TEORIJSKA I PRIMENJENA MEHANIKA 20, pp 43–51, 1994, UDK 534

NONLINEAR INVERSE PROBLEM

H. R. Busby

(Received 28.07.1994.)

1. Introduction

As one constructs models to better reflect true physical behavior it becomes necessary to include nonlinear terms into the equations representing the models. In general nonlinear analysis cannot be encompassed in one general framework but must be discussed on a case by case basis. There are, however, a class of nonlinear models that can be analyzed with a common approach. These are models that are only slightly nonlinear. This technique has been applied to nonlinear inverse problems, using dynamic programming, by solving a sequence of linear problems such that each solution gets closer to the nonlinear one [1]. Newton's method of solving a set of nonlinear equations is an example of such an approach. The success of such methods depends on two factors, one is that the linearized portion of the model dominate the nonlinear terms. The other is to start with a good first solution. Usually one has sufficient knowledge of a system to help in the linearization and also to provide a good estimate of the solution.

Another technique for the nonlinear inverse problem is the method of successive approximations in policy space [2-4]. This technique is based upon gradients in which the approach to an optimal policy is by successive steps. The method of derivation employs the familiar concepts and techniques of dynamic programming. We essentially guess a presumably non-optimal decision sequence $\mathbf{z_k}$. By simple reasoning we derive a set of recurrence relations that can be used to evaluate the effect of a small change in the decision sequence. We then use this information about the effect of decision changes to generate a new, improved, sequence of decisions. The effect of changes in the new sequence are evaluated. This iterative process is continued until no further improvement is possible. Each successive solution we obtain will be feasible for the problem, but not optimal.

2. Formulation

Consider a system of n nonlinear difference equations

$$y_i(t+h) = g_i(y_1(t), y_2(t), \dots, y_n(t), z_1(t), z_2(t)) \quad i = 1, 2, \dots, n \tag{1}$$

where z_1 and z_2 are known forcing functions to be determined. Consider only two forces just to help get the pattern. The error criterion will be linear given by the expression

$$E_N = \sum_{j=1}^{N} (\mathbf{Q}\mathbf{y}_j - \mathbf{d}_j, \mathbf{Q}\mathbf{y}_j - \mathbf{d}_j) + (b\mathbf{z}_j, \mathbf{z}_j) \tag{2}$$

where $\mathbf{Q}$ is a matrix which relates the state variables to the measurements $\mathbf{d}$. Following [2] define

$$f_n(y_1, y_2, \ldots, y_n) = \text{The value of E starting at state } \mathbf{y}_n \text{ and using the guessed policy at } \mathbf{z}_j. \tag{3}$$

Applying dynamic programming gives

$$f_n(\mathbf{y}_n) = (\mathbf{Q}\mathbf{y}_n - \mathbf{d}_n, \mathbf{Q}\mathbf{y}_n - \mathbf{d}_n) + (b\mathbf{z}_n, \mathbf{z}_n) + f_{n+1}(\mathbf{g}_n) \tag{4}$$

which can also be written as

$$\begin{aligned} f_n(y_1(t), y_2(t), \ldots, y_n(t)) =& (\mathbf{Q}\mathbf{y}_n - \mathbf{d}_n, \mathbf{Q}\mathbf{y}_n - \mathbf{d}_n) + (b\mathbf{z}_n, \mathbf{z}_n) + \\ & + f_{n+1}(g_1, g_2, \ldots, g_n) \end{aligned} \tag{5}$$

where

$$\begin{aligned} g_1 &= g_1(y_1(t), y_2(t), \ldots, y_n(t), z_1(t), z_2(t)) \\ g_2 &= g_2(etc) \end{aligned}$$

