XVI IMEKO World Congress
Measurement - Supports Science - Improves Technology - Protects Environment ... and Provides Employment - Now and in the Future
Vienna, AUSTRIA, 2000, September 25-28

PARAMETER SELECTION IN INDIRECT MEASUREMENTS

A.G. Polak and J. Mroczka
Chair of Electronic and Photonic Metrology
Department of Electronics, Wroclaw University of Technology
ul. B. Prusa 53/55, 50-317 Wroclaw, Poland

Abstract: An alternative method of dealing with ill-conditioned problems in model identification is introduced. It consists in estimation of parameters selected on the ground of the sensitivity analysis and fixing remaining ones as constants. Statistical analysis of the resulting estimator in terms of estimate bias and variance is presented together with a realistic simulation example of the maximum expiration model. Results of simulation show that selection of parameters improves numerical conditioning of model identification and decreases variance of estimates.
Keywords: indirect measurement, ill-conditioned problem, parameter selection

1 INTRODUCTION

The idea of indirect measurements consists in the assumption that there is a casual relationship between unobserved phenomena and measured variables. This dependence can be usually described by a mathematical model. If the model is created by the analysis and description of physical processes going on in the system under investigation (so-called physical modelling [1]) the model parameters correspond to physical features of the system. Identification of such a model yields an estimate of the parameter vector and then enables drawing quantitative conclusions about the system properties not measured directly.

Identification of models of natural objects, especially medical or biological, delivers some kinds of difficulties. The mains of them are so-called ill-posed and ill-conditioned problems [2] arising e.g. from existence of random disturbances, overparametrization and/or collinearities in parameter influence on the measured quantities. Moreover, the physical models of natural systems are usually non-linear in parameters (NLP), so iterative algorithms of identification have to be used.

Many methods dealing with the problems mentioned above have been elaborated so far. Between them the most popular and efficient are: regularisation [2, 3, 4, 5], principal component regression [2, 7] and partial least squares [6, 7]. The common feature of these methods is handling the number of parameters „used" during estimation considerably smaller than the number of „offered" ones. It should be mentioned that they are mainly used with „black-box" type models.

The aim of this paper is to present another way of ill-conditioned problem solving by selection of these parameters which should be estimated. The idea takes advantage of a fact that usually only a few model parameters influence the model output (i.e. the measured variables) considerably. It clearly means that estimation of these parameters only is sufficient for proper solution of the inverse problem. On the other hand, any attempt to estimate the other parameters leads to very high estimate variances, since measurement errors are „transferred" into large adjustment of those parameter values in this instance. The sensitivity analysis is a method of finding the considered subset of parameters.

2 THEORY

For notation simplicity a multiple-input single-output (MISO) NLP model will be considered. Let us denote its output as a vector **y** of n instant values sampled in time and the p-dimensional parameter vector as q. Suppose also that the model is error-free and the measurement data vector **z** is disturbed by an additive, non-correlated white noise with an $n \times n$ covariance matrix **R**.

2.1 Selection of parameters

The matrix h of the model output sensitivity to the parameters is defined as follows

$$\mathrm{h}^{n\times p} \doteq \frac{\P\mathbf{y}}{\P\mathrm{q}} . \qquad (1)$$

A better reference to parameter selection, instead of the sensitivity h, is the relative sensitivity **S**:

$$
\begin{aligned}
&s(i,j) \doteq \frac{\P y_i}{\P q_j}\frac{q_j}{y_i}, \\
&\mathbf{S}^{n\times p} = \left[\mathrm{diag}(\mathbf{y})\right]^{-1}\mathrm{h}\mathrm{Q},
\end{aligned}
\tag{2}
$$

where $\mathrm{Q} = \mathrm{diag}(\mathrm{q})$. The relative sensitivities are *n*-dimensional vectors, so a norm should be used to compare them. The vector **m** of Euclidean norms of the relative sensitivities is used:

$$
\mathbf{m} = \left[\|\mathbf{s}_1\|_2 \ \ \|\mathbf{s}_2\|_2 \ \ \ldots \ \ \|\mathbf{s}_p\|_2\right]^{\mathrm{T}} = \left[\mathrm{diag}\left(\mathbf{S}^{\mathrm{T}}\mathbf{S}\right)\right]^{1/2}. \tag{3}
$$

Sorting elements of the vector **m** from the highest to the lowest value one gets parameters put in order for eventual estimation. Let us choose the first *s* parameters (vector q_S) for estimation and accept the remaining ones (vector q_C) as constants.

