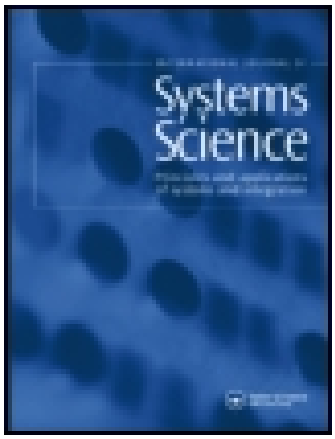



International Journal of Systems Science
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tsys20

# Gradient minimax techniques for system modelling

J. W. BANDLER [a], N. D. MARKETTOS [a] & T. V. SRINIVASAN [a]
[a] Department of Electrical Engineering and Communications Research Laboratory, McMaster University, Hamilton, Ontario, Canada
Published online: 25 Apr 2007.

To cite this article: J. W. BANDLER, N. D. MARKETTOS & T. V. SRINIVASAN (1973) Gradient minimax techniques for system modelling, International Journal of Systems Science, 4:3, 317-331, DOI: 10.1080/00207727308920018

To link to this article: http://dx.doi.org/10.1080/00207727308920018

INT. J. SYSTEMS SCI., 1973, VOL. 4, NO. 3, 317–331

# Gradient minimax techniques for system modelling

J. W. BANDLER, N. D. MARKETTOS
and T. V. SRINIVASAN

Department of Electrical Engineering and
Communications Research Laboratory,
McMaster University, Hamilton, Ontario, Canada

[Received 18 July 1972]

Some recently proposed gradient methods for minimax or near minimax approximation are applied to producing optimal second-order and third-order models of a high-order system. The Fletcher–Powell method, a more recent method by Fletcher and a method by Jacobson and Oksman are employed with least $p$th approximation, using large values of $p$, as proposed by Bandler and Charalambous and critically compared with the grazor search technique of minimax approximation by Bandler *et al.* The solutions obtained are shown to satisfy the necessary conditions for a minimax optimum.



## 1. Introduction

The purpose of determining low-order models for high-order systems is to simplify preliminary design and optimization of such systems. System modelling using least-squares approximation has been investigated recently by Bandler *et al.* (1972 a). They compared, in particular, the relative effeciencies of three gradient minimization methods as applied to least squares problems. The system modelled was a seventh-order system which represents the control system for the pitch rate of a supersonic transport aircraft.

In the present work, the maximum value of the error between the step responses of the above seventh-order system and the model is effectively minimized. This may be accomplished either by directly minimizing the maximum error (minimax), or by least $p$th approximation techniques which, by selecting a large enough value for $p$, gives, for practical purposes, a minimax solution. The direct minimax method, called the grazor search technique, has been recently proposed by Bandler, Srinivasan and Charalambous (1972 c). The strategy is based on steepest descent directions found by linear programming. The least $p$th approximation approach was based on a paper by Bandler and Charalambous (1971) where very large values of $p$—up to $10^{12}$—have been successfully used. For the present problem the value of $p$ was chosen as 1000, on the basis of acceptable almost minimax results and reasonable computer central processing time. A comparison is made between three gradient methods (Fletcher and Powell 1963, Fletcher 1970, Jacobson and Oksman 1970) a description of which is given by Bandler *et al.* (1972 a). The grazor search technique is also compared with the least $p$th approximation approach.

## 2. Statement of the problem

It is required to find a transfer function of a model of a given order, the

response of which is the best approximation to the response of the actual system to a particular input for a specified error criterion.

In general the transfer function of a given order $n$ may be written as

$$H(s)=\frac{b_m s^m+b_{m-1}s^{m-1}+\ldots+b_1 s+b_0}{s^n+a_{n-1}s^{n-1}+\ldots+a_1 s+a_0}, \tag{1}$$

where $m \leqslant n$ for physical systems. For this work the input is a unit step and the criterion chosen is to directly or indirectly minimize the maximum error over a specified time interval. The problem, therefore, is the determination of the parameter vector $\mathbf{x}$, defined as

$$\mathbf{x}=[a_0\ a_1\ \ldots\ a_{n-1}\ b_0\ b_1\ \ldots\ b_m]^{\mathrm{T}}, \tag{2}$$

such that the maximum error is minimized.

It is often desirable to impose constraints on the model. For instance, it will be logical to demand that the steady-state value of the model be equal to that of the system. So, a steady-state constraint will be imposed on all the problems in this work.

### 3. Least $p$th approximation

Least $p$th approximation is the minimization of the $p$th norm of the sampled errors defined as

$$F(\mathbf{x})=\left(\sum_{i\in I}|e_i(\mathbf{x})|^p\right)^{1/p}, \tag{3}$$

where $\mathbf{x}$ is a $k$-vector and $e_i(\mathbf{x})$ is the error between the system and model responses at some sample point $i$ of a finite set $I$, relating to all the sampled points. It is assumed that $e_i(\mathbf{x})$ is continuous with continuous partial derivatives with respect to $\mathbf{x}$ for all $i$. $p \geqslant 1$.

