A Prediction Error Adaptive Kalman filter for on-line spectral measurement correction and concentration estimation

Daniel Sbarbaroa\*, Tor Arne Johansenb, Jorge Yañezc

aDepartement of Electrical Engineering, Universidad de Concepción, Concepción, Chile

bDepartment of Engineering Cybernetics, Norwegian University of Science and Technology, Trondheim, Norway

cDepartment of Analytical and Inorganic Chemistry, Faculty of Chemical Sciences, Universidad de Concepcion, Concepcin, Chile

dsbarbar@udec.cl

Abstract

Spectral measurements offer real-time insights into the composition and concentration of species in process samples. However, they are sensitive to external factors such as temperature, pressure, and particle size distribution, all of which have a significant impact on the precision of spectroscopic measurements. In this study, we introduce an integrated discrete-time nonlinear model considering the dynamic aspects of the process, alongside a physics-based sensor model. Additionally, we propose an innovative application of two alternative Prediction Error Adaptive Kalman Filters to estimate both concentrations and sensor model parameters. The simulation of a simple ternary mixing process enables a comparison of the key characteristics of the proposed adaptive Kalman filters with a standard Extended Kalman Filter. The simulation results show that both Prediction Error Kalman filters can estimate concentrations and sensor parameters with minimal error, even in the presence of temperature variations and measurement noise. However, the proposed filters offer advantages in terms of ease of tuning and convergence.

* 1. Introduction

Spectroscopic instrumentation plays an important role in the context of Process Analytical Technologies (PAT) since it provides vital information for monitoring and controlling industrial processes. It is widely used in various process industries and is capable of fast measurements with no sample preparation, making it ideal for on-line applications. The sensing principles are based on the interaction between light and matter, acquired by opto-electronic components. To estimate relevant information from spectroscopic measurements, it is necessary to model their dependencies. Data-driven models, based on a linear assumption, have been successfully applied to estimate concentrations of constituents in a mixture from spectral information. However, these models do not always properly represent nonlinear interactions, resulting in limited extrapolation capabilities. Kriesten et al. (2008) proposed the Indirect Hard Modeling (IHM) approach, a physics-based spectral method that incorporates the physicochemical information of the elements. Under this approach, pure components are parametrized as peak models, accounting for nonlinear effects such as peak shifts and deformations.

As spectroscopic signals can be influenced by the characteristics of the transmission media between the analyzed sample or changes in their properties, it is necessary to correct these effects for accurate concentration estimation. Several data-driven approaches can be employed to compensate for temperature in Near-Infrared spectra (Hageman et al., 2005). The main limitations of these approaches are their requirement for a set of informative data for calibration, which is performed offline. In this contribution, we propose extending the Indirect Hard Modeling (IHM) approach to achieve a compact and interpretable parametric representation of the temperature's effect on the mixture spectra.

Thus, the sensor model combined with the process model results in a nonlinear state-space representation. It is noteworthy that standard Kalman filters have been utilized in analytical chemistry for solving various problems, such as multicomponent curve resolution and concentration estimation, among others (Rutan, 1987). However, in this case, since the model's structure is nonlinear, standard Kalman filters cannot be applied. Extended Kalman Filters can be applied, but their convergence cannot be guaranteed. In a previous work (Sbarbaro et al., 2023), a nonlinear transformation was proposed to address the nonlinearity problem and use a standard Kalman filter. This approach leads to an over-parametrized model. The main drawback of this approach lies in the increased dimension of the problem, and tuning can be challenging as parameters can no longer be associated with the noise covariances.

In this study, we explore two alternative approaches based on prediction error methods, making use of a minimal set of variables. The first approach, referred to as the Dual EKF (Jakoby, 1987), involves the utilization of two EKFs and the total derivative of the prediction error. This enables the separate estimation of concentration and sensor parameters, effectively reducing computational time. The second approach introduces a Recursive Prediction Error EKF, utilizing a single EKF for concentration estimation, with parameter estimation achieved through the minimization of a predictive error metric (Riva et al., 2015). The paper organization is as follows: Section 2 describe the model, Section 3 presents the problem and the proposed Kalman filters, Section 4 illustrates and compares their performance with respect a standard EKF. Finally, in section 5, some conclusion and future works are outlined.

