METHODS FOR PARAMETER RANKING IN NONLINEAR, MECHANISTIC MODELS

Berit Floor Lund [*,1] **Hans E. Berntsen** [**]
Bjarne A. Foss [*]

[*] *Norwegian University of Science and Technology*
[**] *SINTEF*
Institute of Engineering Cybernetics, NTNU, 7491
Trondheim, Norway
berit.floor.lund@sintef.no

Abstract: The paper addresses efficient methods for parameter sensitivity analysis and ranking in large, nonlinear, mechanistic models requiring examination of many points in the parameter space. The paper shows how orthogonal decomposition and permutation of the sensitivity derivative is an intuitive and structured method for automatic ranking of the parameters within a candidate set. Provided the model error is Gaussian, and with the problem on a triangularized form, the additional variance associated with each parameter can easily be found. Ranking according to additional variance is therefore another option. The methods are tested on an industrially used simulator model. *Copyright © IFAC 2005*

Keywords: Sensitivity analysis, physical models, nonlinear models, parameter estimation, identifiability.

1. INTRODUCTION

Analysis of the sensitivity derivative of a model is a necessary step when designing model based parameter estimation algorithms. This paper addresses methods for efficient manual and automated sensitivity analysis for nonlinear models on the general form

$$\hat{y} = g(\theta, u)$$

where $\hat{y} \in \mathbb{R}^{n_y}$ is the model output vector, $\theta \in \mathbb{R}^{n_\theta}$ the parameter vector and $u \in \mathbb{R}^{n_u}$ is a measured input vector which may be present in the model. The sensitivity derivative of the model outputs to the parameter vector is defined as

$$S = W^{-\frac{1}{2}} \frac{d\hat{y}(\theta, u)}{d\theta} \in \mathbb{R}^{n_y \times n_\theta}$$

where $W^{-\frac{1}{2}}$ is a diagonal weighting matrix. Each column of the sensitivity derivative will express one parameter's sensitivity in all outputs, and can therefore be viewed as the sensitivity direction for the corresponding parameter.

The focus in this paper is on situations when the sensitivity derivative is on numerical form, as opposed to situations when an analytical expression for S can be found. Since the model is nonlinear, global identifiability can generally not be proven. To increase the probability that the model is identifiable over the whole parameter space, the sensitivity derivative must be checked for as many parameter vector values as possible. If the model is also nonlinear in the input vector u, then u will affect the sensitivity derivative, and S needs to be calculated for a wide range of values of u as well. This requires efficient automated sensitivity analysis methods.

In sensitivity analysis, the following properties of the sensitivity derivative are of interest:

- Many large values in a column are an indication of high sensitivity to the corresponding parameter.
- Large differences in the norms of the columns indicate large differences in sensitivity or poor scaling.
- Degree of linear dependence between columns

Many linear transformations of S or $S'S$ will reveal these properties. A search into relevant literature databases covering the control engineering domain verify that eigenvector transformation of $S'S$ is the most commonly used method. Eigenvector transformation can be used for manual parameter ranking through ordering the eigenvalues of $S'S$ according to size, and inspecting the corresponding eigenvectors to determine which parameters are the most significant contributors to this particular direction. Generally this may not be a trivial task since the eigenvector decomposition produces linear combinations of the original directions of the problem (columns of S).

A large condition number, $\kappa = \sqrt{\frac{\max eig(S'S)}{\min eig(S'S)}}$, indicates either weak individual sensitivity or linear dependence within S. Eigenvalue transformation can therefore be used for automatic ranking of the parameters, by removing (combinations of) columns from S and calculate the condition number of the corresponding sub-matrices of $S'S$. As an automated method this will give a trial and error method.