It is desirable to increase the value of $p$ as much as possible, since the larger the value of $p$ the nearer to minimax will the solution be. This is shown by

$$\lim_{p\to\infty} F(\mathbf{x})=M(\mathbf{x}), \tag{4}$$

where

$$M(\mathbf{x}) \triangleq \max_{i\in I}|e_i(\mathbf{x})|. \tag{5}$$

There are, however, two computational limitations. One is that if $|e_i(\mathbf{x})|>1$, when using large values of $p$, the numbers tend to become too large for the computer to handle. The other is that if $|e_i(\mathbf{x})|<1$, again when raised to a large power, the numbers tend to zero and most of the information is lost.

A normalization proposed by Bandler and Charalambous (1971), permitted the use of extremely large values of $p$. The objective function to be minimized may be rewritten as

$$F(\mathbf{x})=M(\mathbf{x})\left(\sum_{i\in I}\left|\frac{e_i(\mathbf{x})}{M(\mathbf{x})}\right|_p\right)^{1/p}, \tag{6}$$

in which case at least one of the numbers raised to $p$ will be equal to 1.

In (6), if the set $I$ is replaced by $J$, an index set relating only to the extrema of the error function, considerable economies in computing time will result at a slight risk of creating false optima. As can be seen from the expression of the gradient, namely,

$$\nabla F(\mathbf{x}) = \left( \sum_{i \in I} \left| \frac{e_i(\mathbf{x})}{M(\mathbf{x})} \right|_p \right)^{1/p-1} \left( \sum_{i \in I} \left| \frac{e_i(\mathbf{x})}{M(\mathbf{x})} \right|_{p-2} \frac{e_i(\mathbf{x})}{M(\mathbf{x})} \nabla e_i(\mathbf{x}) \right) \tag{7}$$

the coefficients of $\nabla e_i(\mathbf{x})$ will, for most points and large enough $p$, be very small, thus contributing very little to $\nabla F(\mathbf{x})$. The manner by which $\nabla e_i(\mathbf{x})$ is derived is indicated for the present examples in the paper by Bandler *et al.* (1972).

The following values of $p$ have been used : 10, $10^2$, $5 \times 10^2$, $10^3$, $10^4$, $10^6$, $10^9$ and $10^{12}$. It was found that, although agreement in significant figures amongst the extrema increased as $p$ increased, the computer central processing time increased considerably for values of $p$ above $10^3$. Thus for comparison of the minimization techniques $p = 10^3$ was considered suitable.

## 4. The grazor search method

A new algorithm called the grazor search method has been recently developed (Bandler *et al.* 1972 c) in which gradient information of one or more of the largest extrema in the error function is used to produce a downhill direction by solving a suitable linear programming problem. A linear search follows to find a minimum in that direction, and the procedure is repeated. This procedure is repeated with as many extrema as necessary until a minimax solution is reached to some desired accuracy. The method is guaranteed to converge under certain conditions.

Suppose we have the problem of minimizing

$$M(\mathbf{x}) = \max_{i \in I} y_i(\mathbf{x}), \tag{8}$$

where $\mathbf{x}$ and $I$ are as defined earlier and the $y_i$ are real, non-linear, differentiable functions generally. Let $\hat{y}_l(\mathbf{x})$, $l = 1, \ldots, n_r$ be the largest local discrete maxima (ripples) of $y_i(\mathbf{x})$, for $i \in I$, in descending magnitude. The grazor search method consists of solving a linear programme at the point $\mathbf{x}^j$

$$\text{maximize } \alpha_{k_r+1}(\mathbf{x}^j) \geqslant 0, \tag{9}$$

subject to

$$-\nabla^T \hat{y}_m(\mathbf{x}^j) \sum_{l=1}^{k_r} \alpha_l{}^j \nabla \hat{y}_l(\mathbf{x}^j) \leqslant -\alpha_{k_r+1}{}^j, \quad m = 1, \ldots, k_r, \tag{10}$$

$$\alpha_l{}^j \geqslant 0, \quad l = 1, \ldots, k_r, \tag{11}$$

$$\sum_{l=1}^{k_r} \alpha_l{}^j = 1, \tag{12}$$

where $\hat{y}_l(\mathbf{x}^j)$, $l = 1, \ldots, k_r$ are the highest ripples under consideration ($k_r \leqslant n_r$). When $k_r = 1$, we obtain the steepest descent direction for $\hat{y}_1(\mathbf{x}^j)$.

We next define

$$\Delta \mathbf{x}^j = -\sum_{l=1}^{k_r} \alpha_l{}^j \nabla \hat{y}_l(\mathbf{x}^j), \tag{13}$$

which is normalized to

$$\Delta \mathbf{x}_n{}^j = \Delta \mathbf{x}^j / \| \Delta \mathbf{x}^j \|. \tag{14}$$

Starting at $\mathbf{x}^j$, a step $\alpha^j \Delta \mathbf{x}_n{}^j$ is taken for $\alpha^j = \alpha_0{}^j$; if no improvement in $M$ results, $\alpha^j$ is reduced by factors of $\beta$ until a better point is obtained or $\alpha^j < \hat{\alpha}$. Let $\alpha^{j*}$ produce the first improved point from $\mathbf{x}^j$. Then $\Delta \mathbf{x}^0 = \alpha^{j*} \Delta \mathbf{x}_n{}^j$ is defined.

