Antonio Espuña, Moisès Graells and Luis Puigjaner (Editors), Proceedings of the 27th European Symposium on Computer Aided Process Engineering – ESCAPE 27
October 1st - 5th, 2017, Barcelona, Spain
http://dx.doi.org/10.1016/B978-0-444-63965-3.50248-8

Unscented Kalman Filter. Application of the robust approach to polymerization processes

Jhovany Tupaz[b], Mariano Asteasuain[a,b], Mabel Sánchez[a,b*]

[a]*DIQ, Universidad Nacional del Sur (UNS), Av. Alem 1253, Bahía Blanca 8000, Argentina*
[b]*PLAPIQUI, Universidad Nacional del Sur-CONICET, Camino La Carrindanga km 7, Bahía Blanca 8000, Argentina*
msanchez@plapiqui.edu.ar

Abstract

The control of polymerization processes has central importance because operational conditions affect the processing and end-use properties of the product. The nonlinear controllers based upon rigorous models make use of the on-line state estimates obtained from the available measurements. For polymerization processes, the Unscented Kalman Filter has shown a rewarding performance for state estimation. Because the presence of outliers distorts the behaviour of the filter, Robust Statistics-based approaches have been proposed to reduce their detrimental effect on variable estimates. Until now, only Huber type M-estimators have been used as loss function of the estimation problem. In this work, the ability of other types of M-estimators to improve estimate robustness without introducing numerical problems is analysed. The performances of the M-estimators are compared for a copolymerization process within the framework of a filtering technique based on the Unscented Transformation, which uses a reformulation of the covariance of measurements errors.

Keywords: Outliers, Robust Filtering, M-estimators, Copolymerization, Unscented Kalman Filter

1. Introduction

The dynamic operation of polymerization processes is an extremely complex task. They usually exhibit highly exothermic reactions and changes in viscosity which cause complex heat-transfer dynamics and flow patterns. Moreover, those processes are strongly nonlinear.

The problem of state estimation in nonlinear systems has been covered extensively in the past. A widespread estimation technique in process control is the Extended Kalman Filter, but this strategy may present linearization errors when the nonlinearity is strong. The Unscented Kalman Filter (UKF) has been developed for this type of processes. It applies the Unscented Transformation (Julier and Uhlmann, 2004), which is a way to calculate the statistics of a random variable when it undergoes a nonlinear transformation. The state estimation of polymerization processes has been successfully addressed using the UKF (Galdeano et al., 2011).

Because the presence of outliers distorts the quality of state estimates, nonlinear filtering techniques based on Robust Statistics are devised to obtain reliable estimates

within the framework of the UKF. All of them use the Huber function as an M-estimator. Some of them apply a linearized approximation of the nonlinear measurement function (Wang et al., 2010), but the linearized method bears many drawbacks, such as, low accuracy, cumbersome derivation and evaluation of the Jacobian matrices, etc. Others strategies use the nonlinear measurement function directly and a reformulation of the covariance of the measurement errors (Chan et al., 2012).

In this work, the ability of redescending M-estimators to be used as loss functions for the derivative-free UKF algorithm is analyzed. The study is carried out for the state estimation of the methyl methacrylate (MMA) and vinyl acetate (VA) copolymerization process. Results show the disadvantages of using the Huber M-estimator, and the effect of parameter settings on the efficiency and robustness of the state estimates.

2. Robust Unscented Kalman Filtering

Let us consider that a dynamic system can be represented by the following state space model

$$x_k = f(x_{k-1}) + w_{k-1} \tag{1}$$

$$y_k = h(x_k) + v_k \tag{2}$$

where x_k is the n-dimensional vector of states at time step k, y_k is the m-dimensional measurement vector, f(.) is the process model function, h(.) is the measurement model function, w_{k-1} is the process noise caused by disturbances and modelling errors, and v_k is the measurement noise.