An alternative method for automatic ranking proposed in this paper utilizes the original directions of S, and reduces the column space of S into an orthogonal set of vectors in a successive manner. This provides an intuitive and structured way of ranking the parameters and is carried out as follows. From the non-selected set of columns in S, select the column with the highest norm, form a unit vector, and remove this direction (by projection and subtraction) from the non selected set of columns. The procedure is repeated until all columns have been selected. The order of selection is stored in a permutation matrix. This economical QR decomposition with permutation of the sensitivity derivative will give a triangular form of $S'S$. The same form can be found by LDL' decomposition and permutation of $S'S$. Successive orthogonalization with permutation is described by example in section 3 of this paper.

If the deviation between the actual and simulated output vectors, i.e. the model error $\varepsilon = y - \hat{y}(\theta)$, can be assumed to be a stochastic process with Gaussian properties, then $tr(S'S)^{-1}$ gives an estimate of the lower limit of the parameter covariance (Söderström and Stoica 1989).

(Berntsen 1977) showed that if $S'S$ is already on a triangular form, a particularly easy form of the individual variance contribution of each parameter can be found, see section 4. This can be applied to an already ordered set of parameters or used as another method to rank the parameters.

Manual inspection of $S'S$ can in simple cases give sufficient information to rank the parameters. Manual inspection is however also useful in more complex cases to gain initial insight into important features of the problem, and provide a basis to understand or second-guess the ranking done by automated procedures. A transformation of $S'S$ which is particularly useful for manual inspection is presented in section 2.

All three methods have been applied to an example 21×6 sensitivity derivative in section 5. Manual inspection has been used initially to reveal the most important properties of the example matrix. Next, the parameters have been ranked through successive orthogonalization and by smallest additional variance. The methods have also been applied to a larger, industrial example, and the results of this analysis have been summarized in section 6. Conclusions are given in section 7.

2. MANUAL INSPECTION OF $S'S$

In the following a transformation of $S'S$ giving a particularly useful form for manual inspection, is demonstrated. The following is valid for any sensitivity derivative S with dimension $n_y \times n_\theta$, $n_y \geq n_\theta$, but for pedagogical reasons, an $n_y \times 3$ matrix, $S = \begin{bmatrix} a & b & c \end{bmatrix}$, is used. a, b, c are the column vectors.

Cross multiplication of S gives

$$S'S = \begin{bmatrix} \|a\|^2 & \langle a, b \rangle & \langle a, c \rangle \\ \langle b, a \rangle & \|b\|^2 & \langle b, c \rangle \\ \langle c, a \rangle & \langle c, b \rangle & \|c\|^2 \end{bmatrix}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product and $\|\cdot\|$ the Euclidian norm of a vector. The squared norms on the diagonal give a measure of the total sensitivity of each individual parameter. Large inner products in the off diagonal elements indicate that the two vectors have large elements in the same places (linear dependence), and/or a large norm of the vectors involved. To separate the information about the impact of each individual parameter from the information about linear dependence, the norms ($S(i,i) > 0$) are extracted and $S'S$ to written on the form shown in equation (1).

In (1) the first and last matrices contain information about the "strength" of each parameter, as sensitivity vector length. The off diagonal terms of the middle matrix can be recognized from the

$$S'S = D'CD = \begin{bmatrix} \|a\| & 0 & 0 \\ 0 & \|b\| & 0 \\ 0 & 0 & \|c\| \end{bmatrix} \begin{bmatrix} 1 & \dfrac{\langle a,b \rangle}{\|a\|\,\|b\|} & \dfrac{\langle a,c \rangle}{\|u\|\,\|w\|} \\ \dfrac{\langle b,a \rangle}{\|b\|\,\|a\|} & 1 & \dfrac{\langle b,c \rangle}{\|b\|\,\|c\|} \\ \dfrac{\langle c,a \rangle}{\|c\|\,\|a\|} & \dfrac{\langle c,b \rangle}{\|c\|\,\|b\|} & 1 \end{bmatrix} \begin{bmatrix} \|a\| & 0 & 0 \\ 0 & \|b\| & 0 \\ 0 & 0 & \|c\| \end{bmatrix} \quad (1)$$

Cauchy Schwarz inequality, provided $\|x\|, \|y\|, > 0$, then

$$-1 \leq \frac{\langle x,y \rangle}{\|x\|\,\|y\|} \leq 1 \quad (2)$$

(2) can in $\mathbb{R}^2$ be recognized as the cosine of the angle between the vectors. Strong linear dependence gives values close to -1 or 1, whereas near orthogonality gives values close to zero.

