Transactions of the 15th International Conference on
Structural Mechanics in Reactor Technology (SMiRT-15)
Seoul, Korea, August 15-20, 1999

A Tikhonov Regularization Scheme for System Identification of Continuous Structures

Hyunwoo Park[1), Soobong Shin[2) and Haesung Lee[1)

1) Seoul National University, Korea
2) Dong- A University, Korea

ABSTRACT

This paper presents a system identification scheme with a regularization technique for continuous structures. The Tikhonov regularization is employed to overcome the instabilities caused by ill-posedness of nonlinear inverse problems. The singular values decomposed from the sensitivity matrix of responses are utilized to investigate the characteristics of the nonlinear inverse problem and the role of the regularization function. To determine regularization factors, three different schemes - the L-curve method (LCM), the generalized cross validation (GCV), and the variable regularization factor scheme (VRFS) - are employed. The regularization effect of the LCM, the GCV and the VRFS is presented through two numerical examples.

1. INTRODUCTION

System identification (SI) methods have been widely used for the last few decades in the area of structural engineering to refine an analytical model for a structure and to detect damage in a structural system. However, most of the applications have been limited to discrete structures such as trusses or frames, and very limited works on continuous structures are available.

It is known that SI schemes based on the minimization of least squared errors between measured and analytically computed responses have suffered from inherent instabilities caused by the ill-posedness of inverse problems [1, 2]. In particular, when measured data are polluted with noise or when a finite element model used for the SI cannot represent the actual situation properly, the instabilities become very severe.

This paper shows that the results from the conventional SI using the output error estimator (OEE) may be unstable due to noise in measurement data by means of the singular value decomposition (SVD). To overcome the instabilities of SI methods, the Tikhonov regularization scheme [2], in which a positive definite regularization function is added to the OEE, is employed. Suitable regularization factors have to be determined in order to obtain a physically meaningful and numerically stable solution from a SI scheme with the regularization.

Well-established schemes such as the L-curve method (LCM) [3] and the generalized cross validation (GCV) [4] have been proposed for linear inverse problems to obtain well-balanced regularization factors. For nonlinear inverse problems, however, rigorous schemes to determine regularization factors are rarely available. Lee *et al.* present a simple but very effective scheme referred to as the variable regularization factor scheme (VRFS) for shape identification problems. In this paper, an iterative method to utilize the LCM and GCV in nonlinear inverse problems is proposed. The performance of the VRFS is also investigated when a Tikhonov-type regularization function is employed. Numerical simulation studies using the LCM, the GCV, and the VRFS are presented and discussed.

2. SYSTEM IDENTIFICATION WITH REGULARIZATION

2.1 System Identification Scheme

The unknown system parameters of the finite body are identified by minimizing least squared errors between calculated and measured displacements at some discrete observation points located on the exterior boundary of the finite body.

$$\underset{\mathbf{x}}{\text{Minmize}}\, \Pi_E = \frac{1}{2}\sum_{i=1}^{nlc}\left\|\mathbf{u}_i^c(\mathbf{x}) - \mathbf{u}_i^m\right\|^2 \quad \text{subject to } \mathbf{R}(\mathbf{x}) \le 0 \tag{1}$$

where $\mathbf{u}_i^c$, $\mathbf{u}_i^m$, $\mathbf{R}$ and nlc are calculated displacement vector by the analytical model, measured displacement vector at observation points for the i-th load case, constraint vector for the system parameters, and the number of load cases, respectively, and $\|\cdot\|$ represents the Euclidean norm of a vector. It is assumed that elastic material properties of some part of the finite body are different from original, known material properties, which are referred as baseline values. The baseline values represent *a priori* information on the given finite body.

The error function defined in Eq. (1) is normalized by the Euclidean norm of the measured displacement vectors of all load cases, and system parameters and constraints are normalized with respect to the corresponding baseline value.

$$\underset{\xi}{\text{Minmize}}\, \pi_E = \frac{1}{2}\left\|\widetilde{\mathbf{U}}(\xi) - \overline{\mathbf{U}}\right\|^2 \quad \text{subject to } \Xi(\xi) \le 0 \tag{2}$$

where $\widetilde{\mathbf{U}}$ and $\overline{\mathbf{U}}$ are vectors obtained by arranging the vectors of the normalized calculated displacements and the normalized measured displacements for each load case in a row. ξ and Ξ are normalized system parameters and constraints with respect to the baseline values respectively.