To determine the first order effect of a change in the forcing terms at time $t(n)$ one needs to evaluate $\frac{\partial f_n}{\partial z_1}, \frac{\partial f_n}{\partial z_2}$. By partial differentiation of Eq.(5) with respect to z we have

$$\begin{bmatrix} \frac{\partial f_n}{\partial z_1} \\ \frac{\partial f_n}{\partial z_2} \end{bmatrix} = 2b \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}_n + \begin{bmatrix} \frac{\partial g_1}{\partial z_1} & \frac{\partial g_2}{\partial z_1} & \frac{\partial g_3}{\partial z_1} & \cdots & \frac{\partial g_n}{\partial z_1} \\ \frac{\partial g_1}{\partial z_2} & \frac{\partial g_2}{\partial z_2} & \frac{\partial g_3}{\partial z_2} & \cdots & \frac{\partial g_n}{\partial z_2} \end{bmatrix} \begin{bmatrix} \frac{\partial f_{n+1}}{\partial y_1} \\ \frac{\partial f_{n+1}}{\partial y_2} \\ \vdots \\ \frac{\partial f_{n+1}}{\partial y_n} \end{bmatrix} \tag{6}$$

and also that

$$\begin{bmatrix} \frac{\partial f_n}{\partial y_1} \\ \frac{\partial f_n}{\partial y_2} \\ \vdots \\ \frac{\partial f_n}{\partial y_n} \end{bmatrix} = 2\mathbf{Q}^T(\mathbf{Q}\mathbf{y}_n - \mathbf{d}_n) + \begin{bmatrix} \frac{\partial g_1}{\partial y_1} & \frac{\partial g_2}{\partial y_1} & \cdots & \frac{\partial g_n}{\partial y_1} \\ \frac{\partial g_1}{\partial y_2} & \frac{\partial g_2}{\partial y_2} & \cdots & \frac{\partial g_n}{\partial y_2} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial g_1}{\partial y_n} & \frac{\partial g_2}{\partial y_n} & \cdots & \frac{\partial g_n}{\partial y_n} \end{bmatrix} \begin{bmatrix} \frac{\partial f_{n+1}}{\partial y_1} \\ \frac{\partial f_{n+1}}{\partial y_2} \\ \vdots \\ \frac{\partial f_{n+1}}{\partial y_n} \end{bmatrix} \tag{7}$$

The iterations for each time step are

$$\mathbf{z}_n(\text{new}) = \mathbf{z}_n(\text{old}) + \delta z$$

Recall that each stage we have

$$\Delta f_n = \frac{\partial f_n}{\partial z_1}\Delta z_1 + \frac{\partial f_n}{\partial z_2}\Delta z_2$$

then choose

$$\Delta z_1 = k\,\frac{\partial f_n}{\partial z_1}, \quad \Delta z_2 = k\,\frac{\partial f_n}{\partial z_2} \tag{8}$$

$$\Delta f_n = k\left[\left(\frac{\partial f_n}{\partial z_1}\right)^2 + \left(\frac{\partial f_n}{\partial z_2}\right)^2\right] \tag{9}$$

k is determined from Eq.(9) by setting Δf_n to be an incremental decrease in the current f_n. This in turn determines Δz_1 and Δz_2. The initial conditions are applied at the end point N^*.

$$\left[\frac{\partial f_{N^*}}{\partial \mathbf{y}}\right] = 2\mathbf{Q}^T(\mathbf{Q}\mathbf{y}_{N^*} - \mathbf{d}_{N^*}) \tag{10}$$

Thus to obtain a solution we start with some reasonable guess to the forcing terms. Since the forcing terms are known Eq. (1) can be integrated using the known initial conditions and storing the **y**'s. Now Eq.'s (6) and (7) can be integrated backwards using the initial conditions, Eq.(10). At each step we compute k and $\Delta\mathbf{z}$ from Eq.'s (8) and (9). With a new forcing term we repeat the process and check for convergence.

3. Application to nonlinear systems

Let the differential equation be of the form

$$\dot{\mathbf{y}} + \mathbf{K}\mathbf{y} + \mathbf{n}(\mathbf{y}) = \mathbf{P}\mathbf{f}(t) \tag{11}$$

where all the linear terms have been grouped into $\mathbf{K}\mathbf{y}$ leaving the nonlinear terms in $\mathbf{n}(\mathbf{y})$. Expanding $\mathbf{n}(\mathbf{y})$ with a Taylor's series expansion about some state $\mathbf{y}_0$ gives

$$\mathbf{n}(\mathbf{y}) = \mathbf{n}(\mathbf{y}_0) + \mathbf{A}(\mathbf{y} - \mathbf{y}_0) \tag{12}$$

where $\mathbf{A}$ represents the Jacobian matrix of $\mathbf{n}$ evaluated at $\mathbf{y}_0$. This gives

$$\dot{\mathbf{y}} + (\mathbf{K} + \mathbf{A})\mathbf{y} = -(\mathbf{y}_0) + \mathbf{A}\mathbf{y}_0 + \mathbf{P}\mathbf{f}(t) \tag{13}$$

