a form of system (2).

C. Traditional KF

We will present, first, some notation in order to write the instrumental KF:

- Observed data vector: $y_t^* := (y_0^T, y_1^T, \ldots, y_t^T)^T$.
- $z_{t|t-1} := E(z_t|y_{t-1}^*)$ and $y_{t|t-1} := E(y_t|y_{t-1}^*)$.
- $\Delta y_t := y_t - y_{t|t-1}$ and $\Delta z_t := z_t - z_{t|t-1}$.
- $\Sigma_{t|t} := \mathrm{Cov}(\Delta z_t|y_t^*)$.

The standard KF Procedure [1] has two main hypothesis:

- **H1:**
  - The noise perturbations must be white noise:
  $$E(\xi_t) = E(\zeta_t) = 0, \quad t = 0, 1, \ldots$$
  .
  - Given finite second moments:
  $$E(\xi_t \xi_t^T) = Q_t, \quad E(\zeta_t \zeta_t^T) = R_t$$
  .
  - The perturbations are uncorrelated:
  $$E(\xi_t \xi_s^T) = E(\zeta_t \zeta_s^T) = 0, \quad \forall s \neq t$$
  $$E(\xi_t \zeta_s^T) = 0, \quad \forall t, s = 0, 1, \ldots$$
  .
  - The states are uncorrelated:
  $$E(z_t z_s^T) = 0, \quad \forall t \neq s$$
- **H2:**
  - Normality of the perturbations:
  $$\xi_t \sim \mathcal{N}(0, Q_t), \quad \zeta_t \sim \mathcal{N}(0, R_t), \quad \forall t = 0, 1, \ldots$$

The KF procedure is now presented:

1) **Initialization:**

$$z_{0|0} = z_0 \in \mathbb{R}^n,\ \Sigma_{0|0} = 0$$

2) **Prediction:**

$$\begin{aligned} z_{t|t-1} &= F_t z_{t-1|t-1} \\ \Sigma_{t|t-1} &= F_{t-1}\Sigma_{t-1|t-1}F_{t-1}^T + Q_{t-1} \end{aligned}$$

3) **Correction:**

$$\begin{aligned} z_{t|t} &= z_{t|t-1} + M_t^{\text{opt}}\Delta y_t \\ \Sigma_{t|t} &= \Sigma_{t|t-1} - M_t^{\text{opt}} H_t \Sigma_{t|t-1} \end{aligned}$$

where $M_t^{\text{opt}} = \Sigma_{t|t-1}H_t^T\left(H_t\Sigma_{t|t-1}H_t^T + R_t\right)^{-1}$.

D. Extended KF

The main idea behind the extension is to apply the standard KF in nonlinear systems. Let us consider a system in the form

$$\begin{aligned} x_{t+1} &= f(x_t) + \xi_t \\ y_t &= H_t x_t + \zeta_t \\ x_0 &\in \mathbb{R}^n,\ t = 0, 1, \ldots \end{aligned} \tag{3}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ is a given vector function. In order to apply the KF, we wish to linearize the function $f(\cdot)$ around a reference point x_t^{ref}, following the common linearization scheme using the jacobian of $f(\cdot)$.

Let

$$\tilde{F}_t := \left.\frac{\partial f(x)}{\partial x}\right|_{x=x_t^{\text{ref}}}$$

with $\tilde{F}_t$ model (2) can be constructed and applied to the KF.

III. Instrumental Extended KF

In this work we will use a simple ARMA model, with correlated state and one unknown stationary parameter a. The model in consideration is

$$\begin{aligned} z_{t+1} &= a z_t + \xi_t + d\xi_{t-1} \\ y_t &= h z_t + \gamma_t \\ z_0 &\in \mathbb{R},\ t = 0, 1, \ldots \end{aligned} \tag{4}$$

and the objective is find an estimator $\tilde{a}$ for the unknown parameter a using the data vector y_t^*.

First, we assume that $\tilde{a}$ is not a constant, even when the real paramater a is stationary. This assumption takes into consideration that the data vector can be constantly changing and the estimation will change depending on each observation. Another important assumption is that this estimator $\tilde{a}_t$ has small linear changes, hence, this estimator takes the form

$$\tilde{a}_t = b\tilde{a}_{t-1} + \chi_{t-1} \tag{5}$$

with $0 < b < 1$ and $\chi_t \sim N(0, \sigma)$. Given the dynamics for $\tilde{a}_t$, we propose the complete dynamics for the ARMA model as

$$\begin{aligned} z_{t+1} &= \tilde{a}_t z_t + \xi_t + d\xi_{t-1} \\ \tilde{a}_{t+1} &= b\tilde{a}_t + \chi_t \\ z_0 &\in \mathbb{R},\ \tilde{a}_0 \in \mathbb{R},\ t = 0, 1, \ldots \end{aligned} \tag{6}$$

Note that the dynamics have a nonlinear, particularly bilinear, term in the first equation. Let $x_t := (z_t,\ \tilde{a}_t)^T$ and let $x_t^{\text{ref}} := (z_t^{\text{ref}},\ \tilde{a}_t^{\text{ref}})^T$ be a given reference point, the linearization yields the matrix

$$\tilde{F}_t = \left.\begin{bmatrix} \tilde{a} & z \\ 0 & b \end{bmatrix}\right|_{x=x_t^{\text{ref}}}$$

Initially, the standard KF scheme could be applied, but one of the assumptions that is used to develop the KF procedure is that the state z_t cannot be correlated with $z_{t-s}\ \forall s$; clearly, z_t and z_{t-1} are correlated due to the delay in ξ_t therefore, the state x_t is correlated with x_{t-1}.

Taking into consideration these facts, the standard procedure cannot assure a “adequate” estimation of the state z_t. In order to address this problem, we suggeest an implementation with instrumental variables.

Let $\tilde{x}_t := x_{2t}$, these transformation assures that the new state $\tilde{x}_t$ is not correlated with $\tilde{x}_{t-1}$; evidently, this transformation implies a new sampling in the available data, namely, we need to define a new data vector

$$\tilde{y}_t^* := \begin{cases} (y_0^T, y_2^T, \ldots, y_t^T)^T, & t \text{ even} \\ (y_0^T, y_2^T, \ldots, y_{t-1}^T)^T, & t \text{ odd} \end{cases}$$

Now, given the instrumental variables transformation and the linearization procedure for model (6), the standard KF procedure can be applied.

IV. Discussion and Future Work

Given the proposed extension of the standard KF for nonlinear ARMA models, it can be easily extended for higher order noise delays (making larger steps in the sampling of the available data) or more complex structures for the parameters (matrices for example).

The suggested procedure is designed to satisfy the hypothesis for the standard KF, but it pays with the usage of available data (take half or less available data); on the other hand, a direct application of the KF cannot assure an adequate estimation of the state and we enter a situation where the two approaches have advantages and disadvantages, and it is believed that the two approaches performance depend on the concrete problem to solve.

As future work, first, we consider important to develop a computational implementation of the proposed scheme, compare it with a direct application of the standard KF (even when **H1** is violated) and compare both approaches with some known time-series data.

On the other hand, the estimation of parameter a needs to be contrasted with an standard application of classical method, namely, a least-square estimation and analyze the effectiveness of each approach in different scenarios.

References

[1] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.