Effect of error propagation in successive parameter estimation

V. Kumar¹, N. Kerimoglu², J.W. Thybaut¹, A. Mhamdi², G.B. Marin¹ and W. Marquardt²

¹ Laboratory of Chemical Technology, Ghent University, Ghent, 9000, Belgium
² AVT-Process Systems Engineering, RWTH Aachen University, Aachen, 52056, Germany

MaCKiE 2013, Chennai, India, 04-06/02/2013
**Introduction**

- **Physical meaning**
  - Scientific background
  - Model complexity
- **Mathematical representation**
  - Experimental details
  - Convergence

Input → Model → Output

Parameter estimation

Experimental data

Model complexity

Experimental data

- $\beta_1$
- $\beta_2$

**Cost function**

-1
-0.5
0
0.5
1

-1
-0.5
0
0.5
1
balancing the “pillars”

- Ill-conditioning of the system
  - Identifiability of parameters

- How?
  - Reduce complexity
  - Fix parameters
    - A-priori determined
    - Literature
objective

• Uncertainty in the determined parameter(s)
  – Measurements errors
  – Theoretical calculations

To study the propagation of error in the fixed parameters in successive parameter estimation
• Introduction
• Mathematical background
  – Linear regression analysis
• Case study
  – Well-conditioned
  – Ill-conditioned
  – Reaction kinetics example
• Conclusions
All parameters estimated

\[ \tilde{Y} = X\beta + \xi \quad \xi = N(0, V(\xi)) \]

\[ \hat{\beta} = \text{argmin}_{\beta} [(\tilde{Y} - X\beta)^2] \]

using linear transformations;

\[ \hat{\beta} = \left( X^T X \right)^{-1} X^T \tilde{Y} \]

\[ \hat{\beta} = \beta + \varepsilon \quad \varepsilon = N(0, V(\varepsilon)) \]

\(\hat{\beta}\) is the true estimate of \(\beta\) with an error, \(\varepsilon\)

linear regression analysis

Mathematically,

$$B = AY$$

$$V(B) = V(AY)$$

$$V(B) = AV(Y)A^T$$

$$\hat{\beta} = (X^TX)^{-1}X^T\tilde{Y}$$

$$V(\hat{\beta}) = (X^TX)^{-1}X^TV(\xi)X^T(X^TX)^{-1}$$

In case of constant variance in the measurement error

$$V(\xi) = I\sigma^2$$

by applying matrix operations to the above equation;

$$V(\hat{\beta}) = (X^TX)^{-1}\sigma^2$$
A subset of parameters is fixed

\[ \tilde{Y} = X_1 \beta_1 + X_2 \beta_2 + \xi \]

\[ \hat{\beta}_1 = \beta_1 + \delta \quad \delta = N(0, V(\delta)) \]

\[ \hat{\beta}_2 = \arg\min_{\beta_2} \left[ \left( (\tilde{Y} - X_1 \beta_1) - X_2 \beta_2 \right)^2 \right] \]

using linear transformations;

\[ \hat{\beta}_2 = \left( X_2^T X_2 \right)^{-1} X_2^T (\tilde{Y} - X_1^T \beta_1) \]

\[ \hat{\beta}_2 = \beta_2 + \varepsilon \quad \varepsilon = N(0, V(\varepsilon)) \]

Propagation of variances measurement error and Variances in fixed parameters into the estimated parameters

\[ V(\hat{\beta}_2) = \left( X_2^T X_2 \right)^{-1} X_2^T \left[ V(\tilde{Y} - X_1^T \beta_1) \right] X_2 \left( X_2^T X_2 \right)^{-1} \]
linear regression analysis

Variances of the modified measurements,

\[ V(\tilde{Y} - X_1^T \beta_1) = V(\tilde{Y}) - 2 \text{cov}(\tilde{Y}, X_1^T \beta_1) + X_1 V(\beta_1) X_1^T \]