Using the Crank Nicolson method Eq. (13) can be put into discrete form as

$$[\mathbf{I} + (\mathbf{K} + \mathbf{A})h/2)]\,\mathbf{y}_{i+1} = [\mathbf{I} - (\mathbf{K} + \mathbf{A}h/2)]\,\mathbf{y}_i - h\mathbf{n}(\mathbf{y}_0) + h\mathbf{A}\mathbf{y}_0 + h\mathbf{P}\mathbf{f} \tag{14}$$

where h is the timestep. Now let $\mathbf{y}_0 = \mathbf{y}_i$, then we can write our equation as

$$\mathbf{y}_{i+1} = \mathbf{M}_i\mathbf{y}_i + \mathbf{g}_i + \mathbf{P}_i''\mathbf{f} \tag{15}$$

where

$$\mathbf{M}_i = [\mathbf{I} + (\mathbf{K} + \mathbf{A}_i)h/2)]^{-1} [\mathbf{I} - (\mathbf{K} + \mathbf{A}_i)h/2)]$$
$$\mathbf{g}_i = -[\mathbf{I} + (\mathbf{K} + \mathbf{A}_i)h/2)]^{-1} h(\mathbf{n}(\mathbf{x}_i) - \mathbf{A}_i\mathbf{x}_i) \tag{16}$$
$$\mathbf{P}''_i = [\mathbf{I} + (\mathbf{K} + \mathbf{A}_i)h/2)]^{-1} \mathbf{P}\mathbf{f}_i$$

It is important that these integration formula remain stable. It should be noted that as $h \Rightarrow 0$, Eq. (15) reduces to Eq. (11) and $\mathbf{A}_i$ is not involved. Here $\mathbf{A}_i$ will be evaluated as the Jacobian Matrix of $\mathbf{n}(\mathbf{y})$. With these definitions of $\mathbf{M}_i$, $\mathbf{g}_i$, and $\mathbf{P}''_i$ the previous equations based on successive approximation can be used.

4. Illustrative numerical example

Consider the example of a spring, mass, and dashpot system with a nonlinear spring. This example is taken from [5] where it was used in conjunction with a parameter estimation problem. The nonlinear equation is given by

$$\ddot{x} + 3\dot{x} + 2x + 0.5\,x^3 = f(x) \tag{17}$$

The initial conditions are $x(0) = 2.0$, and $\dot{x}(0) = 0.0$. The inverse problem we are interested in solving is one where measurements have been taken on the displacement x and we wish to estimate the forcing function term $f(t)$. The above equations can be represented in vector-matrix form as

$$\begin{bmatrix} \dot{x} \\ \dot{\nu} \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ \nu \end{bmatrix} + \begin{bmatrix} 0 \\ 0.5x^3 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} f(t) \tag{18}$$

This equation can be solved as outlined above. The unknown function is chosen as $5\sin(t)$ which was used in a direct integration method to generate simulated measurements of the displacement. The measurements were corrupted with a normally distributed noise with $\sigma = 0.10$. These data are shown in Figure 1 which also shows the estimated displacement after 75 iterations and Figure 2 which compares the forcing term. The original function $5\sin(t)$ is shown in the figure to show that the nonlinear inverse method produces a reasonable estimate. All of the inverse solutions were obtained using a smoothing parameter $b = 5.\times 10^{-3}$. Figure 3 demonstrates how the error is reduced after each iteration. However, for some problems the convergence rate may be slow and thus may out require the selection and adjustment of several convergence parameters.

