A HYBRID L-M/BFGS METHOD FOR SYSTEM IDENTIFICATION

T. Ting
University of Bridgeport
Bridgeport, CT 06601

INTRODUCTION

Structural system identification is concerned with the correlation of analysis and test results. The most practical way to improve such a correlation is to directly adjust structural parameters and, in turn, modify the mathematical model of the system. This is normally done by minimizing an error measure associated with the differences (or residue vector) between the two sets of results with respect to design variables. The most common methods employed for this purpose can be classified as least-squares methods, where the sum of squares of differences is treated as an objective function to be minimized. If the posed least squares problem is unconstrained and the residue is linear with respect to design variables, the solution to the problem can be obtained by simply solving the associated normal equations. However, the residue vector for structural problems are usually nonlinear, while the problems can always be cast into unconstrained ones. The problems should be handled by nonlinear least squares methods.

Nonlinear least squares methods require a number of iterations to converge. For a large-scale math model, the response analysis and the design sensitivity analysis are very costly, so an efficient method is sought to limit the number of function and sensitivity evaluations within one iteration and, at the same time, to posses a rapid convergence rate. Normally, a good approach is to avoid using exhaustive line search to find the corresponding design variable changes in an iteration. This is usually achieved by solving a well-conditioned normal equation in each iteration. Unfortunately, many investigators have indicated that the normal equations are frequently ill-conditioned and the resulting design changes are unacceptable or the process diverges. It has been demonstrated that this type of problem may be effectively overcome (in most cases) by adopting Lenvenberg-Marquardt (L-M) type modification to the normal equations. Speeding up the convergence rate remains a challenge to the researchers in this field.

Although significant progress has been made, it is clear that additional work is required to improve the robustness and efficiency of the present computational methods. One common way to improve the convergence rate is to use Newton-like second order methods. But the associated cost in evaluating the second derivatives and their unstable conditions at remote points from the solution should prohibit us to consider such methods. Instead, some well-developed quasi-Newton's methods, which require only the first-order derivatives and provide better than linear convergence rate, are worthy of investigating. This paper is to propose a hybrid L-M/BFGS algorithm for this purpose and a numerical comparison between this method and the original L-M method is given in the sequel.

THEORETICAL BACKGROUND

Newton's Method

Newton's method begins with the second-order Taylor's series, i.e.,

$$\Phi = \Phi_k + \{g\}_k^T \{\Delta x\}_k + \frac{1}{2}\{\Delta x\}_k^T [H]_k \{\Delta x\}_k$$

where

Φ_k = objective function evaluate at iteration k.

$\{\Delta x\}_k$ = design variable changes = $\{x\} - \{x\}_k$

$\{g\}_k$ = gradient vector at iteration k.

$$= \left\{ \frac{\partial \Phi_k}{\partial x_1} \quad \frac{\partial \Phi_k}{\partial x_2} \quad \cdots \quad \frac{\partial \Phi_k}{\partial x_3} \right\}$$

and $[H]_k$ = the corresponding Hessian matrix at iteration k.

$$= \begin{pmatrix} \frac{\partial^2 \Phi_k}{\partial x_1 \partial x_1} & \frac{\partial^2 \Phi_k}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 \Phi_k}{\partial x_1 \partial x_n} \\ \frac{\partial^2 \Phi_k}{\partial x_2 \partial x_1} & \frac{\partial^2 \Phi_k}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 \Phi_k}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 \Phi_k}{\partial x_n \partial x_1} & \frac{\partial^2 \Phi_k}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 \Phi_k}{\partial x_n \partial x_n} \end{pmatrix}$$

The necessary optimality condition for the minimization of Φ with respect to $\{x\}$ is:

$$\{g\}_k + [H]_k \{\Delta x\}_k = 0$$

Thus,

$$\{\Delta x\}_k = -[H]_k^{-1} \{g\}_k$$

Least Squares Solution to Correlation Problem and L-M Method

Let $\Phi = 1/2 \{\Delta y\}^T \{\Delta y\}$,

and $\{\Delta y\}$ = the difference between experimental and analysis results.

$= \{y_e\} - \{y_a\}$

Now, at the k-th iteration, Φ can be approximated as:

$$\Phi = \Phi_k + \{g\}_k^T \{\Delta x\}_k + \frac{1}{2} \{\Delta x\}_k^T [H]_k \{\Delta x\}_k$$

where

$$\{g\}_k = -[T]_k^T \{\Delta y\}_k$$
$$[H]_k = [T]_k^T [T]_k + \left(\sum_{i=1}^{m} \Delta y_i [H_i] \right)_k$$

and $[T]$ = Sensitivity matrix of $\{y\}_a$ with respect to $\{x\}$

$[H_i]$ = Hessian matrix of function Δy_i

If Newton's method is to be followed strictly, it is necessary to have Hessian matrices $[H_i]$ computed for all the Δy_i, i=1,2,...,m, and this can be impractical or costly.