*Acknowledgement:* This work was supported by Fondecyt project 1221225.

* 1. Model description

This work considers the process of mixing several constituents in a main stream by using static mixer in a main stream, as depicted in figure 1.



Figure 1. Mixing process

* + 1. Process model

The process dynamic describing the concentration of each constituent in the main steam at the output of the static mixer can be represented by a discrete-time state space model

$x\left(k+1\right)=Ax\left(k\right)+Bu\left(k\right)+ω\left(k\right)$ (1)

$w\left(k\right)=Cx\left(k\right)$ (2)

where the vector $w \in R^{m}$ represents the concentration of each constituent,$ x \in R^{n}$ the state of the model representing the dynamic of the concentrations, $u\in R^{m}$ the pump’s speed and $ω \in R^{n}$ an input disturbance. The matrices $A\in R^{n×n} $, $B\in R^{n×m}$, and $C\in R^{m×n}$ define the dynamic characteristic of the model.

* + 1. Sensor model

In a mixture of $m $components, and based on the Lambert-Beer law, the mixture absorbance spectrum consists in a weighted sum of the constituent absorbance spectra. Since the sensor has a limited spectral resolution, it delivers a vector of values $y$ corresponding to the light intensity $y\left(k,λ\right)$ measured at $L$ equally spaced wavelength values of $λ$, therefore the spectral output can be modeled as follows:

$y\left(k\right)=d+\sum\_{i=1}^{m}w\_{i}(k)ϕ\_{i}\left(v(k)\right)+ξ\left(k\right)$ (3)

where $y(k) \in R^{L}$,$d\in R^{L}$,$ξ(k)\in R^{L}$, and $ϕ\_{i}\left(v(k)\right)\in R^{L}$. These vector represent samples values of the baseline offset $d\left(λ\right)$, the absorbance of each constituent $ϕ\_{i}\left(λ,v(k)\right)$ ; which it also depends of some external variable $v\left(k\right)$, and the sensor noise modeled by the random variable$ξ\left(k\right)$. According to the IHM approach, the absorbance of each constituent can be expressed in terms of peak shaped functions as follows:

$ϕ\_{i}\left(λ,v(k)\right)=\sum\_{j=1}^{N\_{i}}α\_{ij}(v(k))ψ\_{ij}(λ)$ (4)

 where $N\_{i}$ is the number of peak functions required to model constituent $i$, $ψ\_{ij}\left(λ\right)$ are peak functions. The effect of the exogenous variable, $v(k)$, over the constituent spectra is modeled as changes in the linear factors. These changes can be parametrized in terms of known functions and a set calibration parameters

$α\_{ij}(v(k))=\sum\_{q=1}^{N\_{α\_{ij}}}β\_{ijq}^{}φ\_{ijq}(v(k))$ (5)

where $φ\_{ijp}(v(k))$ are known functions, for instance polynomials. Collecting all the calibration parameters in a single vector $θ=[β\_{111}^{} β\_{112}^{} \cdots β\_{mN\_{m}N\_{α\_{mN\_{m}}}}^{}]^{T}\in R^{N\_{p}}$ with $N\_{p}=\sum\_{i=1}^{m}\sum\_{j=1}^{N\_{i}}N\_{α\_{ij}}$ the sensor equation can be written in vector notation

$y\left(k\right)=d+\sum\_{q=1}^{N\_{p}}θ\_{q}Ψ\_{q}\left(v(k)\right)w\left(k\right)+ξ\left(k\right)$ (6)

where the matrix $Ψ\_{q}\left(v(k)\right)\in R^{L×m}$ has the cross products between $φ\_{ijp}(v(k))$ and $ψ\_{ij}\left(λ\right)$.