Next, a method based on golden section search to find the $\gamma^{j*}$ corresponding to the constrained minimum value of $M(\mathbf{x}^j + \gamma^j \Delta \mathbf{x}^0)$ is used. The $j$th iteration ends by setting $\mathbf{x}^{j+1} = \mathbf{x}^j + \gamma^{j*} \Delta \mathbf{x}^0$ and $\alpha_0{}^{j+1} = \alpha^{j*} \gamma^{j*}$.

For the system modelling problem, we let

$$y_i(\mathbf{x}) = |e_i(\mathbf{x})|, \quad i \in I. \tag{15}$$

## 5. Results

The system modelled was the seventh-order system used previously by Bandler *et al.* (1972 a). The computer used was a CDC 6400. The time interval over which the approximation was made was 0 to 8 sec. 101 uniformly spaced sample points were chosen over the interval. The steady-state value of the model was set at 0·11706, corresponding to the response of the system at the final sample point.

First the following transfer function for a model was used

$$H(s) = \frac{E a_0}{s^2 + a_1 s + a_0}, \tag{16}$$

where $E$ was the steady-state value and the parameter vector was

$$\mathbf{x}^{\mathrm{T}} = [a_0 \;\; a_1]. \tag{17}$$

The optimum parameters obtained using the grazor search method were

$$a_0 = 3{\cdot}06472,$$

$$a_1 = 2{\cdot}38338,$$

resulting in a four-ripple error curve with a maximum error value

$$M = 3{\cdot}76347 \times 10^{-3}.$$

The response and error curves are shown in figs. 1 (*a*) and 1 (*b*), respectively.

The optimum parameters obtained using least $p$th approximation for $p = 1000$ were

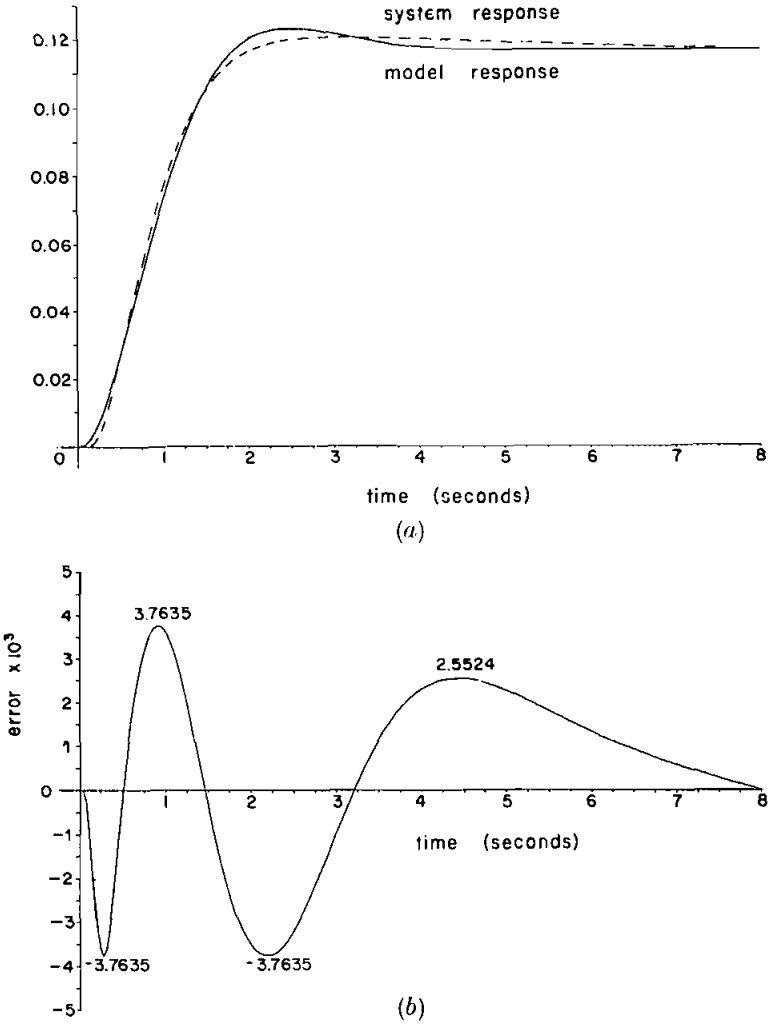
$$a_0 = 3{\cdot}06549,$$

$$a_1 = 2{\cdot}38414,$$

resulting in a similar four-ripple curve with a maximum error value

$$M = 3{\cdot}76510 \times 10^{-3}.$$

Fig. 1




Two parameter optimum (*a*) responses, (*b*) error curve.

