I.A. Karimi and Rajagopalan Srinivasan (Editors), Proceedings of the 11th International Symposium on Process Systems Engineering, 15-19 July 2012, Singapore

Estimation of Primary Variables from Combination of Secondary Measurements: Comparison of Alternative Methods for Monitoring and Control

Maryam Ghadrdan,a Chriss Grimholt,a Sigurd Skogestad,a * Ivar J. Halvorsen,b

a Department of Chemical Engineering; NTNU; Trondheim, Norway
b SINTEF ICT, Trondheim, Norway

Abstract

In this work, we propose a method to estimate the primary variables based on a combination of measurements. Our method is a reformulation of the Loss method. We will compare our method with the other well-known estimators.

Keywords: Static estimator, Loss method, Partial Least Squares, Kalman Filter

1. Introduction

It happens frequently in process control that some important variables are not measurable. Sometimes they are expensive to measure and include delay. The value of primary variable can be inferred by using some secondary variable measurements. The task of soft sensors is the maximal exploitation of transforming the information of secondary measurements into more useful process knowledge.

In multi-component distillation system, a certain temperature profile exists for a specified feed component recovery and feed condition. The idea is to use temperature measurements from several locations in order to estimate the recovery of product composition, which is also good for some variation in feed properties. In this work, we propose an approach for designing a static estimator which is inspired from the loss method by Skogestad (2000). This work is a continuation of the work done by Hori et al. (2005). Measurement noise is included. The optimal static estimator is designed for two categories: "open-loop" performance (estimator used for monitoring) and "closed-loop" performance (estimator used for control). It is optimal in the sense that it gives the smallest prediction error which is defined as the difference between the true value and the estimated value. This approach will be compared with the Partial Least Square (PLS) approach and steady-state kalman filter.

2. Loss method for estimation

The objective is to find a linear combination of measurements such that keeping these constant indirectly leads to nearly accurate estimation with a small loss $\mathbf{L}$ in spite of unknown disturbances, d, and measurement noise, n^x. Figures 1-4 show four different scenarios we have considered. Linear models are assumed for the primary variables y,

*skoge@nt.ntnu.no

measurements x, and secondary variables z. It is also assumed that $\dim(\mathrm{y}) = \dim(\mathrm{x}) = \dim(\mathrm{z})$.

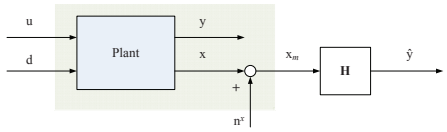

Figure 1: u is a free variable

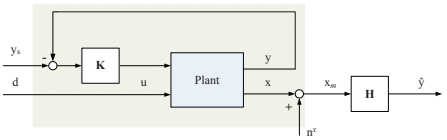

Figure 2: Control of primary variable y

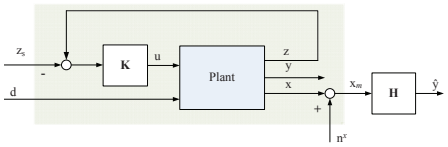

Figure 3: Control of secondary variable z

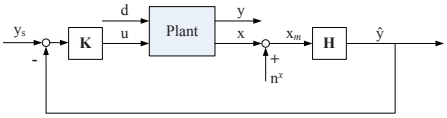

Figure 4: Control of the predicted variable $\hat{y}$

The optimal $\mathbf{H}$ is a matrix which follows the linear relationship $\hat{y} = \mathbf{H}x_m$, and derived by minimizing $\|e(\mathbf{H})\|_{2,exp}$. We assume some expected values for disturbances and noise which come from engineering wisdom. The normalized values of disturbances and noise are used in the calculations. Table 1 shows the final expressions for obtaining $\mathbf{H}$ for different scenarios. The problem is easy for "open-loop" cases. The final expression looks like ordinary least square problems. For "close-loop" case (Figure 4), the Frobenius norm should be minimized subject to $\mathbf{H}\mathbf{G}_x = \mathbf{G}_y$ as constraint. We need to obtain the optimal sensitivity matrix $\mathbf{F}$ which is defined as $\mathbf{F} = \left(\frac{dx_{opt}}{d\mathrm{d}}\right) = \mathbf{G}_x^d - \mathbf{G}_x\mathbf{G}_y^{-1}\mathbf{G}_y^d$, is simply obtained numerically by re-optimizing the model for different disturbances.