The solution of the minimization problem (2) is obtained by solving the following quadratic subproblem iteratively.

$$\underset{\Delta\xi}{\text{Minmize}}\, \Delta\xi^T \mathbf{S}_{k-1}^T\, \mathbf{U}_{k-1}^r + \frac{1}{2}\Delta\xi^T \mathbf{H}_{k-1}\Delta\xi \quad \text{subject to } \Xi(\xi_{k-1} + \Delta\xi) \le 0 \tag{3}$$

where subscript k denotes the iteration count, and $\mathbf{S}_{k-1}$ and $\mathbf{H}_{k-1}$ are the sensitivity matrix of $\widetilde{\mathbf{U}}_{k-1}$ and the Hessian matrix of the error function with respect to the normalized system parameters at the previous iteration, respectively. The displacement residual $\mathbf{U}_{k-1}^r$ is defined as $\mathbf{U}_{k-1}^r = \widetilde{\mathbf{U}}_{k-1} - \overline{\mathbf{U}}$ at the previous iteration, and $\Delta\xi$ is the increment of normalized system parameters at the current iteration step. The Hessian matrix in Eq. (3) is approximated by the Gauss-Newton Hessian to avoid computational complexity.

To simplify the expressions, the subscript representing the iteration count in the incremental formulation presented hereafter is omitted. To limit our attention to the behaviors of the minimization problem (3), it is assumed that all constraints are inactive. The first-order necessary optimality condition for Eq. (3) is given by the following linear equation.

$$\mathbf{S}^T\mathbf{U}^r + \mathbf{S}^T\mathbf{S}\Delta\xi = \mathbf{0} \tag{4}$$

By the singular value decomposition (SVD), the $m\times n$ sensitivity matrix **S** with $m\geq n$ can be written as a product of an $m\times n$ orthogonal matrix **Z**, an $n\times n$ diagonal matrix **Ω**, and the transpose of an $n\times n$ orthogonal matrix **V**. In the definition, m is the total number of measured degrees of freedom for all the applied loads and n is the number of system parameters.

$$\mathbf{S} = \mathbf{Z\Omega V}^T \tag{5}$$

where $\mathbf{\Omega} = diag(\omega_j)$ in which ω_j is a singular value of **S** which has the descending order of $\omega_n = \omega_{\max} \geq \ldots \geq \omega_2 \geq \omega_1 = \omega_{\min} \geq 0$. This paper assumes that the sensitivity matrix always possesses sufficient ranks. Substituting Eq.(5) into Eq. (4), the solution of Eq. (4) can be represented as the following equation.

$$\Delta\xi = -\mathbf{V}diag(\frac{1}{\omega_j})\mathbf{Z}^T\mathbf{U}^r \tag{6}$$

The measured displacement can be theoretically decomposed into the noise-free displacement $\overline{\mathbf{U}}^f$ and the error vector **e** as follows.

$$\overline{\mathbf{U}} = \overline{\mathbf{U}}^f + \mathbf{e} \tag{7}$$

This decomposition cannot be achieved explicitly for actual displacements, and is purely conceptual. However, the decomposition of the displacement vector is very useful to explain and study the effect of errors on solutions of nonlinear inverse problems.

There are two sources of errors in applying a SI method, *i.e.*, measurement error and modeling error. The former represents noise caused by sensitivity of sensors or misreading of test equipment during actual measurements. The latter occurs due to the discrepancy between a real structure and its mathematical model employed in the SI. Substitution of Eq. (7) into Eq. (6) leads to the following expression.

$$\Delta\xi = -\mathbf{V}diag(\frac{1}{\omega_j})\mathbf{Z}^T(\tilde{\mathbf{U}} - \overline{\mathbf{U}}^f) + \mathbf{V}diag(\frac{1}{\omega_j})\mathbf{Z}^T\mathbf{e} \tag{8}$$

The first and the second term on the right-hand-side of Eq. (8) represent the solution increment caused by the displacement residual between the computed solution and noise-free displacement and the disturbed component of the solution increment caused by noise, respectively. The components of $\mathbf{Z}^T\mathbf{e}$ associated with small singular values amplify the deviation of the solution caused by the errors. Since the deviation occurred in each iteration accumulates through optimization iterations, the converged solution for noise-polluted measurements may be quite different from the noise-free solution. Furthermore, the solution is likely to lose physical significance due to the accumulation of solution components amplified by physically meaningless noise. A small change in noise may yield a totally different solution because small singular values amplify the change in measurements, which is a source of the discontinuity characteristics of the SI problems.