Table 1 shows the number of function evaluations required for each of the methods to reach a maximum error value of $3{\cdot}76619 \times 10^{-3}$†. For this problem the Fletcher method and the Jacobson–Oksman method appeared to be the most efficient. In the sixth column of table 1 are the results obtained with the Jacobson–Oksman method when a homogeneous step prediction was used. That is, the scalar $\lambda$ by which the parameter increment was multiplied was selected on the assumption that the function was homogeneous. If, however, the degree of homogeneity in that iteration was negative, then it was assumed that the function was quadratic at that stage, and a quadratic

† Some of the results appeared at the 1972 Princeton Conference. The captions of tables 1 and 2 of that paper are in error. The tables actually indicate the number of function evaluations required to reach a certain value $M$ in the minimization of $M(\mathbf{x})$ and a certain value $F$ in the minimization of $F(\mathbf{x})$.

Table 1. Number of function evaluations required to reach $M = 3{\cdot}76619 \times 10^{-3}$ for the 2-parameter problem

| Starting point **x** | Minimization of $M(\mathbf{x})$ grazor | Minimization of $F(\mathbf{x})$ | | | |
|---|---|---|---|---|---|
| | | Fletcher | Fletcher–Powell | Jacobson–Oksman | |
| | | | | Quadratic step prediction | Homogeneous step prediction |
| 3·0<br>2·0 | 107 | 42 | 59 | 36 | 36 |
| 1·0<br>1·0 | 130 | 78 | 334 | 91 | 127 |
| 1·0<br>4·0 | 165 | 96 | 718 | 834 | † |
| 4·0<br>1·0 | 129 | 64 | False optimum | 41 | 45 |

† Indicates an ARGUMENT TOO LARGE message was given by the computer.

step prediction was used. The results shown in the fifth column were obtained using only a quadratic step prediction. It appeared, for this particular problem, that the quadratic step prediction gave more efficient results. In both cases however, if the predicted value $\lambda$ was larger than a preselected value $\rho$, then this preselected value $\rho$ was taken as the scalar $\lambda$. In other words $\rho$ was the upper limit on $\lambda$.

By allowing the model to have a zero, we have the following transfer function

$$H(s) = \frac{b_1 s + E a_0}{s^2 + a_1 s + a_0}, \tag{18}$$

resulting in a three-variable problem with a parameter vector

$$\mathbf{x}^{\mathrm{T}} = [a_0 \;\; a_1 \;\; b_1]. \tag{19}$$

The optimum parameters obtained using the grazor search method were

$$a_0 = 3{\cdot}83255,$$

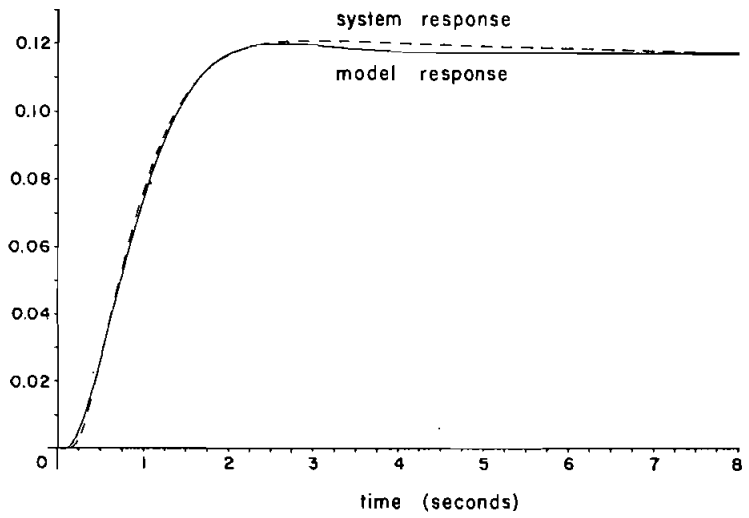
$$a_1 = 3{\cdot}00365,$$

$$b_1 = -1{\cdot}76390 \times 10^{-2},$$

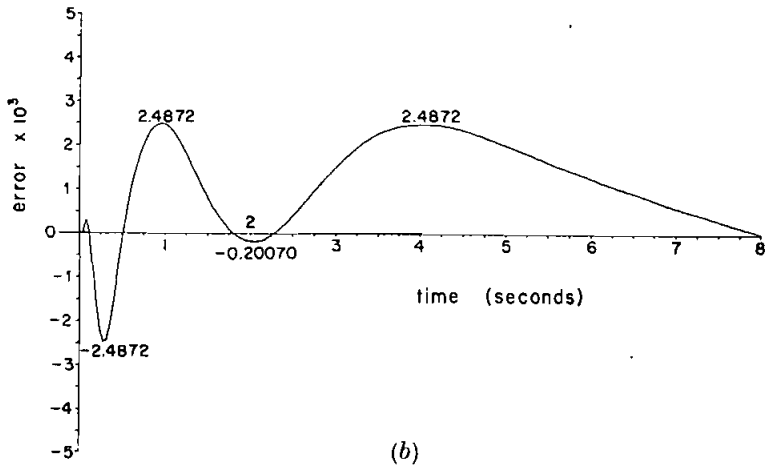
giving a maximum error value

$$M = 2{\cdot}48724 \times 10^{-3}.$$

Fig. 2




(*a*)




(*b*)

Three-parameter optimum (*a*) responses, (*b*) error curve.

The response and error curves are shown in figs. 2 (*a*) and (*b*), respectively.