One method for smoothing the unknown forcing function is that of first-order regularization. The idea is to regulate the first derivative of the forcing function instead of the forces themselves. This is done by adjoining the following differential equation to the dynamic model

$$\dot{z} = r \tag{19}$$

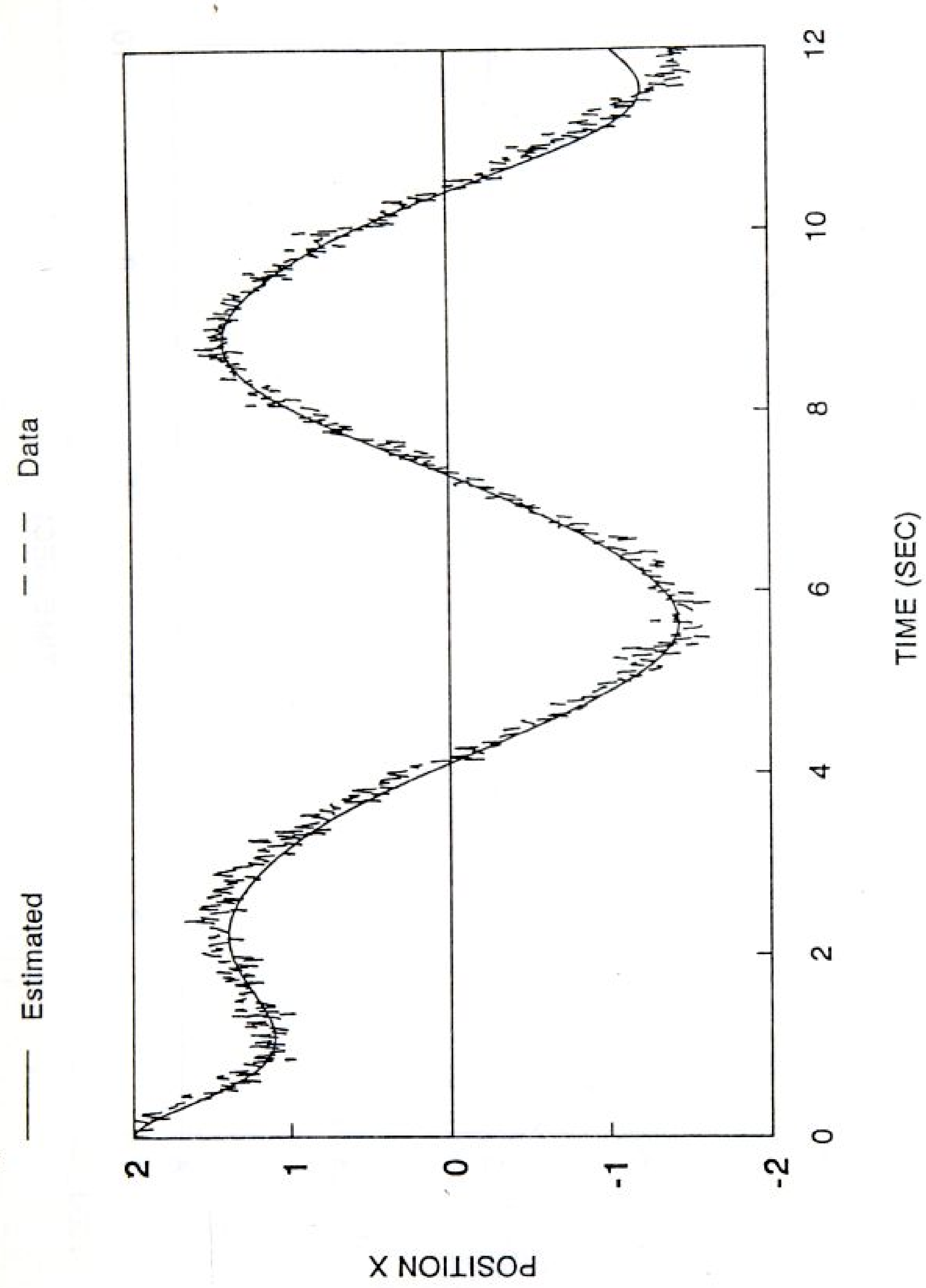

Figure 1. Comparison of Data and Estimated Displacement

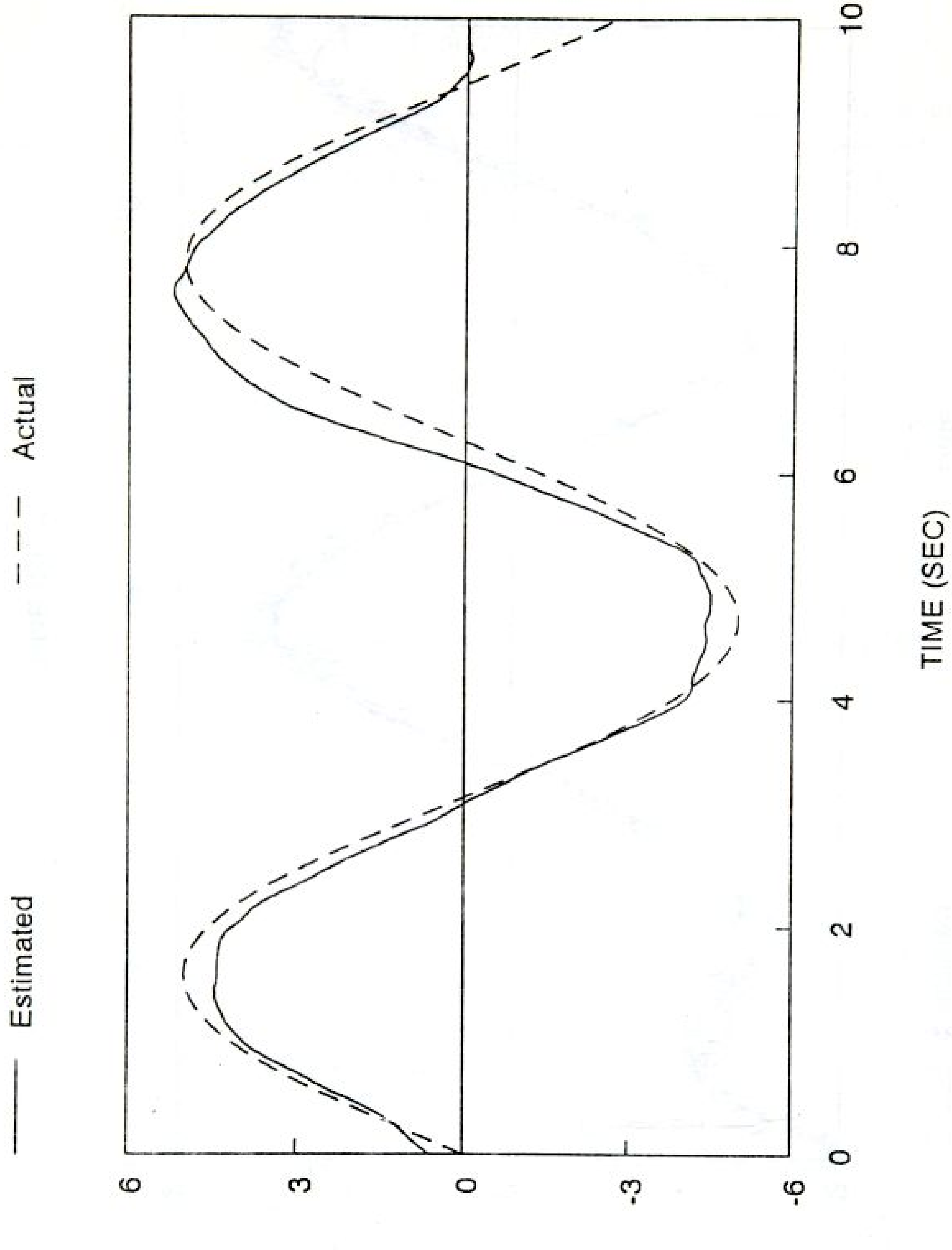

Figure 2. Comparison of Forcing Term

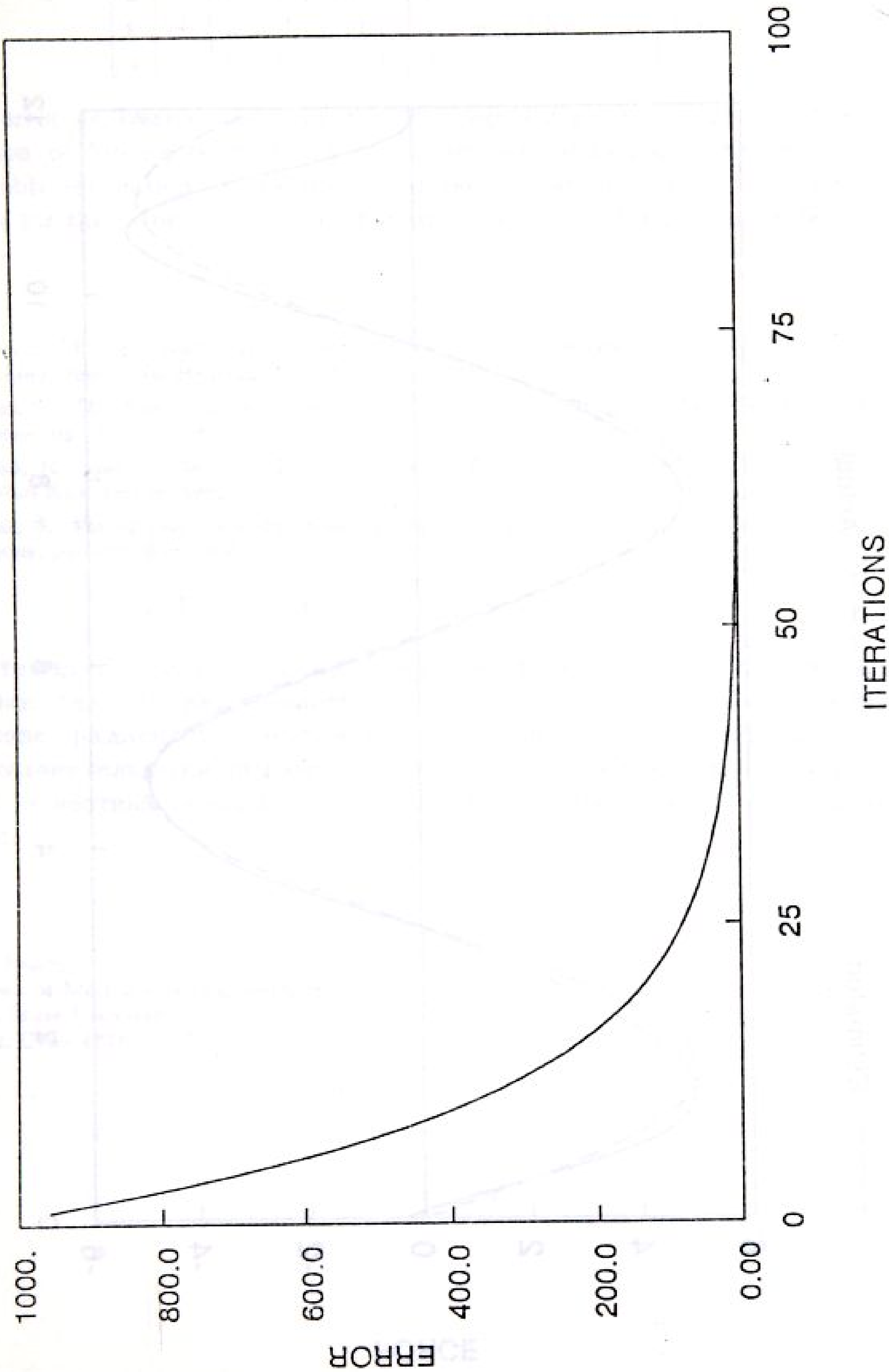

Figure 3. Least Squares Error Versus Number of Iterations