To make a robust inference about the state vector given the model and the measurements obtained until time k ($\hat{x}_{k|k}$), the following nonlinear regression model is formulated at first:

$$\begin{bmatrix} y_k \\ \hat{x}_{k|k-1} \end{bmatrix} = \begin{bmatrix} h(x_k) \\ x_k \end{bmatrix} + \begin{bmatrix} v_k \\ e_{k|k-1} \end{bmatrix} \tag{3}$$

It combines Eq.(2) and the relationship between the true state vector x_k and its prediction taking into account the observations got until time k-1, $\hat{x}_{k|k-1}$. The residual covariance of the nonlinear regression model is

$$M_k = E\left\{ \begin{bmatrix} v_k \\ e_{k|k-1} \end{bmatrix} \begin{bmatrix} v_k \\ e_{k|k-1} \end{bmatrix}^T \right\} = \begin{bmatrix} R_k & 0 \\ 0 & P_{k|k-1} \end{bmatrix} \tag{4}$$

where R_k and $P_{k|k-1}$ are the covariance of the measurements and estimate prediction errors, respectively. The result of the Cholesky factorization of M_k, S_{Mk}, is used to decorrelate the nonlinear regression model by multiplying Eq.(3) by S_{Mk}^{-1}. The final model equation is

$$z_k = g(x_k) + r_k \tag{5}$$

where

$$z_k = S_{M_k}^{-1}\begin{bmatrix} y_k \\ \hat{x}_{k|k-1} \end{bmatrix}, \quad g(x_k) = S_{M_k}^{-1}\begin{bmatrix} h(x_k) \\ x_k \end{bmatrix}, \quad r_k = S_{M_k}^{-1}\begin{bmatrix} v_k \\ e_{k|k-1} \end{bmatrix} \tag{6}$$

An efficient and robust solution of the regression problem can be achieved solving the following optimization problem

$$\hat{x}_k = \arg\min_{x_k} \sum_{i=1}^{m+n} \rho\left(r_k^{(i)}\right) \tag{7}$$

where ρ is the Loss Function (LF) of an M-estimator and $r_k^{(i)}$ is the i-th component of the residual. If $\rho(r_k^{(i)})$ is differentiable, its derivative is the Influence Function (IF), ψ, i.e. ρ'(r) = ψ(r). Furthermore, the weight function w is related to the IF as follows

$$w(r) = \begin{cases} \psi(r) \big/ r & if\ r \neq 0 \\ \psi'(0) & if\ r = 0 \end{cases} \tag{8}$$

The weight matrix W_k is defined using the weights $w(r_k^{(i)})$ as diagonal elements

$$W_k = \begin{bmatrix} w\left(r_k^{(1)}\right) & & 0 \\ & \ddots & \\ 0 & & w\left(r_k^{(m+n)}\right) \end{bmatrix} \tag{9}$$

The atypical measurements receive smaller weights. That matrix is used to reformulate M_k (Chang et al., 2012) as follows

$$\widetilde{M}_k = S_{M_k} W_k^{-1} S_{M_k}^T = \begin{bmatrix} \widetilde{R}_k & 0 \\ 0 & \widetilde{P}_{k|k-1} \end{bmatrix} \tag{10}$$

By considering that $e_{k|k-1}=0$, the reformulation only affects the covariance of the measurement errors, that is, the values of R_k in M_k are replaced by those corresponding to $\tilde{R}_k$ in $\widetilde{M}_k$. Matrix $\widetilde{M}_k$ is applied in the measurement update step of the standard UKF. For the sake of brevity, no more details about the UKF algorithm are provided in this work, but they can be found elsewhere (Julier and Uhlmann, 2004).

3. Robust M-Estimators

Robust Statistics aims at providing reliable estimates not only when the data follow a given distribution exactly, but also when this happens only approximately. Therefore there is a trade-off between efficiency and robustness.

Let us assume that the distribution of an estimate $\hat{\mu}$ is approximately $N(\mu_0, v/N)$ when the number of samples N increases, then $\hat{\mu}$ is asymptotically normal with asymptotic

value μ_0 and asymptotic variance v. The asymptotic efficiency of $\hat{\mu}$ is defined as the ratio v_0/v, where v_0 is the asymptotic variance of the optimal solution of Eq.(7) if the LF is the Least Square estimator, and measures how near $\hat{\mu}$ is to the optimum (Maronna et al., 2006). To ensure a high asymptotic efficiency at the normal distribution, for example 95 %, the parameters of the M-estimator functions can be properly tuning using the jackknife procedure (Llanos et al., 2015).