For $p = 1000$ the optimum parameters obtained were

$$a_0 = 3{\cdot}83592,$$
$$a_1 = 3{\cdot}00605,$$
$$b_1 = -1{\cdot}77277 \times 10^{-2},$$

giving similar response and error curves as in figs. 2 (*a*) and (*b*) and

$$M = 2{\cdot}48794 \times 10^{-3}.$$

Table 2. Number of function evaluations required to reach $M = 2{\cdot}48794 \times 10^{-3}$ for the 3-parameter problem

| Starting point **x** | Minimization of $M(\mathbf{x})$ grazor | Minimization of $F(\mathbf{x})$ | | | | |
|---|---|---|---|---|---|---|
| | | Fletcher | Fletcher–Powell | Jacobson–Oksman | | |
| | | | | Quadratic step prediction | | Homogeneous step prediction |
| | | | | $\rho = 1$ | $\rho = 0{\cdot}5$ | |
| 2·5<br>2·0<br>−2·0 | 149 | 339 | 500 | 279 | ‡ | 339 |
| 1·0<br>1·0<br>−1·0 | 368 | 362 | † | 104 | 276 | 137 |
| 4·0<br>3·0<br>0·01 | 165 | 242 | 184 | 142 | 97 | 260 |
| 3·5<br>1·5<br>−1·0 | 358 | 280 | 342 | 217 | 151 | ‡ |
| 5·0<br>1·0<br>−1·0 | 325 | 193 | † | ‡ | 205 | ‡ |
| 5·0<br>1·0<br>3·0 | 406 | 245 | † | 159 | 119 | ‡ |

† Indicates time limit of 64 sec was reached.
‡ Indicates an ARGUMENT TOO LARGE message was given by the computer.

The number of function evaluations required for the three parameter problem to reach the value $M = 2{\cdot}48794 \times 10^{-3}$ are shown in table 2†. The grazor search technique and the Fletcher method required a smaller number of function evaluations. The Jacobson–Oksman method, with the homogeneous step prediction strategy described earlier, failed in 50% of the cases tried. This failure was due to large variations in parameter values and numbers, in the estimation of the objective function, becoming too large. By using only a quadratic step prediction, a noticeable improvement in efficiency of the algorithm occurred. The results are shown in table 2. In addition to the quadratic step prediction, the value of $\rho$ which was normally kept at 1·0, was now set at 0·5, resulting in improvement from some starting points.

† See footnote on p. 321.

For the final problem the following transfer function was considered

$$H(s)=\frac{b_2s^2+b_1s+Ea_0}{s^3+a_2s^2+a_1s+a_0}, \tag{20}$$

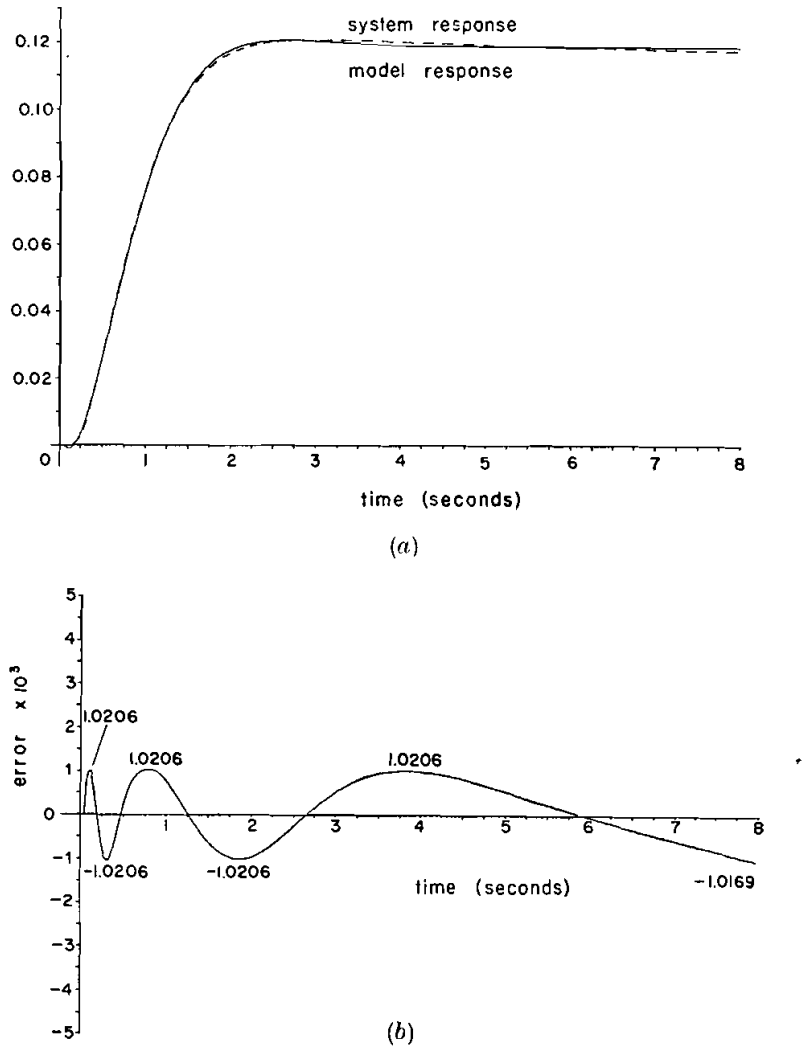
which, for computational efficiency, was used in the form

$$H(s)=\frac{x_5s^2+x_4s+Ex_1x_3}{(s+x_3)(s^2+x_2s+x_1)}, \tag{21}$$

resulting in a five-variable problem with parameter vector

$$\mathbf{x}^{\mathrm{T}}=[x_1\ x_2\ x_3\ x_4\ x_5]. \tag{22}$$

Fig. 3




Five-parameter six-ripple optimum (*a*) responses, (*b*) error curve.