2.2 Estimator of the parameter vector

According to the assumptions, there exist such q_{S0} and q_{C0} that

$$
\mathbf{z} = \mathbf{y}(\mathrm{q}_{S0}, \mathrm{q}_{C0}) + \mathbf{e}, \tag{4}
$$

where **e** is a vector of random measurement errors. During model identification it is necessary to choose a value for the vector q_C, say $\mathrm{q}_C = \mathrm{q}_C^*$.

Using the maximum likelihood criterion

$$
V_S(\mathrm{q}_S) = \frac{1}{2}\left(\mathbf{y}(\mathrm{q}_S, \mathrm{q}_C^*) - \mathbf{z}\right)^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{y}(\mathrm{q}_S, \mathrm{q}_C^*) - \mathbf{z}\right) \tag{5}
$$

an estimate $\hat{\mathrm{q}}_S$ of the q_S vector can be calculated by minimising the V_S value. Assuming that $\mathrm{h}_S(\hat{\mathrm{q}}_S, \mathrm{q}_C^*) \approx \mathrm{h}_S(\mathrm{q}_{S0}, \mathrm{q}_C^*)$ and approximating $\mathbf{y}(\hat{\mathrm{q}}_S, \mathrm{q}_C^*)$ by the two first terms of its Taylor expansion at the point $(\mathrm{q}_{S0}, \mathrm{q}_C^*)$ one gets the estimator formula

$$
\hat{\mathrm{q}}_S = \left[\mathrm{h}_S^{\mathrm{T}}(\mathrm{q}_{S0}, \mathrm{q}_C^*)\mathbf{R}^{-1}\mathrm{h}_S(\mathrm{q}_{S0}, \mathrm{q}_C^*)\right]^{-1}\mathrm{h}_S^{\mathrm{T}}(\mathrm{q}_{S0}, \mathrm{q}_C^*)\mathbf{R}^{-1}\left(\mathbf{z} - \mathbf{y}(\mathrm{q}_{S0}, \mathrm{q}_C^*) + \mathrm{h}_S(\mathrm{q}_{S0}, \mathrm{q}_C^*)\mathrm{q}_{S0}\right). \tag{6}
$$

2.3 Bias of the estimator

The expected value of $\hat{\mathrm{q}}_S$ can be calculated by expanding $\mathbf{y}(\mathrm{q}_{S0}, \mathrm{q}_{C0})$ into the Taylor series at the point $(\mathrm{q}_{S0}, \mathrm{q}_C^*)$ and then taking into account its two first expressions

$$
\mathrm{E}\{\hat{\mathrm{q}}_S\} = \mathrm{q}_{S0} + \left[\mathrm{h}_S^{\mathrm{T}}(\mathrm{q}_{S0}, \mathrm{q}_C^*)\mathbf{R}^{-1}\mathrm{h}_S(\mathrm{q}_{S0}, \mathrm{q}_C^*)\right]^{-1}\mathrm{h}_S^{\mathrm{T}}(\mathrm{q}_{S0}, \mathrm{q}_C^*)\mathbf{R}^{-1}\mathrm{h}_C(\mathrm{q}_{S0}, \mathrm{q}_C^*)\Delta\mathrm{q}_C, \tag{7}
$$

where $\Delta\mathrm{q}_C = \mathrm{q}_{C0} - \mathrm{q}_C^*$. This enables determination of the bias vector $\mathbf{b}_S$

$$
\mathbf{b}_S = \mathrm{E}\{\hat{\mathrm{q}}_S\} - \mathrm{q}_{S0} = \left[\mathrm{h}_S^{\mathrm{T}}(\mathrm{q}_{S0}, \mathrm{q}_C^*)\mathbf{R}^{-1}\mathrm{h}_S(\mathrm{q}_{S0}, \mathrm{q}_C^*)\right]^{-1}\mathrm{h}_S^{\mathrm{T}}(\mathrm{q}_{S0}, \mathrm{q}_C^*)\mathbf{R}^{-1}\mathrm{h}_C(\mathrm{q}_{S0}, \mathrm{q}_C^*)\Delta\mathrm{q}_C. \tag{8}
$$

2.4 Variance of the estimator

The covariance matrix S_S of the vector $\hat{\mathrm{q}}_S$ may by found on the basis of the expression for the expected value of the estimator

$$\mathrm{S}_S = \mathrm{E}\left\{\left(\hat{\mathrm{q}}_S - \mathrm{E}\{\hat{\mathrm{q}}_S\}\right)\left(\hat{\mathrm{q}}_S - \mathrm{E}\{\hat{\mathrm{q}}_S\}\right)^{\mathrm{T}}\right\} = \left[\mathrm{h}_S^{\mathrm{T}}\left(\mathrm{q}_{S0}, \mathrm{q}_C^*\right)\mathbf{R}^{-1}\mathrm{h}_S\left(\mathrm{q}_{S0}, \mathrm{q}_C^*\right)\right]^{-1}. \tag{9}$$

2.5 Total error of estimation

An assessment of the model quality in terms of accuracy of the selected parameter estimates should take into account both the systematic and random errors introduced by the proposed method of model identification.