* 1. The estimation problem and adaptive Kalman filters

The model describing both the process dynamics and the sensor is

$x\left(k+1\right)=Ax\left(k\right)+Bu\left(k\right)+ω\left(k\right)$ (7)

$w\left(k\right)=Cx\left(k\right)$ (8)

$y\left(k\right)=\sum\_{i=1}^{N\_{p}}θ\_{i}Ψ\_{i}(v\left(k\right))w\left(k\right)+ξ(k)$ (9)

To estimate the unknown calibration parameters, they are modelled as

$θ\left(k+1\right)=θ\left(k\right)+ω\_{θ}\left(k\right)$ (10)

where the zero-mean random variable $ω\_{θ} \in R^{N\_{p}}$ models the parameter temporal uncertainty. The estimation problem can be stated as follows: given the system equations; i.e. Eq. (7), Eq. (8), Eq. (9) and Eq. (10), and the measurements $y\left(k\right)$, $v\left(k\right)$, $u\left(k\right)$ estimates the concentrations $w\left(k\right) $and calibration parameters $θ\_{i}$.

* + 1. Dual Extended Kalman filter

The dual Extended Kalman filter considers two independent Kalman filters. One for the state and one for parameters as defined by the following equations:

|  |  |
| --- | --- |
| State update: | Parameter update: |
| $$\hat{x}^{-}\left(k\right)=A\hat{x}\left(k-1\right)+Bu\left(k-1\right)$$$P\_{x}^{-}\left(k\right)=AP\_{x}\left(k-1\right)A^{T}+Q\_{x}$$E\_{x}\left(k\right)=C\_{x}\left(k\right)P\_{x}^{-}\left(k\right)C\_{x}^{T}\left(k\right)+R$$$K\_{x}\left(k\right)=P\_{x}^{-}\left(k\right)C\_{x}^{T}\left(k\right)E\_{x}\left(k\right)^{-1}$$$\hat{x}\left(k\right)=\hat{x}^{-}\left(k\right)+K\_{x}(k)e\left(k\right)$ $P\_{x}\left(k\right)=(I\_{n}-K\_{x}(k)C\_{x}\left(k\right))P\_{x}^{-}\left(k\right)$  |  $\hat{θ}^{-}\left(k\right)=\hat{θ}\left(k-1\right)$$$P\_{θ}^{-}\left(k\right)=P\_{θ}\left(k-1\right)+Q\_{θ}$$$$E\_{θ}\left(k\right)=H\_{θ}\left(k\right)P\_{θ}^{-}\left(k\right)H\_{θ}^{T}\left(k\right)+R$$ $K\_{θ}\left(k\right)=P\_{θ}^{-}\left(k\right)H\_{θ}^{T}\left(k\right)E\_{θ}(k)^{-1}$$$\hat{θ}\left(k\right)=\hat{θ}^{-}\left(k\right)+K\_{θ}(k)e\left(k\right)$$$$P\_{θ}\left(k\right)=(I\_{N\_{p}}-K\_{θ}(k)H\_{θ}\left(k\right))P\_{θ}^{-}\left(k\right)$$ |

where the error $e\left(k\right)$ and the time varying matrices $C\_{x}\left(k\right)$ and $C\_{θ}\left(k\right)$ are calculated as:

$e\left(k\right)=y\left(k\right)-\sum\_{i=1}^{N\_{p}}\hat{θ}\_{i}Ψ\_{i}(v\left(k\right))\hat{w}\left(k\right)$, (11)

$C\_{x}\left(k\right)=\sum\_{i=1}^{N\_{p}}\hat{θ}\_{i}Ψ\_{i}(v\left(k\right))$, $C\_{θ}\left(k\right)=\left[Ψ\_{1}(v\left(k\right))\hat{w}(k)\cdots Ψ\_{N\_{p}}(v\left(k\right))\hat{w}(k)\right]$ (12)