Table 1: Optimal $\mathbf{H}$ values for different scenarios

$\mathbf{H}_1 = \mathbf{Y}_1\mathbf{X}_1^\dagger$	$\mathbf{Y}_1 = \begin{bmatrix} \mathbf{G}_y\mathbf{W}_u & \mathbf{G}_y^d\mathbf{W}_d & 0 \end{bmatrix}$ $\mathbf{X}_1 = \begin{bmatrix} \mathbf{G}_x\mathbf{W}_u & \mathbf{G}_x^d\mathbf{W}_d & \mathbf{W}_{n^x} \end{bmatrix}$
$\mathbf{H}_2 = \mathbf{Y}_2\mathbf{X}_2^\dagger$	$\mathbf{Y}_2 = \begin{bmatrix} \mathbf{W}_{y_s} & 0 & 0 \end{bmatrix}$ $\mathbf{X}_2 = \begin{bmatrix} \mathbf{G}_x^{cl}\mathbf{W}_{y_s} & \mathbf{F}\mathbf{W}_d & \mathbf{W}_{n^x} \end{bmatrix}$
$\mathbf{H}_3 = \mathbf{Y}_3\mathbf{X}_3^\dagger$	$\mathbf{Y}_3 = \begin{bmatrix} \mathbf{G}_y^{cl}\mathbf{W}_{z_s} & \mathbf{F}'_y\mathbf{W}_d & 0 \end{bmatrix}$ $\mathbf{X}_3 = \begin{bmatrix} \mathbf{G}_x^{cl}\mathbf{W}_{z_s} & \mathbf{F}'_x\mathbf{W}_d & \mathbf{W}_{n^x} \end{bmatrix}$
$\mathbf{H}_4$	$\min_{\mathbf{H}} \left\|\mathbf{H}\begin{bmatrix} \mathbf{F}\mathbf{W}_d & \mathbf{W}_{n^x} \end{bmatrix}\right\|_F$ s.t. $\mathbf{H}\mathbf{G}_x = \mathbf{G}_y$

Note that $\mathbf{W}$'s are the diagonal scaling matrices which contain the standard deviations of the elements. We can consider u to be any variable from the process. $\mathbf{G}_y$ and $\mathbf{G}_y^d$ become so trivial if we choose u = y ($\mathbf{G}_y = \mathbf{I}$ and $\mathbf{G}_y^d = 0$), and $\mathbf{G}_x^d$ will be equal to the sensitivity matrix $\mathbf{F}$.

3. Partial Least Square (PLS) Method

In chemometrics, Partial Least Squares (PLS) regression has become an established tool for modeling linear relations between multivariate measurements. This method is used to compress the predictor data matrix X, into a set of latent variable or factor scores. The orthogonal factor scores are used to fit a set of observations to dependent variables Y. The main attraction of the method is that it finds a parsimonious model even when the

predictors are highly collinear or linearly dependent. The final fitting equation will be

$$\mathbf{Y} = \mathbf{BX} + \mathbf{B_0} \tag{1}$$

with B and B_0 as optimization variables. B_0 is close to zero because of centering the data. The main drawback of this method is that there are several realization of the same method which do not lead to the same result for a specific problem (for PLS2 cases).

For data preparation we have two ways: If we have data, small directions in the measurement space should be deleted by SVD. We should ensure that all important directions are sufficiently exposed. We can also use exactly the same data that we get in loss method. $\mathbf{X}$ and $\mathbf{Y}$ in PLS method are the first and second row of $\mathbf{Y}_{all}$ matrix respectively.

$$\mathbf{Y}_{all} = \begin{bmatrix} \mathbf{Y} \\ \mathbf{X} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_y & 0 \\ \mathbf{G}_x & \mathbf{X}_{opt} \end{bmatrix} \tag{2}$$

where $\mathbf{X_{opt}} = \begin{bmatrix} \mathbf{FW}_d & \mathbf{W}_{n^x} \end{bmatrix}$.