The condition number of the middle matrix in (1) is known as "collinearity index", see (Brun *et al.* 2000) and (Belsley 1991). (Brun *et al.* 2000) uses the collinearity index for reduction of the parameter set. He removes columns and combinations of columns from S and calculates the corresponding collinearity index. This is combined with inspection of different norms of the columns of the sensitivity derivative in order to decide on which parameter set to use. A similar analysis was also made in a very initial stage of this the identification problem design, see (Lund *et al.* 2004). The method is however largely manual, and not very well suited for analysis of a large number of sensitivity derivative matrices.

3. RANKING USING ORTHOGONALIZATION OF S

In the following parameter ranking is made utilizing the directions already present in the sensitivity derivative S. S is as defined in section 2.

Assume vector a has the largest norm and therefore corresponds to the parameter having the largest overall sensitivity. This direction is chosen, and a unit vector, $q_1 = \frac{a}{\|a\|}$, is formed and removed from b and c by projection onto q_1 and subtraction, giving

$$\widetilde{b} = b - \frac{q_1'b}{q_1'q_1} q_1 = b - (q_1'b)q_1$$
$$\widetilde{c} = c - (q_1'c)q_1$$

$\widetilde{b}$ and $\widetilde{c}$ are now orthogonal to q_1. Next, the vector out of $\widetilde{b}$ and $\widetilde{c}$ with the largest norm is selected. Assuming b originally was pointing in nearly the same direction as a, then a large component was subtracted from b when forming $\widetilde{b}$. Assume for now that $\widetilde{c}$ has the larger norm, and therefore is selected as the second unit vector, $q_2 = \frac{\widetilde{c}}{\|\widetilde{c}\|}$. The q_2 direction is removed from vectors of the

remaining set, which is now only $\widetilde{b}$, to form the new vector $\bar{b}$

$$\bar{b} = \widetilde{b} - (q_2'\widetilde{b})q_2$$

and set $q_3 = \frac{\bar{b}}{\|\bar{b}\|}$. The above example can be summarized into the expression $SE = QR$

$$S \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} \begin{bmatrix} q_1'a & q_1'c & q_1'b \\ 0 & q_2'c & q_2'b \\ 0 & 0 & q_3'b \end{bmatrix}$$

where S is $n_y \times 3$, with $n_y \geq 3$. E is a 3×3 permutation matrix. Q is an orthonormal $n_y \times 3$ and R is an upper triangular 3×3 matrix. The procedure can be characterized as "economy" QR decomposition with permutation. The term "economy" is used since full QR decomposition produces a full $n_y \times n_y$ Q-matrix, and $n_y \times n_\theta$ R-matrix. The order of selection of the vectors is expressed by the permutation matrix.

The parameters have now been selected according to their strength in linearly independent directions. The main advantage of this method is that is utilizes the directions already present in the sensitivity derivative, and that the method provides a structured order of selection as opposed to trial and error methods.

The result can be further refined by extracting the diagonal of R , $R = D\bar{R}$ and find the cross product $E'S'SE$

$$E'S'SE = \overline{R}'D'Q'QD\overline{R} = \overline{R}'D^2\overline{R}$$

since $Q'Q = I$, and $D'D = D^2$. The expression now has the triangularized form

$$E'S'SE = \begin{bmatrix} 1 & 0 & 0 \\ \dfrac{q_1'c}{q_1'a} & 1 & 0 \\ \dfrac{q_1'b}{q_1'a} & \dfrac{q_2'b}{q_2'c} & 1 \end{bmatrix} D^2 \begin{bmatrix} 1 & \dfrac{q_1'c}{q_1'a} & \dfrac{q_1'b}{q_1'a} \\ 0 & 1 & \dfrac{q_2'b}{q_2'c} \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

with

$$D^2 = \begin{bmatrix} (q_1'a)^2 & 0 & 0 \\ 0 & (q_2'c)^2 & 0 \\ 0 & 0 & (q_3'b)^2 \end{bmatrix}$$

The elements of D contain the squared lengths of the orthogonal fractions of the original sensitivity vectors.