The relative systematic error b_r of the parameter estimate equals

$$b_r\left(\hat{q}_k\right) \doteq \frac{b_S\left(\hat{q}_k\right)}{q_k} \tag{10}$$

and the total relative systematic error of the whole parameter vector estimate:

$$m_b \doteq \sqrt{\frac{1}{s}\sum_{k=1}^{s} b_r^2\left(\hat{q}_k\right)} = s^{-1/2}\left(\mathbf{b}_S^{\mathrm{T}}\mathrm{Q}_S^{-1}\mathrm{Q}_S^{-1}\mathbf{b}_S\right)^{1/2}, \tag{11}$$

where $\mathrm{Q}_S = \mathrm{diag}\left(\mathrm{q}_S\right)$. The relative random error d_r of the parameter estimate has been defined as

$$d_r\left(\hat{q}_k\right) \doteq \frac{s_S\left(\hat{q}_k\right)}{q_k}, \tag{12}$$

where $s_S^2\left(\hat{q}_k\right)$ is the *k*-th parameter estimate variance, and then the total relative random error is

$$m_s \doteq \sqrt{\frac{1}{s}\sum_{k=1}^{s} d_r^2\left(\hat{q}_k\right)} = s^{-1/2}\left[\mathrm{tr}\left(\mathrm{Q}_S^{-1}\mathrm{S}_S\mathrm{Q}_S^{-1}\right)\right]^{1/2}. \tag{13}$$

Finally, the total relative mean squared error of the estimates can be calculated by use of the following formula:

$$\Delta_S \doteq \sqrt{m_b^2 + m_s^2} = s^{-1/2}\left[\mathbf{b}_S^{\mathrm{T}}\mathrm{Q}_S^{-1}\mathrm{Q}_S^{-1}\mathbf{b}_S + \mathrm{tr}\left(\mathrm{Q}_S^{-1}\mathrm{S}_S\mathrm{Q}_S^{-1}\right)\right]^{1/2}. \tag{14}$$

3 SIMULATION EXAMPLE

Simulation studies have been performed with the reduced model for maximum expiration including the flow limiting mechanism [8, 9]. It describes the process of maximally strong and quick expiration of air from the human lung. The model is discrete, non-linear and it has 12 degrees of freedom. The conditions of slight mixed restrictive and obstructive disorders of the respiratory system have been chosen, characterised by vital capacity (VC) of 3.5 litters, forced expiratory volume in the first second ratio to VC of 70%, peak expiratory flow of 4.2 l/sec and airway resistance of 0.47 kPa·s/l (see Fig. 1A).

The sensitivity matrix h has been calculated numerically in the similar way to this reported in [9] by model simulation with parameters changed by $\pm\Delta q$. Then the relative sensitivities have been computed (Fig. 1B) and the model parameters have been sorted beginning from the most to the least important one according to the Euclidean norms of the sensitivity vectors.

Since the values of the not estimated parameters q_C are not generally known, they must be fixed during model identification. Evaluating margins of the possible shift in a given parameter value in the real system (equivalent to pathological changes in the lung) and taking q_C^* in the middle of such limits, the worst situation is when $\Delta\mathrm{q}_C$ is a half of the parameter range. This situation has been taken into account during simulations.