We need to know the expected "optimal variation" in $\mathbf{X}$ as given by the matrix $\mathbf{X}_{opt}$. Here "optimal" means that y is constant (see the second column in $\mathbf{Y}_{all}$). In addition, we also need to obtain $\mathbf{G}_x$ and $\mathbf{G}_y$ from the data, which means that the data must contain "non-optimal" variations in u, and not only contain optimal data where $\mathrm{u} = \mathrm{u}_{opt}\,(\mathrm{d})$- see the first column in $\mathbf{Y}_{all}$. This is called Closed-Loop Regressor (CLR) (Skogestad et al., 2011). CLR suffers from the same weakness as LS, giving poor results for ill-conditioned matrices and underdetermined systems. Performing a principal component analysis on the $\mathbf{X}$ data will remove the weaker directions containing noise resulting in a well-conditioned matrix. Then, CLR can be applied to the data. We call this "truncated CLR".

4. Kalman filtering

The Kalman filter estimates process states by using a form of feedback control. The linearity of state dynamics and observation process, as well as the normal distribution of noise in state dynamics and measurements are the assumptions of kalman filter. A linear difference equation $x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1}$ with a measurement that is $z_k = \mathbf{C}x_k + v_k$, define the linearized process. The random variables w_k and v_k represent the process and measurement noise respectively. They are assumed to be independent of each other and with normal distributions.

$$p\,(w) \sim \mathcal{N}\,(0, Q) \tag{3}$$

$$p\,(v) \sim \mathcal{N}\,(0, R) \tag{4}$$

The objective is to minimize the estimation error. By writing a posteriori state estimate as a linear combination of an a priori estimate and the difference between actual measurement and measurement prediction weighted by kalman gain, K is calculated to minimize the a posteriori estimation error covariance. Since the focus of our work is on chemical processes, the time scales at which the sensor noise characteristics change are much larger than the time scale at which we study the system. Thus we assume the system and noise covariances are time-invariant. In addition, as mentioned previously, our proposed estimator is categorized as static estimator. So, the steady-state of kalman filter is interesting. The steady-state kalman gain is calculated as $K_\infty = P_\infty^- H^T \left(HP_\infty^- H^T + R\right)$. Figure 5 shows the block diagram of kalman filter estimation.