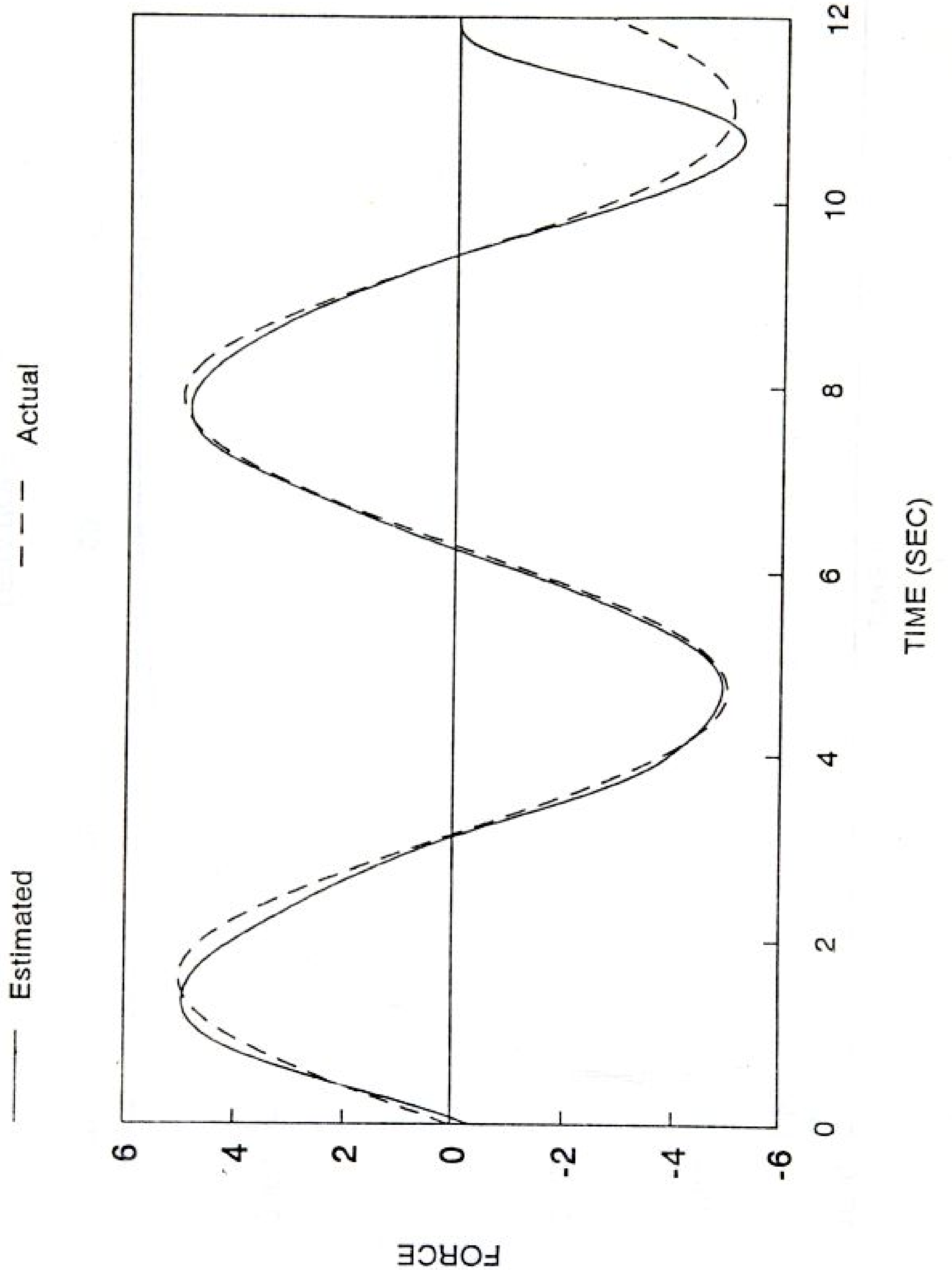

Figure 4. Comparison of Forcing Term Using First-Order Regularization

The new system becomes

$$\begin{bmatrix} \dot{x} \\ \dot{\nu} \\ \dot{z} \end{bmatrix} + \begin{bmatrix} 0 & -1 & 0 \\ 2 & 3 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ \nu \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ 0.5x^3 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ r \end{bmatrix} \tag{20}$$

and the error expression now uses r in the regulating term. Figure 4 shows the comparison of forcing terms using first-order regularization. The results show a reasonable estimation. Using first-order regularization however, requires more iterations for the error to converge. For this example 250 iterations were needed.

REFERENCES

[1] Trujillo, D. M., and Wallis, R., *Determination of Heat Transfer from Components During Quenching*, Industrial Heating, pp. 22-24, July 1989.

[2] Dreyfus, S., *The Numerical Solution of Variational Problems*, J. of Math. Analysis and Applications, pp. 30-45, (1962).

[3] Bellman, R., and Dreyfus, S., *Applied Dynamic Programming*, Princeton University Press, Princeton New Jersey, 1962.

[4] Dreyfus, S., *Variational Problems with Inequality Constraints*, J. of Math. Analysis and Applications, pp. 297-308, 1962.

НЕЛИНЕЫНАЯ ОБРАТНАЯ ЗАДАЧА

Метод последовательных приближенииы в пространстве стратегииы использован для решения нелинеыноы обратноы задачи. Метод основан на анализе градиентов; приближение к оптимальноы стратегии достигает ся последовательными шагами. Для обоснования метода используются хорошо известные концепции и подходы динамического программирования.

Henry R. Busby
Department of Mechanical Engineering
The Ohio State University
Columbus, Ohio 43210-1107
USA