The optimum parameters obtained using the grazor search method were

$$x_1 = 4{\cdot}34547,$$
$$x_2 = 3{\cdot}36809,$$
$$x_3 = 1{\cdot}08248 \times 10^{-1},$$
$$x_4 = 5{\cdot}14475 \times 10^{-1},$$
$$x_5 = -3{\cdot}56180 \times 10^{-2},$$

resulting in a six-ripple error curve with a maximum error value

$$M = 1{\cdot}02062 \times 10^{-3}.$$

Table 3. Number of function evaluations required to reach the shown value of 1000 M for the 5-parameter problem

| Starting point $\mathbf{x}$ | Minimization of $M(\mathbf{x})$ grazor | Minimization of $F(\mathbf{x})$ | | |
|---|---|---|---|---|
| | | Fletcher | Jacobson–Oksman quadratic step prediction | |
| | | | $\rho = 1$ | $\rho = 0{\cdot}5$ |
| 3·0<br>3·0<br>1·5 | 437 | 530 | 886 | 778 |
| 0·5<br>−0·1 | 1·2139 | 1·0207 | 1·0206 | 1·0206 |
| 1·5<br>3·0<br>2·5 | 782 | 768 | 931 | 325‡ |
| 1·0<br>0·1 | 1·2473 | 1·0207 | 1·0206 | 45·086 |
| 4·0<br>3·0<br>0·1 | 489 | 177 | 114‡ | 108 |
| 0·5<br>−0·03 | 1·0206 | 1·0207 | 1·5061 | 1·0206 |
| 3·0<br>5·0<br>0·2 | 634 | 862 | 248 | 350 |
| 0·3<br>−0·1 | 1·1720 | 1·0207 | 1·0206 | 1·0207 |
| 5·0<br>4·0<br>0·5 | 817 | 484 | 17‡ | 582 |
| 1·0<br>−0·5 | 1·0337 | 1·0207 | 19·660 | 1·0207 |
| Least-squares optimum | 537<br>1·2472 | 799<br>1·0206 | 263‡<br>1·8954 | 1208†<br>1·0283 |

† Indicates time limit of 128 sec was reached.
‡ Indicates an ARGUMENT TOO LARGE message was given by the computer.

The response and error curves are shown in figs. 3 (*a*) and (*b*), respectively.

The optimum parameters obtained using $p = 1000$ were

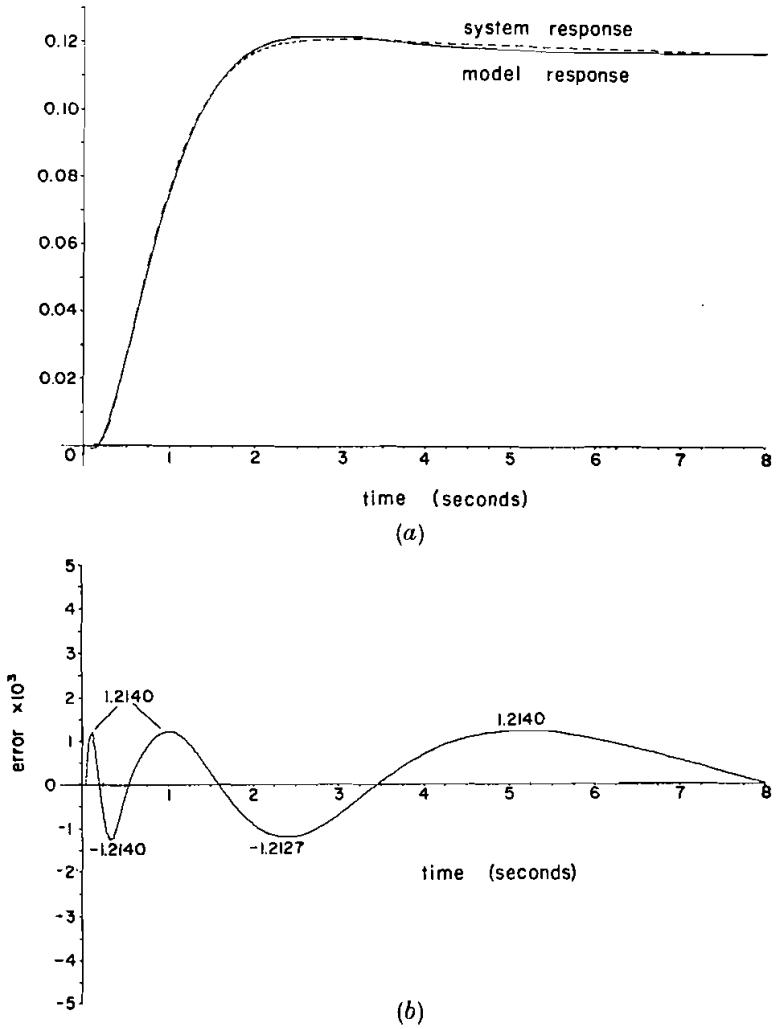
$$x_1 = 4{\cdot}34682,$$
$$x_2 = 3{\cdot}36738,$$
$$x_3 = 9{\cdot}96086 \times 10^{-2},$$
$$x_4 = 5{\cdot}14728 \times 10^{-1},$$
$$x_5 = -3{\cdot}56154 \times 10^{-2},$$