2.2 Regularization Scheme

The regularization can be interpreted as a process of mixing the *a priori* information into the *a posteriori* solution. The *a priori* information is taken into account in the problem statement of inverse problems by adding a regularization function that contains the *a priori* information. The following regularization function proposed by Tikhonov is used in the current study.

$$\Pi^{R} = \frac{1}{2}\lambda^{2}\left\|\mathbf{x}-\mathbf{x}_{0}\right\|^{2} \tag{9}$$

where λ and $\mathbf{x}_0$ denote the regularization factor and the baseline properties of a structure, respectively. By adding the normalized regularization function with respect to the baseline properties to the minimization problem of Eq. (2), the regularized system identification problem is written in the normalized form.

$$\underset{\xi}{\text{Minimize}}\ \pi = \frac{1}{2}\left\|\tilde{\mathbf{U}}(\xi)-\overline{\mathbf{U}}\right\|^{2} + \frac{1}{2}\lambda^{2}\left\|\xi-\mathbf{1}\right\|^{2} \quad \text{subject to } \Xi(\xi)\le 0 \tag{10}$$

where **1** denotes a column vector which has unit values in all the components. The regularization factor determines the effect of regularization on the system identification, *i.e.*, the influence of the *a priori* information on the solution of Eq. (10). The quadratic subproblem of Eq. (10) is defined as

$$\begin{aligned}&\underset{\Delta\xi}{\text{Minmize}}\left[\Delta\xi^{T}\mathbf{S}^{T}\mathbf{U}^{r}+\frac{1}{2}\Delta\xi^{T}\mathbf{S}^{T}\mathbf{S}\Delta\xi\right]+\lambda^{2}\left[\Delta\xi^{T}(\xi-\mathbf{1})+\frac{1}{2}\Delta\xi^{T}\Delta\xi\right]\\&\text{subject to } \Xi(\xi+\Delta\xi)\le 0\end{aligned} \tag{11}$$

The incremental solution of Eq. (11) is obtained by use of the SVD.

$$\Delta\xi^{R} = -\mathbf{V}diag\left(\frac{\omega_{j}^{2}}{\omega_{j}^{2}+\lambda^{2}}\frac{1}{\omega_{j}}\right)\mathbf{Z}^{T}\mathbf{U}^{r} - \mathbf{V}diag\left(\frac{\lambda^{2}}{\omega_{j}^{2}+\lambda^{2}}\right)\mathbf{V}^{T}(\xi-\mathbf{1}) \tag{12}$$

where $\Delta\xi^{R}$ is the regularized solution for the increments of the system parameters at the current iteration. As the regularization factor λ increases, the first term of Eq.(12) vanishes, and the *a priori* information governs the regularized solution (12). On the other hand, as the regularization factor decreases, the second term of Eq.(12) vanishes, and the *a posteriori* solution governs the regularized solution (12).

By decomposing the measured displacements into the noise-free components and error components using Eq. (7), the following expression is obtained.

$$\Delta\xi^{\kappa} = \left\{ -\mathbf{V}diag\left(\frac{\omega_j^2}{\omega_j^2+\lambda^2}\frac{1}{\omega_j}\right)\mathbf{Z}^T\mathbf{U}^r - \mathbf{V}diag\left(\frac{\lambda^2}{\omega_j^2+\lambda^2}\right)\mathbf{V}^T(\xi-\mathbf{1}) \right\} - \mathbf{V}diag\left(\frac{\omega_j^2}{\omega_j^2+\lambda^2}\frac{1}{\omega_j}\right)\mathbf{Z}^T\mathbf{e} \tag{13}$$

The terms in the curly bracket and the last term in Eq. (13) represent the solution increment of Eq. (11) for noise-free measurements and solution increments contributed by errors in measurements, respectively. Components of $\mathbf{Z}^T\mathbf{e}$ associated with small singular values, which are responsible for the discontinuity and the deviation from the noise-free solution as explained before, are mostly suppressed in the solution as the regularization factor becomes large.