The total differential of the output is calculated as follows

$H\_{θ}\left(k\right)=\left.\frac{∂\hat{y}}{∂θ}\right|\_{\hat{x}^{-}\left(k\right),\hat{θ}^{-}\left(k\right)}=C\_{θ}\left(k\right)+C\_{x}\left(k\right)\frac{∂\hat{x}^{-}\left(k\right)}{∂θ}$ (13)

where $\frac{∂\hat{x}^{-}\left(k\right)}{∂θ}=A\frac{∂\hat{x}\left(k-1\right)}{∂θ}$ and assuming that $K\_{x}(k-1)$ weakly depends on $θ$as suggested by Plett (2004),

$\frac{∂\hat{x}\left(k-1\right)}{∂θ}=\frac{∂\hat{x}^{-}\left(k-1\right)}{∂θ}-K\_{x}\left(k-1\right)C\_{θ}\left(k-1\right).$ (14)

The initialization step considers $\hat{θ}\left(0\right)= 0$ , $P\_{θ}\left(0\right)=γ\_{1}I$, $P\_{x}\left(0\right)=γ\_{2}I$ where $γ\_{1}$, $γ\_{2}$ are positive real constants, and $Q\_{θ}$,$Q\_{x}$,$R$represent the noise covariance matrices for the parameters, the states and the measurement respectively.

* + 1. Prediction Error Adaptive Kalman filter

The PE Adaptive EKF estimates parameters by minimizing a cost function of the prediction error. To this end the sensitivities of the predictions with respect to the parameter must be calculated. The state and the parameters updates are defined by the following equations:

|  |  |
| --- | --- |
| State update | Parameter update |
| $$\hat{x}^{-}\left(k\right)=A\hat{x}\left(k-1\right)+Bu\left(k-1\right)$$$P\_{x}^{-}\left(k\right)=AP\_{x}\left(k-1\right)A^{T}+Q\_{x}$$$E\_{x}\left(k\right)=C\_{x}\left(k\right)P\_{x}^{-}\left(k\right)C\_{x}^{T}\left(k\right)$$$K\_{x}\left(k\right)=P\_{x}^{-}\left(k\right)C\_{x}^{T}\left(k\right)E\_{x}\left(k\right)^{-1}$ $\hat{x}\left(k\right)=\hat{x}^{-}\left(k\right)+K\_{x}(k)e\left(k\right)$ $P\_{x}\left(k\right)=(I\_{n}-K\_{x}(k)C\_{x}\left(k\right))P\_{x}^{-}\left(k\right)$  | $$\hat{θ}\left(k\right)=\hat{θ}\left(k-1\right)+L\left(k\right)e(k)$$$$V\left(k\right)=λV\left(k-1\right)+(1-λ)( e\left(k\right)e\left(k\right)^{T})$$$$L\left(k\right)=O\left(k-1\right)S\_{\hat{y}}\left(k\right)^{T}\left(λV\left(k\right)+S\_{\hat{y}}\left(k\right)O\left(k-1\right)S\_{\hat{y}}\left(k\right)^{T}\right)^{-1}$$$$O\left(k\right)=\left(I\_{N\_{p}}-L\left(k\right)S\_{\hat{y}}\left(k\right)\right)O\left(k-1\right)\left(I\_{N\_{p}}-L\left(k\right)S\_{\hat{y}}\left(k\right)\right)$$ $+L\left(k\right)V\left(k\right)L\left(k\right)^{T}$ |

The error, $e\left(k\right)$, is calculated as Eq. (11). The initialization step requires the following initial conditions $\hat{θ}\left(0\right)= 0$ , $V\left(0\right)= 0$ and $P\_{x}\left(0\right)=γ\_{1}I$ where $γ\_{1}$ is a positive real constant and $Q\_{x}$,$R$matrices representing the noise covariance matrices for the states and the measurement respectively. The parameter $λ$ represents a forgetting factor and it is a positive real number. The output sensitivity $S\_{\hat{y}}\left(k\right)$ can be written as

$S\_{\hat{y}}\left(k\right)=C\_{x}\left(k\right)S\_{\hat{x}}\left(k\right)^{-}+C\_{θ}\left(k\right)$ (15)

where the sensitivity of the predicted state $\hat{x}^{-}\left(k\right) $and matrix $P\_{x}^{-}(k)$ are respectively