giving curves similar to those of figs. 3 (*a*) and (*b*) and a maximum error

$$M = 1{\cdot}02063 \times 10^{-3}.$$

Some runs with the Fletcher–Powell method, on the five-parameter problem, indicated the method was the slowest and since this was already

Fig. 4




Five-parameter five-ripple solution (*a*) responses, (*b*) error curve.

established in the previous problems, further runs of the Fletcher–Powell method were considered unnecessary. The Jacobson–Oksman method, used with a homogeneous step prediction, as explained earlier, failed from each starting point. With a quadratic step prediction however and with $\rho = 1{\cdot}0$ the method succeeded in locating the optimum in 50% of the cases tried. By reducing $\rho$ to 0·5 the method failed to locate the optimum only once. The results are shown in table 3.

The Fletcher method reached a unique six-ripple solution in all the cases tried, although there was a large variation in the number of function evaluations required. The grazor search technique reached the six-ripple solution in one of the cases shown, while in some of the other cases it terminated in a five-ripple solution.

In some instances, the real pole of the model had the tendency to move to the right-hand side of the $s$-plane and since this would produce an unstable model, the last parameters giving stable results were taken as the final values. In all cases, however, the real pole seems to lie very close to the $j\omega$ axis and any constraint, although easily implemented in the form of square transformation, would have made the pole go to zero.

It was further noted that when the Fletcher method, used with $p = 1000$, was started from one of the five-ripple solutions where the grazor search technique terminated, a direction was found which decreased $F(\mathbf{x})$ while temporarily increasing $M(\mathbf{x})$ and the method converged towards the six-ripple minimax solution, though slowly. When the same procedure was repeated with $p = 10^6$, the algorithm failed to move from that point. Figures 4 (*a*) and (*b*) show the response and error curves for a five-ripple solution obtained by the grazor search method.

## 6. Conditions for a minimax optimum

If $\hat{y}_l(\mathbf{x})$, $l = 1, 2, \ldots, k_r$ are taken for practical purposes, as equal, then for $\mathbf{x}$ to satisfy the necessary conditions for a minimax optimum (Bandler 1971), there exist non-negative multipliers $u_l$, $l = 1, \ldots, k_r$ such that

$$\sum_{l=1}^{k_r} u_l \nabla \hat{y}_l(\mathbf{x}) = \mathbf{0}, \tag{23}$$

$$\sum_{l=1}^{k_r} u_l = 1. \tag{24}$$

These conditions were applied to the final parameter values arrived at through optimization by the grazor search method, and found to be satisfied in all cases, as indicated below.

(*a*) *2-parameter solution:* $n_r = 4$, $k_r = 3$

| $l$ | Time instant | Error maximum ($1000\,\hat{y}_l$) | Multiplier ($u_l$) |
|---|---|---|---|
| 1 | 0·24 | 3·76347 | 0·75047 |
| 2 | 0·88 | 3·76347 | 0·16519 |
| 3 | 2·16 | 3·76347 | $8{\cdot}4342 \times 10^{-2}$ |
| 4 | 4·40 | 2·55235 | — |

$$\sum_{l=1}^{k_r} u_l \nabla \hat{y}_l = [0{\cdot}0 \;\; 0{\cdot}0]^T$$

$$\sum_{l=1}^{k_r} u_l = 1{\cdot}0,$$

(*b*) 3-*parameter solution:* $n_r = 4$, $k_r = 3$

| $l$ | Time instant | Error maximum ($1000\,\hat{y}_l$) | Multiplier ($u_l$) |
|---|---|---|---|
| 1 | 4·00 | 2·48724 | 0·90758 |
| 2 | 0·24 | 2·48724 | $4{\cdot}2744 \times 10^{-2}$ |
| 3 | 0·96 | 2·48724 | $4{\cdot}9680 \times 10^{-2}$ |
| 4 | 2·00 | $2{\cdot}00700 \times 10^{-1}$ | — |

$$\sum_{l=1}^{k_r} u_l \nabla \hat{y}_l = [0{\cdot}0 \;\; 0{\cdot}0 \;\; 1{\cdot}1 \times 10^{-5}]^T$$

$$\sum_{l=1}^{k_r} u_l = 1{\cdot}0.$$

(*c*) 5-*parameter solution*

(i) Six-ripple solution : $n_r = 6$, $k_r = 6$.