3. DETERMINATION OF AN OPTIMAL REGULARIZATION FACTOR

Two different schemes are possible in nonlinear inverse problems to determine the regularization factor: a fixed scheme and a variable scheme. In the fixed scheme, the regularization factor is fixed throughout iterations, and the optimal regularization factor is determined based on the converged solution of a nonlinear inverse problem. In the variable scheme, the optimal parameter is defined at each iteration, and varies throughout iterations.

In this paper, the variable schemes are employed to determine regularization factors at each iteration of a quadratic subproblem by applying one of the three following schemes; the L-curve method (LCM) proposed by Hansen [3], the generalized cross validation proposed by Wahba and coworkers [4], and the variable regularization factor scheme (VRFS) proposed by Lee and the coworkers [1].

3.1 The L-Curve Method

The L-curve is a log-log plot of the regularization function versus the error function for various regularization factors. Hansen showed for linear inverse problems that the plot formed a 'L' shaped curve, and that the optimal regularization factor corresponds to the sharp edge of the curve where the curvature of the curve becomes maximum [3]. For nonlinear inverse problems, the L-curve is defined at each iteration for the linearized error function. The curvature of the L-curve is given as

$$\kappa(\lambda) = \frac{\rho'\eta'' - \rho''\eta'}{((\rho')^2 + (\eta')^2)^{1.5}} \tag{14}$$

where ρ and η denote the linearized error function and the regularization function in a log scale, respectively, and the superscript $'$ denotes the differentiation of a variable with respect to λ. Since ρ and η are continuous functions of λ and expressed explicitly for λ, the derivatives in Eq. (14) are obtained analytically. The optimal regularization factor that yields the maximum curvature of the L-curve is calculated precisely by a one-dimensional line search.

3.2 Generalized Cross Validation

The GCV proposed by Wahba and coworkers [4] has been employed in numerous linear inverse problems successfully. The GCV is based on statistical consideration that a good value of the regularization parameter should predict missing data accurately. In the GCV, an optimal regularization factor λ is determined by the minimization of the following equation.

$$V(\lambda)=\frac{1}{n}\frac{\left\|\mathbf{I}-\mathbf{A}(\lambda)\mathbf{U}^r\right\|^2}{\left[\frac{1}{n}\mathrm{Trace}(\mathbf{I}-\mathbf{A}(\lambda))\right]^2} \tag{15}$$

where, $\mathbf{A}(\lambda)=\mathbf{S}(\mathbf{S}^T+n\lambda\mathbf{I})^{-1}\mathbf{S}^T\mathbf{U}^r$, and n is the number of the system parameters.

3.3 Variable Regularization Factor Scheme

Recently, the variable regularization factor scheme (VRFS) is proposed by Lee *et al.* for nonlinear inverse problems to identify shapes of inclusions in finite bodies [1]. The VRFS is based on an argument that the regularization function should be smaller than the error function to prevent the regularization function from dominating the optimization process.

In the VRFS, the regularization factor is defined by an inequality condition between the error function and the regularization function as follows.

$$\left\|\tilde{\mathbf{U}}(\xi)-\bar{\mathbf{U}}\right\|^2 \geq \lambda^2\left\|\xi-\mathbf{1}\right\|^2 \tag{16}$$

When the regularization function becomes larger than the error function by the solution of the current iteration, the regularization factor is reduced by multiplying a prescribed reduction factor β ranging from 0 to 1. Lee *et al.* [1] demonstrated that identification results are relatively insensitive to moderate values of the reduction factor around 0.1. The VRFS with β = 0.1 has been successfully applied to shape identification problems.[1] One of the advantages of the VRFS is that the VRFS method can be easily applied to any type of regularization functions.

4. NUMERICAL SIMULATION STUDY

Behaviors of the proposed scheme are investigated for the identification of an inclusion in a square plate when the measurement data are polluted with random noise. The Young's modulus of each element group is taken as the system parameter. To focus on discussion on the regularization scheme, it is assumed that element groups are predefined.