$S\_{\hat{x}}\left(k\right)^{-}=AS\_{\hat{x}}\left(k-1\right)$, $S\_{P}\left(k\right)^{-}=AS\_{P}\left(k-1\right)(I\_{N}\_{p}⊗P\_{x}\left(k-1\right)A^{T})$, (16)

where $⊗ $is the Kronecker product, andthe Kalman gain sensitivity

$S\_{K}\left(k\right)=S\_{P}\left(k\right)^{-}\left(I\_{N}\_{p}⊗C\_{x}\left(k\right)^{T}E\_{x}(k)^{-1}\right)+P\_{x}^{-}(k)S\_{C}\left(k\right)^{T}\left(I\_{N}\_{p}⊗E\_{x}(k)^{-1}\right)+P\_{x}^{-}(k)C\_{x}^{T}\left(k\right)\frac{∂E(k)^{-1}}{∂θ}$ (17)

$\frac{∂E(k)^{-1}}{∂θ}=-E(k)^{-1}\frac{∂E(k)^{}}{∂θ}\left(I\_{N}\_{p}⊗E(k)^{-1}\right)$ (18)

$\frac{∂E(k)^{}}{∂θ}=S\_{C}\left(k\right)^{T}\left(I\_{N}\_{p}⊗P\_{x}(k)C\_{x}\left(k\right)^{T}\right)+C\_{x}\left(k\right)S\_{P}\left(k\right)^{-}\left(I\_{N}\_{p}⊗C\_{x}\left(k\right)^{T}\right)+C\_{x}\left(k\right)P\_{x}^{-}(k)S\_{C}\left(k\right)^{T}$(19)

and finally, for the update step

$S\_{\hat{x}}\left(k\right)=S\_{\hat{x}}\left(k\right)^{-}+K\left(k\right)S\_{\hat{y}}\left(k\right)+S\_{K}\left(k\right)\left(I\_{N}\_{p}⊗e\left(k\right)\right)$ (20)

$S\_{P}\left(k\right)=\left(I\_{N}\_{p}-K\left(k\right)C\_{x}(k)^{T}\right)S\_{P}\left(k\right)^{-}-S\_{k}\left(k\right)\left(I\_{N}\_{p}⊗C\left(k\right)P\_{x}^{-}(k)\right)-K\left(k\right)S\_{c}\left(k\right)\left(I\_{N}\_{p}⊗P\_{x}^{-}(k)\right)$ (21)

* 1. Simulation results

The mixing of three components having temperature dependant absorbance illustrates the performance of the adaptive Kalman filters. A spectrometer measures the spectra at the output of the mixing stage, as depicted in Figure 1.

* + 1. Absorbance modelling

The model of the absorbance of each constituent considers a linear combination of two Gaussian peaks, where some of them are affected by the measured temperature,$ v\left(k\right)$, as described by the following equations:

$ϕ\_{1}\left(v(k)\right)=θ\_{1}ψ\_{11}+θ\_{2}v(k)ψ\_{12}$ $ϕ\_{2}\left(v(k)\right)=θ\_{1}v(k)ψ\_{21}+θ\_{2}ψ\_{22}$ $ϕ\_{3}\left(v(k)\right)=θ\_{1}ψ\_{31}+θ\_{2}v(k)ψ\_{32}$

thus, vectors $Ψ\_{q}$ in the model Eq. (6) are $Ψ\_{1}=\left[ψ\_{11} 0 0\right]$,$ Ψ\_{2}=\left[v(k)ψ\_{12} 0 0\right] $,$Ψ\_{3}=\left[0 v(k)ψ\_{21} 0\right]$,$ Ψ\_{4}=\left[0 ψ\_{22} 0\right]$,$ Ψ\_{5}=\left[0 0 ψ\_{31}\right]$, and $ Ψ\_{6}=\left[0 0 v(k)ψ\_{32}\right]$. The nominal values of the parameters are $θ=[ 2.2 0.09 0.08 1 1.3 0.06 ]$. The dynamic of the static mixer is characterized by a linear dynamic system. The sampling period is one second. Step changes in the flow-rates of each components and periodic variation in the temperature, as shown in Figure 2, are considered.



