JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 11, No. 5, 1973

On Conditions for Optimality in Least pth Approximation[1] with $p \to \infty$

J. W. BANDLER[2] AND C. CHARALAMBOUS[3]

Communicated by C. T. Leondes

Abstract. This paper presents a theoretical discussion of the necessary and sufficient conditions for optimality in generalized nonlinear least pth approximation problems for $p \to \infty$. In the limit, the conditions for a minimax approximation are derived, as is to be expected. Numerical examples involving the modeling of a linear time-invariant fourth-order system by a second-order model and the design of quarter-wave transmission-line transformers illustrate the results.

1. Introduction

Of great practical importance to network and system designers wishing to approximate a specified response by a network or system response or desiring to meet or exceed certain design specifications is the optimality of their approximation. A number of workers interested in minimax approximations (Refs. 1–3) have independently arrived at similar conditions for nonlinear minimax approximation problems. These are naturally derivable from the Kuhn–Tucker conditions for a constrained optimum because of the close relationship between

[1] This work was supported by the National Research Council of Canada under Grant No. A7239 and by a Frederick Gardner Cottrell Grant from the Research Corporation. This paper was presented at the 9th Annual Allerton Conference on Circuit and System Theory, Urbana, Illinois, October 6–8, 1971. The authors thank Mrs. J. R. Popović for helping to correct Example 4.1.

[2] Associate Professor of Electrical Engineering, Department of Electrical Engineering, McMaster University, Hamilton, Ontario, Canada.

[3] Postdoctoral Fellow, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, Canada.

nonlinear minimax approximations and nonlinear programming. Bandler (Ref. 3), in particular, derived the appropriate conditions in a general form suitable for such problems as filter design.

Because of the widespread interest in nonlinear least pth approximation (Ref. 4), and because of recent results (Ref. 5) that permit least pth objectives in a more generalized sense to be directly applicable to such problems as meeting or exceeding design specifications as in filter design, it is felt that a detailed mathematical discussion of conditions for optimality is highly relevant. Thus, the present paper allows for situations more general than the conventional problem of approximation to a single continuous function on a closed interval.

2. Definitions and Assumptions

Define real error functions related to the *upper* and *lower* specifications, respectively, as (Refs. 5–6)

$$\begin{aligned} e_u(\phi, \psi) &\triangleq w_u(\psi)(F(\phi, \psi) - S_u(\psi)), \\ e_l(\phi, \psi) &\triangleq w_l(\psi)(F(\phi, \psi) - S_l(\psi)), \end{aligned} \tag{1}$$

where $F(\phi, \psi)$ is the approximating function (actual response), $S_u(\psi)$ is an upper specified function (desired response bound), $S_l(\psi)$ is a lower specified function (desired response bound), $w_u(\psi)$ is an upper positive weighting function, $w_l(\psi)$ is a lower positive weighting function, ϕ is a vector containing the k independent parameters (design variables), and ψ is an independent variable (e.g., frequency or time).

In practice, we will evaluate all the functions at a finite discrete set of values of ψ taken from one or more closed intervals. Therefore, we will let

$$\begin{aligned} e_{ui}(\phi) &\triangleq e_u(\phi, \psi_i), \qquad i \in I_u, \\ e_{li}(\phi) &\triangleq e_l(\phi, \psi_i), \qquad i \in I_l, \end{aligned} \tag{2}$$

where it is assumed that a sufficient number of sample points have been chosen so that the discrete approximation problem adequately approximates the continuous problem. I_u and I_l are appropriate index sets.

2.1. Case 1. Specification Violated. In the case when the specification is violated, some of the $e_{ui}(\phi)$ or $-e_{li}(\phi)$ are positive. In an

effort to meet the specification, we can propose the following function to be minimized:

$$U(\phi) = \left(\sum_{i\in J_u} [e_{ui}(\phi)]^p + \sum_{i\in J_l} [-e_{li}(\phi)]^p\right)^{1/p}, \tag{3}$$

where

$$\begin{aligned} J_u &\triangleq \{i \mid e_{ui}(\phi) \geqslant 0, \quad i \in I_u\}, \\ J_l &\triangleq \{i \mid -e_{li}(\phi) \geqslant 0, \quad i \in I_l\}, \end{aligned} \tag{4}$$

and $p > 1$.

The larger the value of p, the more nearly would we expect the maximum error to be emphasized, since

$$\max_i [e_{ui}(\phi), -e_{li}(\phi)] = \lim_{p\to\infty} \left(\sum_{i\in J_u} [e_{ui}(\phi)]^p + \sum_{i\in J_l} [-e_{li}(\phi)]^p\right)^{1/p}. \tag{5}$$

2.2. Case 2. Specification Satisfied. For the case when the specification is satisfied, all the $-e_{ui}(\phi)$ and $e_{li}(\phi)$ will be positive. This time, in an effort to exceed the specification by as much as possible, we can propose the following objective function to be minimized[4]:

$$U(\phi) = -\left(\sum_{i\in I_u} [-e_{ui}(\phi)]^{-p} + \sum_{i\in I_l} [e_{li}(\phi)]^{-p}\right)^{-1/p}, \tag{6}$$

where we assume

$$\begin{aligned} -e_{ui}(\phi) &> 0, \quad i \in I_u, \\ e_{li}(\phi) &> 0, \quad i \in I_l, \end{aligned} \tag{7}$$

and $p \geqslant 1$.