However, since it is often the case that the components Δy_i are small, a good approximation to $[H]_k$ might be

$$[H]_k = [T]_k^T [T]_k$$

Thus, the conventional least squares solution to the correlation problem can be obtained as follows:

$$\{\Delta x\}_k = -[H]_k^{-1}\{g\}_k$$

$$= [T_T T]_k^{-1}[T]_k^T\{\Delta y\}_k$$

In many practical cases the matrix $[T^T T]$ is ill-conditioned and the consequence is that the solution is unreasonable. However, this problem can practically be resolved by applying Levenberg-Marquardt (L-M) method as shown below:

$$\{\Delta x\} = ([T^T T]_k + \nu[I])^{-1}[T]_k^T\{\Delta y\}_k$$

where ν is a reasonably small real value (e.g., typically, $\nu = 1.\text{x}10^{-8}$).

BFGS Method

BFGS method is commonly recognized as the most preferable method among all the quasi-Newton methods. The idea underlying quasi-Newton methods is to use an approximation to the inverse Hessian in place of the true inverse and the improved approximations are built up on the basis of information gathered during the iterative process.

Since the derivation of the BFGS formula for constructing the updated inverse Hessian is extremely laborious, only the formula itself is presented in the following.

At iteration k+1, let

$\{\Delta g\}_k = \{g\}_{k+1} - \{g\}_k$
$\sigma_k = \{\Delta g\}^T_k\{\Delta x\}_k$
$\{q\}_k = [H]^{-1}_k\{\Delta g\}_k$

Then

$$[H]_{k+1}^{-1} = [H]_k^{-1} + \frac{1}{\sigma_k}\left[\frac{1}{\sigma_k}(1 + \{\Delta g\}_k^T\{q\}_k)\{\Delta x\}_k\{\Delta x\}_k^T - \{\Delta x\}_k\{q\}_k^T + \{q\}_k\{\Delta x\}_k^T\right]$$

HYBRID L-M/BFGS ALGORITHM

The iterative procedure of the method can be stated as follows:

(1) Start with an initial FE model. Set iteration number k=0

(2) Perform SOL 63 for normal modes analysis to obtain $\{y_a\}_k$

(3) Use SOL 53 to compute design sensitivity matrix $[T]_k$

(4) Use L-M method to set

$$[H]_k^{-1} = ([T^T T]_k + \nu[I])^{-1}$$

and

$$\{\Delta x\}_k = -[H]_k^{-1}\{g\}_k$$

(5) Update FE model.

(6) Set k=k+1.

(7) Perform SOL 63 for normal modes analysis to obtain $\{y_a\}_k$

(8) Use SOL 53 to compute design sensitivity matrix $[T]_k$

(9) Use BFGS method to update $[H]^{-1}_k$ and set

$$\{\Delta x\}_k = -[H]_k^{-1}\{g\}_k$$

(10) Check for the convergence criterion.

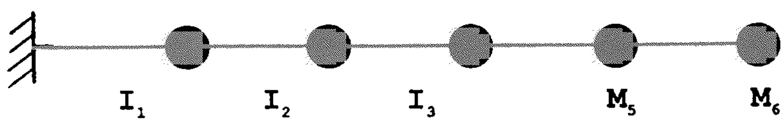
a. stop the procedure if it is met.

b. continue the procedure if it is not met.

(11) Set k=k+1 and go to step (7)

NUMERICAL EXAMPLE

A simple cantilever beam was used for demonstration purpose. The finite element model consists of 5 beam elements and lumped mass elements at nodes. Principal area moments of inertia in the vertical bending plane of the three close-to-wall elements, i.e., I_1, I_2, I_3, and two close-to-tip masses, i.e, M_5 and M_6, were designated as design variables. There were a total of five design variables and their baseline values as well as the perturbed values for generating mock test data are shown in Fig. 1.



BASELINE

$I_1 = I_2 = I_3 = 3.6$
$M_5 = M_6 = 5.0$
(E = 10.x106)*
($I_4 = I_5 = 3.0$)*
($M_2 = M_3 = M_4 = 5.0$)*
* values unchanged

PERTURBED

$I_1 = I_2 = I_3 = 3.0$
$M_5 = M_6 = 5.0$

Fig.1. Baseline and Perturbed Design Variables

Five natural frequencies corresponding to the lowest five modes were selected as correlation data. Table 1 lists the iteration history along with the target values of all the normalized design variables for the Hybrid L-M/BFGS method. Comparative result for the L-M method is shown in Table 2 for comparison purpose. The relative errors of the fundamental frequency verses iteration numbers are depicted in Fig. 2 to demonstrate the convergence pattern of both methods.