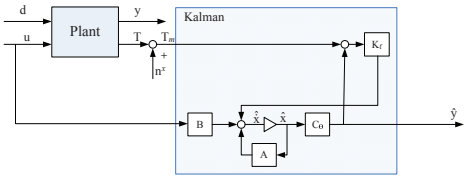

Figure 5: Block diagram of Kalman filter

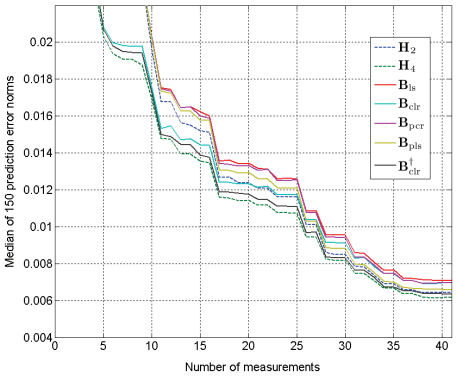

Figure 6: Median Prediction error

The algorithm of Kalman filter requires knowledge of the process noise variance W and the measurement noise variance V (Nakamura, 1982). If state-feedback control is used, the overall controller is optimal because of the separation principle. If an output-feedback controller ($u = \mathbf{K}y$) is used, then it is generally not optimal to use the $\hat{y} = \mathbf{C}\hat{x}$ estimated by Kalman filter. In loss method, the primary variable comes directly from combination of measurements ($c = \mathbf{H}y$). Since the measurements do not contribute similarly in the estimation of primary variable, it is expected that by putting weights on the state error terms, the estimation of primary variable will be improved. Another point is that we should think of what to use the estimation for. KF is said to be used for control,but if R approaches infinity, then it means that there is no control. Mejdell and Skogestad (1993) have shown that the kalman-filter might be better than a simple PCR in open-loop performance, which is because of the recursive nature of the filter, but PCR performs similarly if it is used for closed loop. We can model slowly-varying disturbances by adding states of the noise model. This gives the augmented kalman filter (Brown and C., 1997). Here, we use non-stationary noise. So zero steady-state is not reached.

5. Example

A binary distillation column model - Column A (Skogestad, 1997) - is used to demonstrate the performance of different estimators. There are two inputs, namely the reflux flow and the boilup, and two disturbances, which is the change in feed composition.

Table 2 shows the results of validation for estimators for different scenarios. Calibrating with one scenario and validating with another is mostly applicable to the last scenario. So, the shaded cells are actually showing the more interesting data. As it was expected,

the optimal estimator is obtained when it was calibrated and validated for its intended scenario. Note that all the scenarios are not comparable to each other because of different variances for different scenarios. Since there is no control in the first scenario, a small standard deviation in u was selected to give a small standard deviation in y.

Figure 6 shows the performance of estimators. Calibration data was generated by drawing random values for $\mathrm{u} \sim \mathcal{N}\left(0, 0.005^2\mathbf{I}_2\right)$, $\mathrm{d} \sim \mathcal{N}\left(0, 0.05^2\mathbf{I}_2\right)$, $\mathrm{y}_s \sim \mathcal{N}\left(0, 0.005^2\mathbf{I}_2\right)$ and $z_s \sim \mathcal{N}\left(0, \begin{bmatrix} 0.05^2 & 0.5^2 \end{bmatrix}\mathbf{I}_2\right)$, and calculating the corresponding output variables x_m and y for the respective scenarios (except scenario 4). The median of the prediction error for 150 runs are used to assess the estimators' performances because noise and variation in input variables resulted in a distorted picture of estimator performance by outliers. By increasing the number of regressors, the error decreases. All estimators are trained on calibration data from scenario 2 and validated on scenario 4.

Table 2: The mean prediction error of the model-based estimators applied to four operation scenarios

Caliberation Data \ Validation Data	S1	S2	S3	S4
S1	0.0085	0.2749	0.0215	0.0506
S2	0.0591	0.0093	0.0104	0.0104
S3	0.0599	0.0166	0.0098	0.0132
S4	0.0099	0.0099	0.0099	0.0099

6. Conclusions

In this paper, we introduced a new static estimator. Four scenarios have been used to get the calibration data and were validated for the closed-loop scenario. Our emphasis was on the fact that we should be aware of what we want to use the estimator for.

References

Brown, R. G., C., H. P. Y., 1997. Introduction to Random Signals and Applied Kalman Filtering. John Wiley & Sons.

Hori, E. S., Skogestad, S., V., A., 2005. Perfect steady-state indirect control. Ind. Eng. Chem. Res. 44, 863–867.

Mejdell, T., Skogestad, S., 1993. Output estimation using multiple secondary measurements. AIChE J. 39 (10), 1641–1653.

Nakamura, M., 1982. Relationship between steady state kalman filter gain and noise variances. J. of Systems Science 13, 1153–1163.

Skogestad, S., 1997. Dynamics and control of distillation columns: A critical survey. J. Modeling, Identification and Control 18 (3), 177–217.

Skogestad, S., 2000. Plantwide control: the search for the self-optimizing control structure. J. Proc. Control 10, 487–507.

Skogestad, S., Yelchuru, R., Jäschke, J., 2011. Optimal use of measurements for control, optimization and estimation using loss method: Summary of existing results and some new. In: Selected Topics on Constrained and Nonlinear Control Workbook. STU Bratislava and NTNU Trondheim.