Figure 2 a) absorbance b) flowrates c) sensor output d) temperature

* + 1. Parameter and state estimation results

The initial parameters of all Kalman filters are $Q\_{x}=10^{-2}I$,$R=0.1I$, $P\_{x}=10^{-2}I$,and $λ$=0.98. The observer initial conditions for states and parameters are set to zero. The evolution of the variables for all the algorithms is depicted in Figure 3. EKF and DEKF exhibit similar behaviour, converging faster in the estimation of concentrations but much slower in the parameters compared to PEAKF. The smallest Average Execution Time (AET) is achieved by EKF followed by DEKF. However, their parameter estimation accuracy is not as good as that obtained by PEAKF, as shown in Table 1.

Table 1: Summary results

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Algorithm | $$I(x)$$ | $$I(θ)$$ | Estimated parameters$θ$ | AET (ms) |
| EKF | 2.9518 | 0.3594 | 2.2082 0.0918 0.0816 0.9934 1.2434 0.0627 | 0.4774 |
| DEKF | 5.7662 | 0.3598 | 2.2164 0.0924 0.0818 0.9962 1.2438 0.0623 | 0.7549 |
| PEAKF | 4.1621 | 0.1078 | 2.2040 0.0902 0.0799 0.9905 1.3105 0.0613 | 6.3156 |

$$ I(x)=\frac{1}{1000}\sum\_{i=0}^{1000}\left‖x\left(k\right)-\hat{x}\left(k\right)\right‖$$



EKF

DEKF

PEAKF

Figure 3 Evolution of parameters and concentrations

* 1. Conclusions

The application of Prediction Error Adaptive Kalman filters to estimate both concentrations and sensor model parameters using a discrete-time nonlinear model is investigated. Simulations of a simple ternary mixing process reveal that the proposed adaptive Kalman filters have relative merits compared to the standard EKF. In the simulated system, PEAKF demonstrates the fastest convergence and improved accuracy in terms of sensor parameters, albeit at the expense of increased calculations. DEKF results are comparable to those obtained by the EKF. Thus, the proposed adaptive EKFs offer advantages in terms of ease of tuning and convergence. Future work will involve a more in-depth analysis of DEKF without simplifying assumptions, a comparative analysis with Unscented Kalman Filters, and the robust and efficient implementation of the algorithms.

References

S.C. Rutan, 1987,S. C. Kalman Filtering Approaches for Solving Problems in Analytical Chemistry. J. Chemom. 1, 7–18.

Z. Chen Z, J. Morris, 2009, Process analytical technology and compensating for nonlinear effects in process spectroscopic data for improved process monitoring and control. Biotechnol J., 4(5), 610-9.

J. Hageman, J. Westerhuis, and A. Smilde, 2005, Temperature robust multivariate calibration: An overview of methods for dealing with temperature influences on near infrared spectra, J. Near Infrared Spectrosc., 13(2), 53–62.

W. Jakoby, M. Pandit, 1987, A prediction-error-method for recursive identification of nonlinear systems, Automatica, 23(4), 491-496.

M. H. Riva, D. Beckmann, M. Dagen and T. Ortmaier, 2015, Online parameter and process covariance estimation using adaptive EKF and SRCuKF approaches, IEEE Conference on Control Applications (CCA), Sydney, Australia, 1203-1210, doi: 10.1109/CCA.2015.7320776.

D. Sbarbaro, T. A. Johansen and J. Yañez, 2023, Adaptive Kalman Filter for On-Line Spectroscopic Sensor Corrections, 9th International Conference on Control, Decision and Information Technologies, Rome, Italy, 2109-2114, doi: 10.1109/CoDIT58514.2023.10284429.

G. L. Plett, 2004, Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 3. State and parameter estimation, Journal of Power Sources, 134(2), 277-292.