LU (LDL') decomposition on $S'S$ with permutation to maximize the norm of the diagonal elements in each step, will give the same result as in (3).

4. RANKING ACCORDING TO MINIMUM PARAMETER VARIANCE

In the following it is shown that a simple expression for the individual variance contribution of each parameter can be found if $S'S$ is on triangular form.

The Cramer Rao lower bound (Ljung 1999) states that for any unbiased estimator $\hat{\theta}$ of θ

$$cov\,\hat{\theta} \geq M^{-1}$$

where M is the Fisher information matrix. Assuming the model error $\varepsilon = y - \hat{y}(\hat{\theta}_N)$, with $\varepsilon, y \in \mathbb{R}^{n_y}$, is stochastic with Gaussian properties and covariance W, then the Fisher information matrix is

$$M = \left[\frac{d\varepsilon}{d\theta}\right]' W^{-1} \left[\frac{d\varepsilon}{d\theta}\right]$$

Then $\frac{d\varepsilon}{d\theta} = -\frac{d\hat{y}(\theta)}{d\theta}$ and

$$M = \left[\frac{d\hat{y}(\hat{\theta})}{d\theta}\right]' W^{-1} \frac{d\hat{y}(\hat{\theta})}{d\theta} = S'S$$

An estimate of the variance lower bound is then

$$var\,\hat{\theta} \geq tr(S'S)^{-1}$$

If $S'S$ has been triangularized, and $U = (\bar{R}^{-1})'$ then

$$(S'S)^{-1} = (\bar{R}'D^2\bar{R})^{-1} = UD^{-2}U'$$

Selecting a subset $\beta \in \theta$ consisting of the parameters 1 through p, $p \leq n_\theta$, gives the variance

$$var(\beta_p) = \sum_{i=1}^{p} \frac{\|u_i\|^2}{d_i^2}$$

where u_i is column i in U. Again, a simple example is used to determine the increase in variance going from 2 to 3 parameters. If U_2 is a 2×2 lower triangular matrix with a unit diagonal, then $var(\beta_2) = tr(U_2 D_2^{-2} U_2')$ is

$$var(\beta_2) = tr \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{d_1^2} & 0 \\ 0 & \frac{1}{d_2^2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}$$

$$= tr \begin{bmatrix} \left(\frac{1}{d_1^2} + \frac{a^2}{d_2^2}\right) & \left(\frac{a}{d_2^2}\right) \\ \left(\frac{a}{d_2^2}\right) & \left(\frac{1}{d_2^2}\right) \end{bmatrix} = \frac{1}{d_1^2} + \frac{a^2}{d_2^2} + \frac{1}{d_2^2}$$

Expanding with a third parameter gives $var(\beta_3) = tr(U_3 D_3^{-2} U_3')$

$$tr \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{d_1^2} & 0 & 0 \\ 0 & \frac{1}{d_2^2} & 0 \\ 0 & 0 & \frac{1}{d_3^2} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{bmatrix}$$

$$= \frac{1}{d_1^2} + \frac{a^2}{d_2^2} + \frac{b^2}{d_3^2} + \frac{1}{d_2^2} + \frac{c^2}{d_3^2} + \frac{1}{d_3^2}$$

$$= var(\beta_2) + \frac{u_3'u_3}{d_3^2}$$

where u_3 is the third column added to form U_3 from U_2. Generally (Berntsen 1977)

$$var(\beta_{p+1}) = var(\beta_p) + \frac{\|u_{p+1}\|^2}{d_{p+1}^2} \qquad (4)$$

where $\frac{\|u_{p+1}\|^2}{d_{p+1}^2}$ is an expression for the additional variance introduced to the total variance when adding parameter $p + 1$ to the set.