Fig. 1 illustrates geometry, boundary condition, and the applied traction. The inclusion is denoted by the shadowed region in the figure. The Young's modulus of the square plate is 210 GPa, which is the representative of steel. Two types of inclusions are identified; a soft inclusion (aluminum, E = 70 GPa) and a hard inclusion (tungsten, E = 380 GPa). Random noise in measurement data is assumed as 5% proportional noise for both the aluminum and the tungsten case. The predefined element groups are shown in Fig. 2, and each element group contains 4 elements. Observation points are depicted as solid circles in Fig. 2. The baseline values of all the system parameters are assumed as Young's modulus of steel. The initial values for the optimization of the system parameters are taken as the same as the baseline value. The reduction factor of the VRFS, β=0.1, is used for the simulation studies in the current paper.

Since optimal regularization factors calculated from the GCV are too small for both hard and soft inclusion cases, the estimated system parameters are identical with those from the SI without regularization. Therefore, results from the GCV will not be presented in this paper.

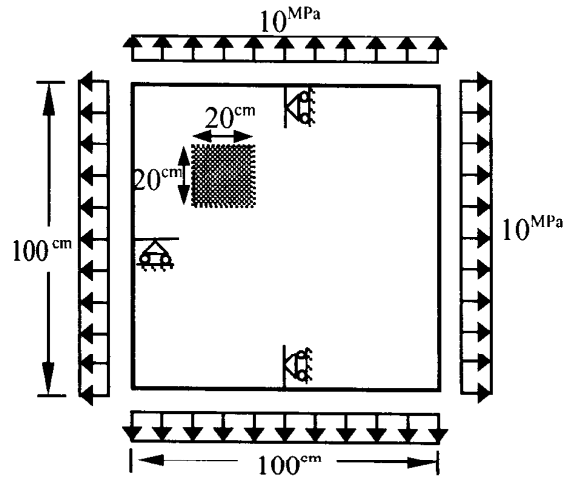

Fig. 1. Geometry and boundary conditions of a square plate

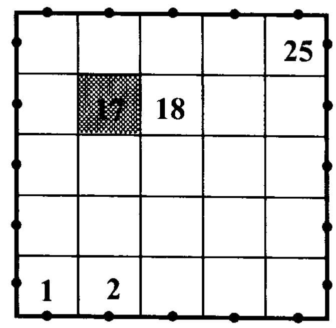

Fig. 2. Predefined group configuration of a square plate

The identified results for a soft inclusion case and a hard inclusion case are shown in Fig. 3, Fig. 4, respectively. For both cases, SI scheme without regularization leads to severely oscillatory results. In this case, it is very difficult to identify the existence of the inclusion because the reduction in the Young's modulus of element group 17 may be caused by an actual inclusion as well as by the oscillatory results. For soft cases, the LCM and the VRFS somewhat reduce the amplitude of oscillation for the element groups in the plate matrix. However, significant oscillations in the estimated results are observed. Since the Young's modulus of element group 17 is reduced prominently for a soft inclusion case compared with the oscillation magnitude of the other element groups, the existence of a soft inclusion may be identified. For a hard inclusion case, the LCM cannot yield the converged solutions, while the VRFS has difficulty to identify the hard inclusion successfully because it yields much lower Young's modulus compared with the actual value.

Regularization factors obtained by the different schemes at each iteration step are compared in Fig. 5 and Fig. 6. By relating the obtained regularization factors of Fig. 5 and Fig. 6