Replacing above expression in the expression for \( V(\hat{\beta}_2) \)

\[
V(\hat{\beta}_2) = \left( X_2^T X_2 \right)^{-1} X_2^T \left[ V(\tilde{Y} - X_1^T \beta_1) \right] X_2 \left( X_2^T X_2 \right)^{-1}
\]

\[
V(\hat{\beta}_2) = \left( X_2^T X_2 \right)^{-1} X_2^T \left[ V(\tilde{Y} - 2 \text{cov}(\tilde{Y}, X_1^T \beta_1) + X_1 V(\beta_1) X_1^T \right] X_2 \left( X_2^T X_2 \right)^{-1}
\]

Writing \( V(\hat{\beta}_2) \) in terms of errors,

\[
V(\hat{\beta}_2) = \left( X_2^T X_2 \right)^{-1} X_2^T \left[ V(\xi) - 2 \text{cov}(\xi, X_1^T \delta) + X_1 V(\delta) X_1^T \right] X_2 \left( X_2^T X_2 \right)^{-1}
\]

No correlation between the error in measurements and the error in fixed parameter(s)

\[ V(\hat{\beta}_2) = \left( X_2^T X_2 \right)^{-1} X_2^T \left[ V(\xi) + X_1 V(\delta) X_1^T \right] X_2 \left( X_2^T X_2 \right)^{-1} \]
• Introduction

• Mathematical background
  – Linear regression analysis

• Case study
  – Well-conditioned
  – Ill-conditioned
  – Reaction kinetic example

• Conclusions
well versus ill conditioned systems

\[ \tilde{Y} = X_1 \beta_1 + X_2 \beta_2 + \xi \]

System 1

\[
\begin{align*}
X_1 &= [3 \quad 8 \quad 2 \quad 11 \quad 7]^T \\
X_2 &= [4 \quad 19 \quad 5 \quad 15 \quad 10]^T \\
\beta &= [2 \quad 3]^T
\end{align*}
\]

System 2

\[
\begin{align*}
X_1 &= [50 \quad 190 \quad 205 \quad 300 \quad 340]^T \\
X_2 &= [4 \quad 19 \quad 5 \quad 15 \quad 10]^T \\
\beta &= [0.2 \quad 3]^T
\end{align*}
\]

\[ X = [X_1 \quad X_2] \]

condition number of \(X^TX\)

75.2 \hspace{1cm} 1791.7

variances (model contribution \((X^TX)^{-1}\))

\[
\begin{bmatrix}
0.0592 & -0.0333 \\
-0.0333 & 0.0201
\end{bmatrix}
\hspace{1cm}
\begin{bmatrix}
1.5837e-05 & -2.7742e-04 \\
-2.7742e-04 & 6.2351e-03
\end{bmatrix}
\]
linear in parameters : well-conditioned

\[ \hat{y} = \begin{bmatrix} 3 & 4 \\ 8 & 19 \\ 2 & 5 \\ 11 & 15 \\ 7 & 10 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \xi \]

\[ \xi = N(0, 2.0) \]

\[ \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \]

\[ \delta = N(0, 0.5) \]

Monte Carlo simulations with 1000 realizations for two set-ups

<table>
<thead>
<tr>
<th></th>
<th>Fixed Parameter Without Error Density</th>
<th>Fixed Parameter With Error Density</th>
<th>Analytical solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean of ( \beta_2 )</td>
<td>2.9992±0.1510</td>
<td>2.9998±0.5636</td>
<td>-</td>
</tr>
<tr>
<td>Variance of ( \beta_2 )</td>
<td>0.0057</td>
<td>0.0794</td>
<td>0.0846</td>
</tr>
<tr>
<td>Mean of ( \beta_1 )</td>
<td>1.9960±0.2623</td>
<td>2.0494±1.6133</td>
<td>-</td>
</tr>
<tr>
<td>Variance of ( \beta_1 )</td>
<td>0.0172</td>
<td>0.6507</td>
<td>0.7017</td>
</tr>
</tbody>
</table>
linear in parameters : ill-conditioned

\[ \hat{y} = \begin{bmatrix} 50 \\ 190 \\ 205 \\ 300 \\ 340 \end{bmatrix} \begin{bmatrix} 4 \\ 19 \\ 5 \\ 15 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \xi \]

\[ \xi = N(0,2.0) \]

\[ \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 0.2 \\ 3 \end{bmatrix} \]

\[ \hat{\beta}_1 = 0.2 + \delta_1 \]

\[ \delta_1 = N(0,0.05) \]

\[ \hat{\beta}_2 = 3 + \delta \]

\[ \delta = N(0,0.5) \]