The parameters can therefore be ordered according to minimum additional variance, and the order of selection can be represented by a permutation matrix, E_v giving the expression

$$E_v' (S'S)^{-1} E_v = UD^{-2}U'$$

5. EXAMPLE

The case used here is the sensitivity derivative from a nonlinear, dynamic model representing the mass transport, chemical reactions and thermodynamic phenomena of a submerged arc silicon furnace (Foss and Wasbø 2001). The model is a differential algebraic equation system, and is called Simod.

Most of the candidate parameters express uncertainties associated with the inputs and the energy distribution to the furnace. These parameters can also be viewed as process disturbances.

The sensitivity derivative, S, has been obtained using a central difference approximation, therefore each element s_{ij} in S is

$$s_{ij} = \frac{1}{w_i^{\frac{1}{2}}} \cdot \frac{d\hat{y}_i}{d\theta_j} \approx \frac{1}{w_i^{\frac{1}{2}}} \cdot \frac{\hat{y}_i(\theta_j + \Delta\theta_j) - \hat{y}_i(\theta_j - \Delta\theta_j)}{2\Delta\theta_j}$$

$\Delta\theta_j$ is the perturbation interval for θ_j which is the parameter value of element j in the parameter vector θ. $w_i^{\frac{1}{2}}$ is the standard deviation of the noise of the i'th element of the actual output vector. The reason for including this standard deviation is to avoid giving a high score to model outputs which correspond to noisy, actual outputs. It is paramount that $\frac{d\hat{y}}{d\theta}$ is scaled according to the natural variation range of the parameters and the outputs as well.

Since global identifiability generally can not be proven for the nonlinear model, sensitivity analysis should be done for as many parameter vector values as possible. For demonstration of the methods presented in the paper, only one parameter vector, $\theta = \begin{bmatrix} 0.2 & 0.6 & 13.7 & 0.03 & 0.92 & 0.1 \end{bmatrix}$, has been used. The scaled sensitivity derivative at this parameter vector value is $S =$

$$
\begin{bmatrix}
-2.204 & -0.08309 & -0.4096 & 0.03563 & 0.07694 & -0.7551 \\
-1.716 & -0.2744 & -5.494 & 0.3501 & 0.1312 & 0.6418 \\
-0.3569 & -0.1394 & 2.019 & 0.04123 & -0.02163 & -0.2508 \\
-2.159 & -0.02855 & -0.3459 & 0.02090 & 0.04908 & -0.6824 \\
-1.243 & 0.5462 & -3.950 & -0.7891 & 0.8500 & -0.1928 \\
-0.3501 & -0.1438 & 1.987 & 0.05719 & -0.01304 & -0.2590 \\
-2.044 & -0.1289 & -0.4124 & 0.03348 & 0.05143 & -0.8040 \\
-0.9978 & -1.964 & -9.796 & -0.5977 & 0.1140 & -1.310 \\
-0.3907 & 0.1841 & 3.203 & -0.1114 & -0.001869 & -0.2088 \\
-2.168 & -0.04300 & -0.1918 & -0.007997 & 0.02347 & -0.8847 \\
-1.681 & 0.2567 & -6.478 & 0.4334 & -0.7772 & -0.3600 \\
-0.3444 & 0.07747 & 2.691 & -0.04052 & 0.002749 & -0.1569 \\
-2.224 & -0.05802 & -0.3568 & 0.03075 & 0.01231 & -0.9006 \\
-1.764 & 0.2127 & -5.409 & -0.01474 & -0.7880 & -0.7065 \\
-0.3106 & -0.04527 & 2.255 & 0.04534 & 0.01331 & -0.1017 \\
-2.270 & -0.08749 & -0.4579 & 0.05717 & 0.02911 & -0.8137 \\
-1.032 & 0.6817 & -6.628 & 0.3682 & -0.5508 & -0.3650 \\
-0.3482 & -0.004523 & 2.392 & -0.003359 & 0.006071 & -0.1333 \\
-2.299 & -0.04337 & -0.2732 & 0.01182 & 0.03450 & -0.7782 \\
-1.264 & -0.1179 & -5.955 & -0.2546 & 0.2906 & -0.2511 \\
-0.4098 & -0.03964 & 2.458 & -0.02861 & 0.008653 & -0.1884
\end{bmatrix}
$$