TABLE 1

L-M/BFGS Correlation Results

DESIGN VAR.	INIT. VALUE	ITERATION 0	1	2	3	4	5	TARGET VALUE
1 (I_1)	1.	.676	.747	.834	.828	.835	.834	.833
2 (I_2)	1.	.799	.833	.846	.836	.833	.833	.833
3 (I_3)	1.	.873	.849	.831	.834	.834	.833	.833
4 (M_5)	1.	1.003	1.142	1.290	1.248	1.250	1.250	1.250
5 (M_6)	1.	1.083	1.157	1.243	1.244	1.252	1.251	1.250

TABLE 2

L-M Correlation Results

DESIGN VAR.	INIT. VALUE	ITERATION 0	1	2	3	TARGET VALUE
1 (I_1)	1.	.676	.823	.833	.833	.833
2 (I_2)	1.	.799	.841	.833	.833	.833
3 (I_3)	1.	.873	.832	.833	.833	.833
4 (M_5)	1.	1.003	1.274	1.250	1.250	1.250
5 (M_6)	1.	1.083	1.246	1.250	1.251	1.250

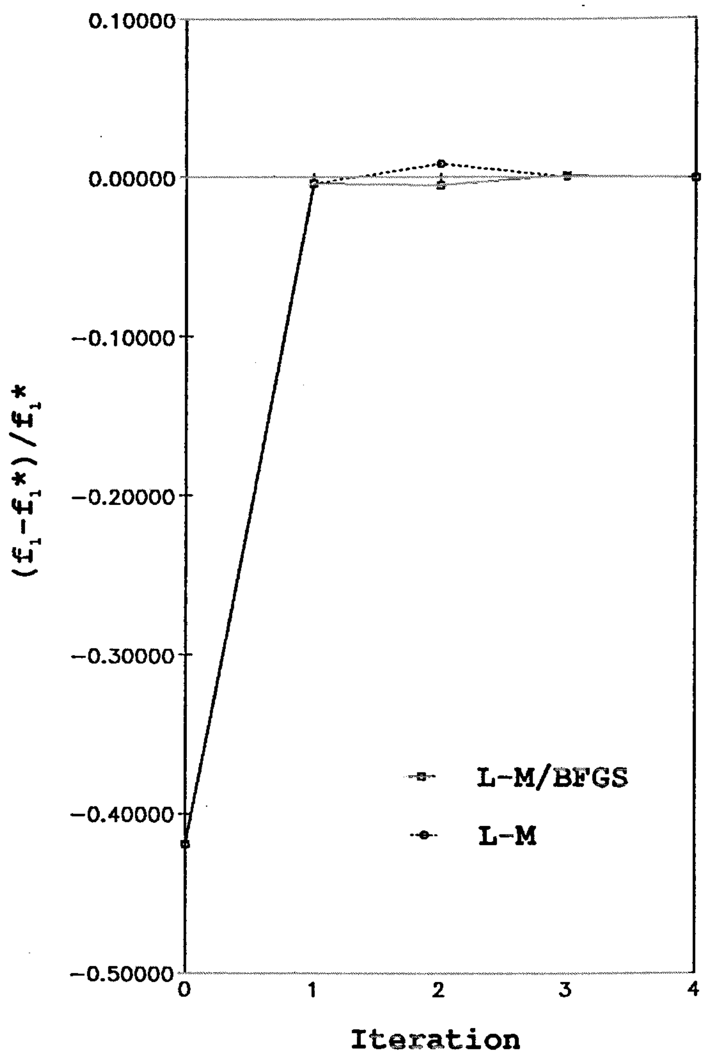


Fig.2. Relative Frequency Error Vs. Iteration

CONCLUDING REMARKS

A hybrid Levenberg-Marquardt/BFGS Algorithm has been developed and numerically demonstrated for solving nonlinear least squares problems and indeed there are many possibilities for such an algorithm. At this moment there is no clear advantage in using the newly developed method to substitute the L-M method for practical system identification. However, such a method reveals its potential in solving any class of problems for which L-M method is unsatisfactory.

Based on the numerical comparison of the two methods, we can conclude that the L-M method outperforms the hybrid L-M/BFGS method, especially, at near solution points. The L-M method rapidly converges to the exact solution in four iterations. Although L-M/BFGS method reaches the near optimal solution quite fast, after this point, it exhibits a relatively slower rate of convergence. Nevertheless, it has a more monotonic convergence pattern when compared to that of L-M method. This suggests that for highly nonlinear situations the hybrid L-M/BFGS method may be more reliable in approaching the optimum.