| $l$ | Time instant | Error maximum ($1000\,\hat{y}_l$) | Multiplier ($u_l$) |
|---|---|---|---|
| 1 | 1·84 | 1·020616 | $3{\cdot}6510 \times 10^{-2}$ |
| 2 | 0·72 | 1·020616 | $8{\cdot}4333 \times 10^{-2}$ |
| 3 | 0·08 | 1·020616 | 0·51806 |
| 4 | 3·76 | 1·020616 | $2{\cdot}7915 \times 10^{-2}$ |
| 5 | 0·24 | 1·020616 | 0·32227 |
| 6 | 8·00 | 1·016870 | $1{\cdot}0910 \times 10^{-2}$ |

$$\sum_{l=1}^{k_r} u_l \nabla \hat{y}_l = [0{\cdot}0 \;\; 0{\cdot}0 \;\; 0{\cdot}0 \;\; 0{\cdot}0 \;\; 0{\cdot}0]^T$$

$$\sum_{l=1}^{k_r} u_l = 1{\cdot}0.$$

(ii) Five-ripple solution : $n_r = 5$, $k_r = 5$.

| $l$ | Time instant | Error maximum ($1000\,\hat{y}_l$) | Multiplier ($u_l$) |
|---|---|---|---|
| 1 | 0·32 | 1·213988 | 0·23428 |
| 2 | 5·12 | 1·213988 | 0·19815 |
| 3 | 0·08 | 1·213988 | 0·39281 |
| 4 | 0·96 | 1·213986 | 0·10217 |
| 5 | 0·32 | 1·212651 | $7{\cdot}2598 \times 10^{-2}$ |

$$\sum_{l=1}^{k_r} u_l \nabla \hat{y}_l = [-1{\cdot}5 \times 10^{-5} \;\; 0{\cdot}0 \;\; 0{\cdot}0 \;\; 0{\cdot}0 \;\; 0{\cdot}0]^{\mathrm{T}}$$

$$\sum_{l=1}^{k_r} u_l = 1{\cdot}0.$$

For cases (*b*) and (*c*, ii), $k_r$ is equal to $k$ and there are $k_r+1$ equations and $k_r$ unknowns for the solution of (23) and (24). The dependent eqns. in (23) are kept aside, while the independent eqns. of (23), together with (24), are solved for the $u_l$, $l=1, \ldots, k_r$. The values of the $u_l$ are now substituted back into the dependent eqns. of (23) to check if the residual values are nearly zero. The non-zero values of the components of

$$\sum_{l=1}^{k_r} u_l \nabla \hat{y}_l$$

for cases (*b*) and (*c*, ii) correspond to the residuals of the dependent equations.

In interpreting these results one may associate (*a*) and (*c*, i) in saying that the main criterion is how close to equal the ripples are and (*b*) and (*c*, ii) in how small the size of the linear combination is in comparison with the sizes of the individual gradient vectors. In the first case we are satisfied with the criterion from a practical point of view, in the second the linear combination is about 2 to 4 orders of magnitude smaller than the gradient vectors.

## 7. Conclusions

The grazor search algorithm is found to be more efficient than the Fletcher–Powell method on the problems chosen. The method proposed by Fletcher appears to be the most efficient of the methods used in that firstly it required, for most cases, a smaller number of function evaluations and secondly it was consistent in reaching the optimum. The Jacobson–Oksman method, although giving good results, appeared to be sensitive to scaling. This is shown by the improvement that occurred when $\rho$ was reduced to 0·5 from 1·0. In the revised edition of the Jacobson–Oksman paper (private communication) a strategy is suggested where $\rho$ is reduced automatically in the algorithm by using a certain criterion. This might improve the method for the type of problems solved in this paper.

## Acknowledgments

This work was supported by grant A7239 and A3374 from the National Research Council of Canada. The authors are grateful to Dr. D. H. Jacobson and Dr. R. Fletcher for making available their computer programmes. The cooperation of Dr. N. K. Sinha in the formulation of the problem is appreciated. Useful suggestions made by Mr. C. Charalambous are also acknowledged.

## References

Bandler, J. W., 1971, *I.E.E.E. Trans. Circuit Theory*, **18**, 476.

Bandler, J. W., and Charalambous, C., 1971, *5th Asilomar Conf. Circuits Systems*, Pacific Grove, California.

BANDLER, J. W., MARKETTOS, N. D., and SINHA, N. K., 1972 a, *Int. J. Systems Sci.*, to be published.

BANDLER, J. W., MARKETTOS, N. D., and SRINIVASAN, T. V., 1972 b, *6th Annual Princeton Conf. Inf. Sci. Systems*, Princeton, N.J.

BANDLER, J. W., SRINIVASAN, T. V., and CHARALAMBOUS, C., 1972 c, *I.E.E.E. Trans. microw. Theory Tech.*, **20**, 596.

FLETCHER, R., 1970, *Computer J.*, **13**, 317.

FLETCHER, R., and POWELL, M. J. D., 1963, *Computer J.*, **6**, 163.

JACOBSON, D. H., 1972, Harvard University, Cambridge, Mass., Private communication.

JACOBSON, D. H., and OKSMAN, W., 1970, *Harv. Univ. Tech. Rep.*, No. 618.