The cross product on the form (1) is first inspected manually. Next the parameters are ranked according to highest squared norm of the orthogonal directions, and according to smallest additional variance.

5.1 Analysis by manual inspection of $S'S$

Transformation of $S'S$ to the form (1), gives for $D =$

$$diag[\, 7.0 \ 2.223 \ 18.320 \ 1.233 \ 1.54 \ 2.792 \,] \qquad (5)$$

and $C =$

$$
\begin{bmatrix}
1.0 & 0.093 & 0.454 & -0.020 & 0.084 & 0.810 \\
0.093 & 1.0 & 0.298 & 0.358 & -0.172 & 0.362 \\
0.454 & 0.298 & 1.0 & 0.136 & 0.209 & 0.370 \\
-0.020 & 0.358 & 0.136 & 1.0 & -0.677 & 0.234 \\
0.084 & -0.172 & 0.209 & -0.677 & 1.0 & 0.122 \\
0.810 & 0.361 & 0.371 & 0.234 & 0.122 & 1.0
\end{bmatrix}
$$

By inspection of D, the third parameter shows the highest individual sensitivity. This parameter also has relatively little linear dependence with the other parameters. The second parameter can therefore most likely be selected independently of the first, and parameter 1 is a likely candidate since it has second largest element in D. From this point on, manual selection is difficult. $C_{1,6}$ indicates that there is a relatively strong linear dependence between parameters 1 and 6. $D_{6,6}$ has approximately the same size as $D_{2,2}$, and with the linear dependence between 6 and 1, it is difficult to know if parameter 2 or parameter 6 is the third choice. Also, element $C_{4,5}$ indicates that there is some degree of linear dependence between the corresponding parameters. From this the conclusion is that parameters 3 and 1 most likely are the first choices. As the third choice, parameter either 2 or 6 is the most likely choice. Parameter 5 may be a choice before 4, but this is uncertain.

5.2 Successive orthogonalization of $S'S$

In the successive orthogonalization of S the QR-function of Matlab has been used, running it in "economy" mode. Negative elements on the diagonal of R have been corrected in R and Q. Ranking order when determined according to maximum orthogonal length is $3, 1, 2, 5, 6, 4$. The ranking order is expressed in the permutation matrix in E. The corresponding values of $D^2 =$

$$diag \begin{bmatrix} 335.6 & 38.7 & 4.49 & 2.13 & 1.84 & 0.28 \end{bmatrix} \qquad (6)$$

It is interesting to compare the values of (6) with the values of (5), squared and ordered according to the ranking caused by the orthogonalization. Due to the orthogonalization, only the first chosen direction will have the same squared norm in both cases, the others will have a more or less reduced norm depending on the degree of linear dependence with already chosen directions.

5.3 Ranking by smallest additional variance

Ranking by minimum additional variance gives the same order as with successive orthogonalization, $3, 1, 2, 5, 6, 4$. The corresponding, additional variance contribution of each parameter according to $D^{-2}diag(UU')$ in (4) are

$$diag \begin{bmatrix} 0.003 & 0.027 & 0.223 & 0.489 & 0.745 & 6.048 \end{bmatrix} \quad (7)$$

Total variance of the set increases 9 times going from 1 to 2, or from 2 to 3 parameters. From 3 to four, or four to five, variance approximately doubles. From 5 to 6, the total variance is dominated by the contribution from the last selected parameter. The last parameter added to the set is parameter 4, which in the inspection of the cross product revealed a low individual sensitivity as well as common directions with other parameters. The results are therefore consistent with the manual inspection.