Up to the present time, only the Huber (Hub) M-estimator has been used as LF for all robust UKF algorithms reported in the literature. The Huber M-estimator is a convex function and therefore unbounded. It has the advantage that the solution of Eq.(7) has a unique local minimum. In consequence, the starting values of the iterative solution procedure may influence the number of iterations but not the final outcome. However, Huber M-estimator is sensitive to very large outliers, and therefore may have a low efficiency for heavy-tailed error distributions.

In contrast, redescending estimators can be made very efficient for heavy-tailed data, but require a good starting point to ensure attaining the "good" solution. The LF of redescending estimators can be of two types. The first one is unbounded but its IF tends to zero at infinity. Therefore, the matrix W_k can be inverted. The second type of LF is bounded but its $\psi(r)$ is strictly equal to zero for |r|>parameter. In this case, the reformulation of the covariance of the measurement errors cannot be used because numerical problems arise. The Welsch (Wel) and Correntropy (Cor) LFs have the required features to apply the aforementioned technique. The IF of the first one asymptotically approaches zero for large values of |r|. The second estimator is the Gaussian kernel function, whose parameter C_{Cor} is the kernel width, and its IF tends quickly to zero for large values of $|r|>C_{Cor}$.

The LFs and their parameters for an asymptotic efficiency equal to 95 % are presented next for the Hub, Wel and Cor M-estimators.

$$\rho_{Hub}(r)=\begin{cases} r^2 & if\ |r|\le c_{Hub} \\ 2c_{Hub}|r|-c_{Hub}^2 & if\ |r|>c_{Hub} \end{cases}, \qquad c_{Hub}=1.40 \tag{11}$$

$$\rho_{Wel}(r)=c_{Wel}^2\left\{1-\exp\left[-\left(r/c_{Wel}\right)^2\right]\right\}, \qquad c_{Wel}=2.98 \tag{12}$$

$$\rho_{Cor}(r)=\left(1/c_{Cor}\sqrt{2\pi}\right)\exp\left[-\left(r^2/2c_{Cor}^2\right)\right], \qquad c_{Cor}=2.05 \tag{13}$$

4. Case study

The process selected as case study is the copolymerization of MMA and VA in a continuous stirred tank reactor with a recycle loop. The fresh stream contains the MMA and VA monomers, azoisobutyronitrile (AIBN) as initiator, benzene (B) as solvent, acetaldehyde as chain transfer agent (CTA). Also some inhibitor (INH) could be present.

The mathematical model of the process comprises a set of differential algebraic equations for the mass and energy balances of the system. A detailed description of the process and its mathematical model can be found elsewhere (Congalidis et al., 1989). The model involves the following state vector:

$$\left[C_{MMA}, C_{VA}, C_{AIBN}, C_B, C_{CTA}, \lambda^*_{MMA}, \lambda^*_{VA}, \gamma_0, \gamma_1, \gamma_2, T_r\right] \tag{14}$$

where C is a molar concentration, λ^* is a molar concentration of the monomer in the copolymer, γ_j the jth-order moment of the copolymer molecular weight distribution, and T_r is the reactor temperature.

Critical process and quality variables, such as the total conversion ($Conv_{Total}$) and the weight-average molecular weight ($\bar{M}_W$) can be expressed in terms of some the states in the state vector. The present work assumes that variables T_r, $Conv_{Total}$ and $\bar{M}_W$ are the on-line measurements available from the process. The first measurement could be obtained by a thermocouple; measurements of conversion and $\bar{M}_W$ could be obtained from on-line sensors that include empirical correlations. (Hashemi et al., 2013). Hence, the measurement vector y_k is:

$$y_k = \left[T_r \quad Conv_{Total} \quad \bar{M}_W\right]^T_k \tag{15}$$

5. Results and Discussion

Ten thousand simulations of the dynamic behaviour of the process were performed. At a certain time, a perturbation occurs in the reactor feed stream leading to a transition from an initial steady state to a final one. The Mean Square Error (MSE) of each measurement is calculated as the average of the differences between its true value and the estimated one using the filtering technique. Table 1 displays the results for the different methodologies when outliers are present and when they are not.