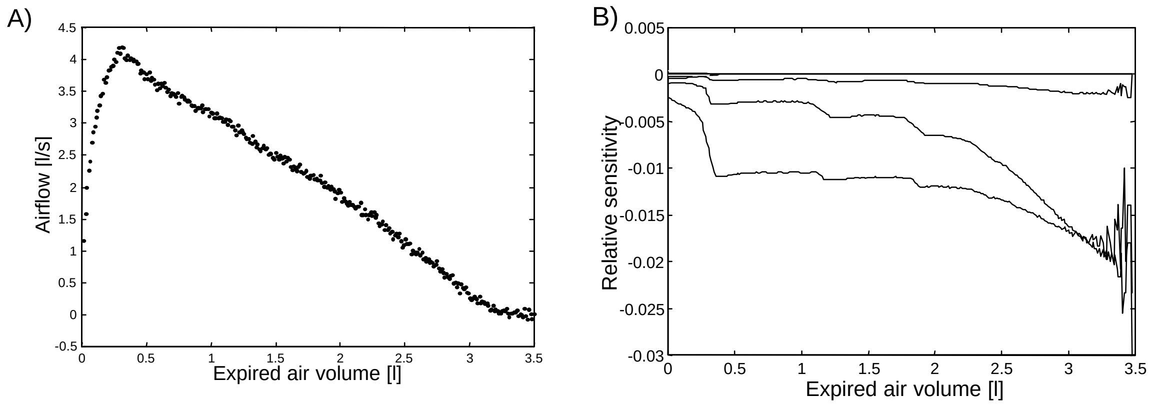

Figure 1. Simulated forced expiration curve with additive noise (A) and relative sensitivities of the model for maximum expiration (B).

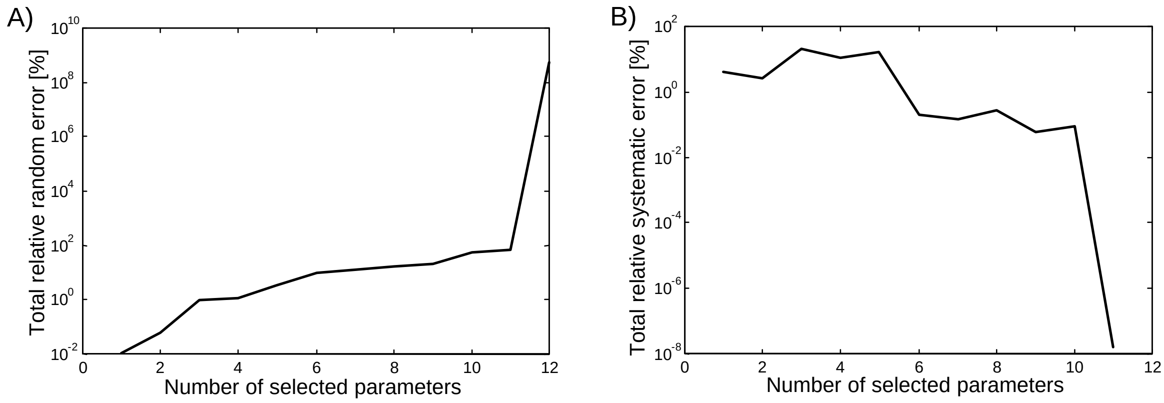

Figure 2. Total relative random (A) and systematic (B) error dependencies on the number of selected parameters.

The total relative systematic error m_b and the total relative random error m_s dependencies on the number of selected parameters are shown in Fig. 2A and 2B, respectively, and analogous dependency of the computed total relative mean squared errors Δ_S is demonstrated in Figure 3.

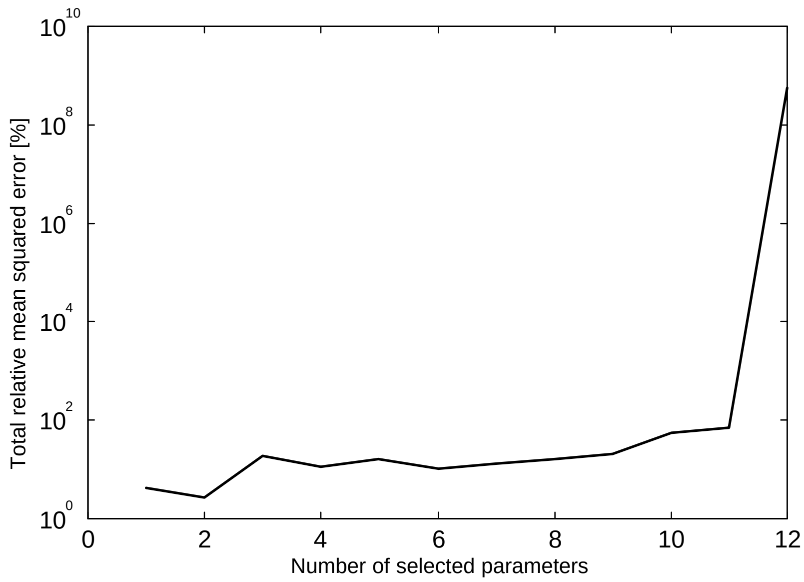

Figure 3. Total relative mean squared error dependence on the number of selected parameters.