5.4 Application to a larger example

A more extensive sensitivity analysis of the candidate parameters of the Simod model was performed, using the successive orthogonalization and smallest additional variance methods. The sensitivity derivative was obtained in a sliding window fashion. Assuming a window of N time instants, the sensitivity derivative at time t is formed by stacking the outputs registered from time $(t - N + 1)$ to t. In the Simod case, 3 outputs are registered at each time instant. Using a window of $N = 7$, with $n_\theta = 6$, each sensitivity derivative will have the dimension 21×6.

The sensitivity derivative was obtained by perturbation around 6 different values of the θ vector. A total of 220 time instants of the real process input vector, u, were used. For $N = 7$, a total of 1290 evaluations of the sensitivity derivative were made. In 230 cases (18%) the two methods gave a different ranking of the parameters. The longer windows, $N = 20$ and $N = 50$, gave different ranking in 25% and 31% of the cases respectively.

Since analysis was made over such a high number of sensitivity derivatives matrices, a mean ranking with corresponding variance was calculated for both methods. The mean ranking generally ordered the parameters the same way. The internal ranking of the parameters varied somewhat depending on N, and also on the value of θ. With respect to the properties of the parameter set, the conclusion still was that the parameter set can be divided into a "good" half consisting of parameters 3, 1 and 2 all giving a low additional variance. The parameters of the "less good" half, parameters 4, 5 and 6, generally add a higher variance when included in the set.

6. CONCLUSION

When parameters in a mechanistic model are chosen as candidates for estimation, the choices are based on model and process knowledge. Therefore, complete linear dependence between sensitivity vectors or zero sensitivity in certain parameters is seldom experienced, and the sensitivity derivative will have full rank. Yet one may still consider if it is worth-while identifying the full set of parameters. Ranking methods are therefore important, and automatic ranking methods even more so if the parameter vector is large or if a large number of parameter values need to be examined. The paper has indicated how successive orthogonalization of the sensitivity derivative can be used efficiently to determine parameter ranking automatically. The method uses the directions already present in the sensitivity derivative, as opposed to eigenvector transformation which generates linear combinations of the original directions. If the innovations are Gaussian, the parameters can also be ranked according to individual variance contributions. If the cross product of the sensitivity derivative is on a triangular form, the additional variance calculations become particularly easy. The methods are demonstrated on a small example in the paper and a summary of a larger example is given. The larger example shows that the mean ranking of the methods give the same parameter selection order, although for individual calculations, the parameters were also ranked differently. A particularly useful transformation for manual inspection of $S'S$ was also demonstrated. This transformation also utilizes the original directions of the sensitivity derivative, and separates the information about sensitivity vector length from linear dependence information.

REFERENCES

Belsley, D. A. (1991). *Conditioning Diagnostics. Collinearity and Weak Data in Regression.* John Wiley and Sons Inc.

Berntsen, H. E. (1977). Modelling, Parameter Estimation and Identifiability in Electrocadiology. PhD thesis. Norwegian Institute of Technology.

Brun, R., P. Reichert and H. R. Kunsch (2000). Practical identifiability analysis of large environmental simulation models. Technical report. EAWAG, ETH.

Foss, B. A. and S. O. Wasbø (2001). An integration scheme for stiff solid-gas reactor models. *Computer Methods in Applied Mechanics and Engineering* **190/45**, 6009–6021.

Ljung, L. (1999). *System Identification. Theory for the User.* Prentice-Hall Inc.

Lund, B. F., B. A. Foss, K. R. Løvåsen and B. E. Ydstie (2004). Sensitivity analysis of a dynamic model for submerged arc silicon furnaces. *Proceedings of the Infacon X Congress.*

Söderström, T. and P. Stoica (1989). *System Identification.* Prentice Hall.