Table 1. Mean Square Errors for different M-estimators

Methodology	MSE without Outliers			MSE with Outliers		
	T_r	$Conv_{Total}$	$\bar{M}_W$	T_r	$Conv_{Total}$	$\bar{M}_W$
CUKF	1.08E-06	4.60E-01	3.75E+04	4.94E-06	1.91E+00	1.83E+05
UKF-Hub	1.43E-06	5.62E-01	5.22E+04	1.51E-06	6.07E-01	5.30E+04
UKF-Wel	1.28E-06	5.11E-01	4.66E+04	1.30E-06	5.25E-01	4.73E+04
UKF-Cor1	1.17E-06	4.87E-01	4.02E+04	1.21E-06	4.92E-01	4.18E+04
UKF-Cor2	1.38 E-06	5.50E-01	5.04E+04	1.44E-06	5.78E-01	5.12E+04

Results are provided for the Classic UKF (CUKF), and the robust derivative-free algorithms that use as M-estimators the Hub (UKF-Hub), Wel (UKF-Wel) and Cor function. Two different approaches are applied to calculate the parameter C_{Cor}. For UKF-Cor1, the jackknife procedure is used to satisfy an asymptotic efficiency equal to 95 %. In contrast, the methodology UKF-Cor2 applies an adaptive technique based on

the residual values to update the parameter value, and assumes that any residual equal to or greater than $2\sqrt{2}$ contributes very little to the estimates (Muñoz et al., 2012).

The results show that when no outliers are present, UKF-Cor1 behaves well in comparison to CUKF, which is the best estimator. The contrary happens for UKF-Hub and UKF-Cor2. Under the presence of outliers, the MSEs for the redescending M-estimators are lower than the one obtained for the Hub function because their IF tend to zero for very large outliers. Also, UKF-Cor1 outperforms UKF-Wel for state estimation, and the methodology to set the parameters used for UKF-Cor1 provides better results than the one used for UKF-Cor2.

6. Conclusion

In this work, the application of redescending M-estimators whose LF is unbounded but its IF tends to zero at infinity is analysed in the framework of derivative-free UKF algorithms. It is remarkable that better estimates are obtained for the copolymerization process in comparison with the ones provided by the current methodologies, which are based on the Hub function. Also no numerical problems arise using that specific type of redescending M-estimators. It is noticeable that the adaptive parameter setting of the Cor function does not yield a real improvement in performance.

References

J.P. Congalidis, J. Richards, and W.H. Ray, 1989, Feedforward and feedback control of a solution copolymerization reactor, AIChE Journal, 35, 6, 891-907.

L. Chang, B. Hu, G. Chang, and A. Li, 2012, Multiple Outliers Suppression Derivative-Free Filter Based on Unscented Transformation, Journal of Guidance, Control, and Dynamics, 35, (5), 1902-1907.

R. Galdeano, M. Asteasuain, and M.C. Sánchez, 2011, Nonlinear State Estimation with Delayed Measurements. Application to Polymer Processes, 21st European Symposium on Computer Aided Process Engineering, 29, 592-596.

R. Hashemi, D. Kohlmann, S. Engell, 2013, Optimizing Control of a Continuous Polymerization Reactor, 10th IFAC International Symposium on Dynamics and Control of Process Systems, 46, (32), 726-731.

J. Julier, J.K. Uhlmann, 2004, Unscented filtering and nonlinear estimation, Proceedings of the IEEE, 92, 3, 401-422.

C.E. Llanos, M.C. Sánchez, and R.A. Maronna, 2015, Robust Estimators for Data Reconciliation, Industrial and Engineering Chemistry Research, 54, (18), 5096-5105.

R.A. Maronna, R.D. Martin, and V. Yohai, 2006, Robust Statistic: Theory and Methods, John Wiley and Sons Ltd., England.

J.C. Munoz, J.Chen, 2012, Removal of the effects of outliers in batch process data throug maximum correntropy estimator, Chemometrics and Intelligent Laboratory Systems, 111, (1), 53-58.

D. Martinez Prata, M. Schwaab, E.L. Lima, and J.C. Pinto, 2013, Simultaneous Robust Data Reconciliation and Gross Error Detection through Particle Swarm Optimization for an Industrial Polypropylene Reactor, Chemical Engineering Science, 65, (17),4943-4954.

X. Wang, N.Cui and J.Guo, 2010, Huber-based unscented filtering and its application to vision-based relative navigation, IET Radar, Sonar and Navigation, 4, (1), 134-141.