Monte Carlo simulations with 1000 realizations for two set-ups

<table>
<thead>
<tr>
<th></th>
<th>Fixed Parameter Without Error Density</th>
<th>Fixed Parameter With Error Density</th>
<th>Analytical solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean of ( \beta_2 )</td>
<td>2.9973±0.1612</td>
<td>3.0096±1.7087</td>
<td>-</td>
</tr>
<tr>
<td>Variance of ( \beta_2 )</td>
<td>0.0065</td>
<td>0.7299</td>
<td>0.7726</td>
</tr>
<tr>
<td>Mean of ( \beta_1 )</td>
<td>0.2000±0.0001</td>
<td>0.1997±0.0087</td>
<td>-</td>
</tr>
<tr>
<td>Variance of ( \beta_1 )</td>
<td>1.5000E-09</td>
<td>1.8900E-05</td>
<td>5.0888E-04</td>
</tr>
</tbody>
</table>
Reaction:

\[
A \xrightarrow{k_1} B \xrightarrow{k_2} C \xrightarrow{k_3} D
\]

Data generation:

\(k_1 = 2.5 \text{ sec}^{-1}\)
\(k_2 = 3.5 \text{ sec}^{-1}\)
\(k_3 = 5.6 \text{ sec}^{-1}\)

Measurement error: 3\% in output

Volume: 2L

Flow varies from 0.05 – 10 L/sec

\[
C_A^0 = 2 \text{ mol/L} \quad C_B^0 = 1 \text{ mol/L} \quad C_C^0 = 0.75 \text{ mol/L} \quad C_D^0 = 0.25 \text{ mol/L}
\]

Reactor model:

\[
F_i^{\text{in}} - F_i^{\text{out}} + R_i V = 0 \quad \text{where } i = A, B, C \text{ and } D
\]
fixed parameter: small variances

\[ k_3 = 5.6 + \delta \quad \delta = N(0,0.3) \]

\[ A \xrightarrow{k_1} B \xrightarrow{k_2} C \xrightarrow{k_3} D \]
fixed parameter: large variances

\[ k_3 = 5.6 + \delta \quad \delta = N(0, 2.5) \]
\[ k_3 = 5.6 + \delta \quad \delta = N(0, 2.5) \]

\[ A \xrightarrow{k_1} B \xrightarrow{k_2} C \xrightarrow{k_3} D \]
\[ k_3 = 5.6 + \delta \quad \delta = N(0, 2.5) \]

\[ A \xrightarrow{k_1} B \xrightarrow{k_2} C \xrightarrow{k_3} D \]

- \( k_3 < 5.6 \), third step is rate-determining
  - Under predicted products and hence, over predicted reactants

- \( k_3 > 5.6 \), other steps are rate-determining
  - No major change in parity

\[ k_3 = 0.06 \rightarrow 12.96 \text{ sec}^{-1} \]
fixed parameter at wrong value

\[ k_3 = 5.6 + \delta \]
\[ \delta = N(0, 2.5) \]

\[ k_3 = 3.0 + \delta \]

\[ k_3 = 8.0 + \delta \]
• Propagation of errors in the fixed parameters has been studied successfully for linear cases.
• Variances of the estimated parameters are amplified significantly because of the ill conditioning, while for well-conditioned system the propagation is not so pronounced.
• Kinetic example(s) are limited by the reaction behaviour with larger variance in the fixed parameters, while in case of smaller variances, statistics dominates.
• The uncertainties in the fixed parameters should be accounted for in optimal experimental design.
The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme FP7/2007-2013 under grant agreement n° 238013.
Thank you!