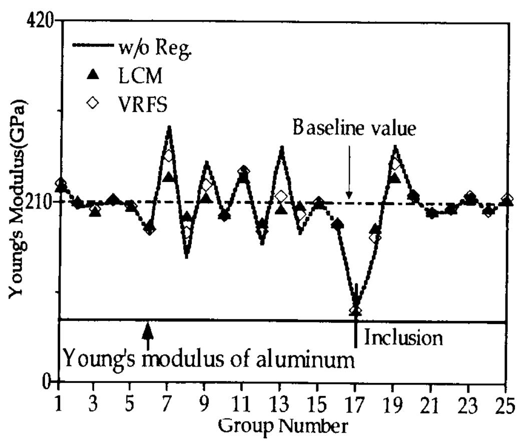

Fig. 3. Estimated Young's modulus - aluminum, 5% noise-

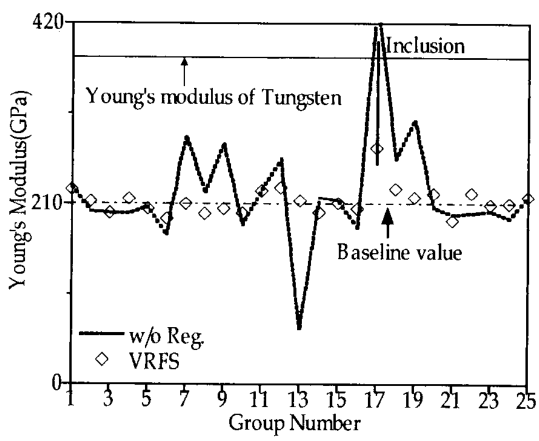

Fig. 4. Estimated Young's modulus - tungsten, 5% noise-

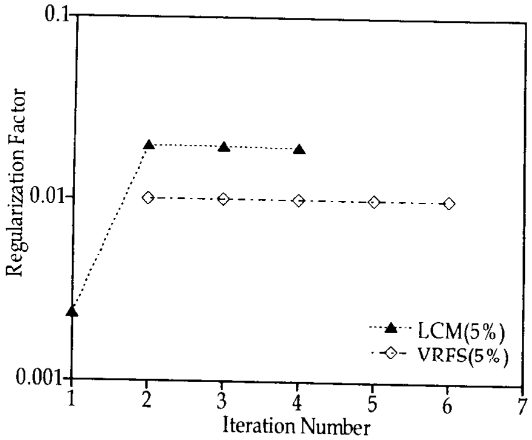

Fig. 5. Optimal regularization factors
- aluminum, 5% noise-

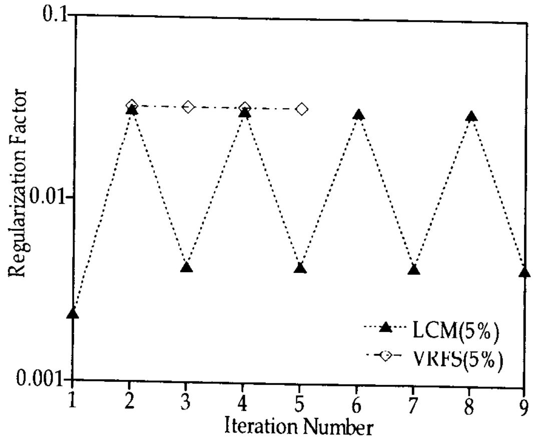

Fig. 6. Optimal regularization factors
- tungsten, 5% noise-

to the corresponding identified results in Fig. 3 and Fig. 4, it is easily observed that the identified results are greatly influenced by the magnitude of the regularization factor. For a hard inclusion case, the regularization factor by the LCM oscillates periodically at each iteration, which leads to divergence of the optimization iteration.

5. CONCLUSIONS

The effect of a regularization function is investigated through theoretical formulations combined with singular values decomposed from the sensitivity matrix of the response. It is shown that the error component associated with a small singular value can be suppressed in the incremental solution when the regularization function is added. The LCM, the GCV, and the VRFS that are applicable to nonlinear inverse problems, are reviewed and their theoretical backgrounds are discussed. From numerical simulation study for the identification of an inclusion in the square plate, it is revealed that more rigorous scheme to determine regularization factors is needed since the aforementioned methods yield too smeared results in some cases, and can not effectively control oscillations of estimated results in other cases.

REFERENCES

1. Lee, H. S., Kim, Y. H., Park, C. J., and Park, H. W.,. "A new spatial regularization scheme for the identification of geometric shapes of inclusions in finite bodies." to appear in *Inter. J. for Numerical Methods in Engineering*,1999
2. Groetsch, C. W., *The theory of Tikhonov regularization for Fredholm equations of the first kind*, Pitman Advanced Publishing, Boston, 1984
3. Hansen, P. C., "Analysis of discrete ill-posed problems by means of the L-curve", *SIAM review*, vol.34, no. 4, 1992, pp. 561-580
4. Golub, G. H., Heath, M., and Wahba, G., "Generalized cross-validation as a method for choosing a good ridge parameter", *Technometrics*, vol. 21, 1978, pp. 215-223