Identification of selected parameters forces inversion of the matrix $\left[\mathrm{h}_S^{\mathrm{T}}\left(\mathrm{q}_{S0},\mathrm{q}_C^*\right)\mathbf{R}^{-1}\mathrm{h}_S\left(\mathrm{q}_{S0},\mathrm{q}_C^*\right)\right]$. This process may be ill-conditioned that leads to unforeseeable errors of numerical computations. Performed simulations have proved that numerical conditioning depends on the number of estimated parameters (see Tab. 1).

Table 1. Dependence of the condition number, defined as the ratio of the largest singular value of the matrix to be inverted to the smallest one, on the number of selected parameters.

Number of selected parameters	1	2	3	4	5	6	7	8	9	10	11	12
Condition number	1.00	2.50×10^{2}	3.81×10^{11}	7.86×10^{11}	8.13×10^{21}	8.66×10^{21}	6.07×10^{22}	2.76×10^{29}	3.21×10^{29}	3.30×10^{29}	7.57×10^{29}	6.95×10^{30}

4 CONCLUSIONS

An alternative method of dealing with ill-conditioned problems in model identification has been introduced. The main idea is to estimate only the parameters selected on the ground of the sensitivity analysis and fixing remaining ones as constants. Statistical analysis of the resulting estimator shows that the parameter selection reduces its variance (see also Fig 2A). Simultaneously the estimator bias is changing with the number of selected parameters. This relationship is not so trivial as the previous one, since the bias depends both on the covariance matrix and the vector $\Delta\mathrm{q}_C$ of imprecisely chosen values for the „constant" parameters (compare Eq. (8) with Eq(9)). The effect of this co-influence can be seen in Fig. 2B as non-monotonic alternation of the total relative systematic error of estimation. The total relative mean squared error (Fig. 3) can serve as a criterion for choosing the number of estimated parameters. For example, one may expect the least total error while estimating only two (mechanical properties of the airways) of twelve parameters in the case of maximum expiration model, although the sensitivity analysis has shown importance also of the third one (elasticity of lung tissues) (see Fig. 1B).

The realistic simulation example indicates that the parameter selection alone may be an inefficient method of error reduction in indirect measurements. Though it improves numerical conditioning of estimation (Tab. 1) and considerably decreases random errors in some instances (Fig. 2A), the total error can be still large (see Fig. 3). What is worth checking in the future is combining the parameter selection with a regularisation procedure. It is possible that these two methods together will give much better results than they give separately. Another problem is „dynamic" selection of estimated parameters during the identification process. In each iteration not only the temporal value of the parameter vector is changed but also the sensitivity matrix (NLP model), which is used for parameter selection. One can try to elaborate an adaptive method of determining the number of parameters estimated in the next step of identification in order to overcome this problem.

REFERENCES

[1] L. Ljung, T. Glad, *Modeling of Dynamic Systems*, Prentica Hall, Englewood Cliffs, 1994.
[2] T.A. Johansen, On Tikhonov regularization, bias and variance in nonlinear system identification. *Automatica* **33** (3) (1997) 441-446.
[3] A.E. Hoerl, R.W. Kennard, Ridge regression: biased estimation for non-orthogonal problems. *Technometrics* **12** (1970) 55-67.
[4] D.W. Marquardt, Generalized inverses, ridge regression, biased linear and non-linear estimation, *Technometrics* **12** (1970) 591-612.
[5] A.N. Tikhonov, V.Y. Arsenin, *Solutions of Ill-posed Problems*, Winston, Washington DC, 1977.
[6] R.R. Hocking, The analysis and selection of variables in linear regression, *Biometrics* **32** (1976) 1-49.
[7] I.S. Helland, Partial least squares regression and statistical models, *Scand. J. Statis*. **17** (1990) 97-114.
[8] A.G. Polak, A forward model for maximum expiration, *Comp. Biol. Med*. **28** (6) (1998) 613-625.
[9] A.G. Polak, J. Mroczka, A metrological model for maximum expiration. *Measurement* **23** (4) (1998) 265-270.

AUTHORS: A.G. POLAK and J. MROCZKA, Chair of Electronic and Photonic Metrology, Department of Electronics, Wroclaw University of Technology, ul. B. Prusa 53/55, 50-317 Wroclaw, Poland, Phone: +48 71 3206581, Fax: +48 71 3214277, E-mail: polak@kmeif.pwr.wroc.pl