The larger the value of p the more nearly would we expect the minimum error to be emphasized, since

$$\min_i [-e_{ui}(\phi), e_{li}(\phi)] = \lim_{p\to\infty} \left(\sum_{i\in I_u} [-e_{ui}(\phi)]^{-p} + \sum_{i\in I_l} [e_{li}(\phi)]^{-p}\right)^{-1/p}. \tag{8}$$

2.3. Assumptions. It is assumed that a minimum exists in a closed and bounded region of points ϕ and that $e_{ui}(\phi)$ and $e_{li}(\phi)$ are continuous for all i with continuous partial derivatives, at least in the

[4] Since, in this paper, the conditions for optimality are of interest, it is convenient to use slightly different objective functions from those proposed in an earlier work (Ref. 5).

neighborhood of the minimum. Then, the objective functions proposed are continuous with continuous partial derivatives in the neighborhood of the minimum.

3. Theorems

Theorem 3.1. At an optimum point $\check{\phi}_\infty$ for a minimax approximation problem,

$$\sum_{i\in K_u} u_{ui}\,\nabla e_{ui}(\check{\phi}_\infty) = \sum_{i\in K_l} u_{li}\,\nabla e_{li}(\check{\phi}_\infty),$$

$$\sum_{i\in K_u} u_{ui} + \sum_{i\in K_l} u_{li} = 1,$$

$$u_{ui} \geqslant 0, \qquad i\in K_u,$$

$$u_{li} \geqslant 0, \qquad i\in K_l,$$

where

$$\nabla \triangleq [\partial/\partial\phi_1, \partial/\partial\phi_2, \ldots, \partial/\partial\phi_k]^T,$$

and where $e_{ui}(\check{\phi}_\infty)$ for $i\in K_u$ and $-e_{li}(\check{\phi}_\infty)$ for $i\in K_l$ are the equal maxima.

Proof for Case 1. Differentiating Eq. (3), we have

$$\begin{aligned}\nabla U(\phi) &= \Big(\sum_{i\in J_u}[e_{ui}(\phi)]^p + \sum_{i\in J_l}[-e_{li}(\phi)]^p\Big)^{1/p-1}\\ &\quad\times\Big(\sum_{i\in J_u}[e_{ui}(\phi)]^{p-1}\,\nabla e_{ui}(\phi) - \sum_{i\in J_l}[-e_{li}(\phi)]^{p-1}\,\nabla e_{li}(\phi)\Big)\\ &= \Big(\sum_{i\in J_u}[e_{ui}(\phi)]^p + \sum_{i\in J_l}[-e_{li}(\phi)]^p\Big)^{1/p}\\ &\quad\times\Big(\sum_{i\in J_u}\frac{[e_{ui}(\phi)]^p}{\sum_{i\in J_u}[e_{ui}(\phi)]^p + \sum_{i\in J_l}[-e_{li}(\phi)]^p}\,\frac{\nabla e_{ui}(\phi)}{e_{ui}(\phi)}\\ &\quad-\sum_{i\in J_l}\frac{[-e_{li}(\phi)]^p}{\sum_{i\in J_u}[e_{ui}(\phi)]^p + \sum_{i\in J_l}[-e_{li}(\phi)]^p}\,\frac{\nabla e_{li}(\phi)}{[-e_{li}(\phi)]}\Big).\end{aligned} \tag{9}$$

The necessary conditions for an optimum of $U(\phi)$ are that

$$\nabla U(\check{\phi}_p) = 0, \tag{10}$$

where $\check{\phi}_p$ denotes the optimum parameter vector for particular values of p. Let

$$\mu_i(p) = \frac{[e_{ui}(\check{\phi}_p)]^p}{\sum_{i\in J_u}[e_{ui}(\check{\phi}_p)]^p + \sum_{i\in J_l}[-e_{li}(\check{\phi}_p)]^p}, \tag{11}$$

$$\lambda_i(p) = \frac{[-e_{li}(\check{\phi}_p)]^p}{\sum_{i\in J_u}[e_{ui}(\check{\phi}_p)]^p + \sum_{i\in J_l}[-e_{li}(\check{\phi}_p)]^p}, \tag{12}$$

$$u_{ui} = \lim_{p\to\infty} \mu_i(p), \tag{13}$$

$$u_{li} = \lim_{p\to\infty} \lambda_i(p). \tag{14}$$

Then, assuming that the above limit exists, it is clear that, as $p \to \infty$, the necessary conditions for optimality applied to Eq. (9) yield

$$u_{ui} \begin{cases} = 0, & i \notin K_u, \\ \geqslant 0, & i \in K_u, \end{cases} \tag{15}$$

$$u_{li} \begin{cases} = 0, & i \notin K_l, \\ \geqslant 0, & i \in K_l, \end{cases} \tag{16}$$

and

$$\sum_{i\in K_u} u_{ui} + \sum_{i\in K_l} u_{li} = 1. \tag{17}$$

Therefore, as $p \to \infty$,

$$\sum_{i\in K_u} u_{ui} \nabla e_{ui}(\check{\phi}_\infty) = \sum_{i\in K_l} u_{li} \nabla e_{li}(\check{\phi}_\infty), \tag{18}$$

and the theorem is proved.

Proof for Case 2. Following a similar procedure to the one used for Case 1, the same conditions can be derived, but in this case we let

$$\mu_i(p) = \frac{[-e_{ui}(\check{\phi}_p)]^{-p}}{\sum_{i\in I_u}[-e_{ui}(\check{\phi}_p)]^{-p} + \sum_{i\in I_l}[e_{li}(\check{\phi}_p)]^{-p}}, \tag{19}$$

$$\lambda_i(p) = \frac{[e_{li}(\check{\phi}_p)]^{-p}}{\sum_{i\in I_u}[-e_{ui}(\check{\phi}_p)]^{-p} + \sum_{i\in I_l}[e_{li}(\check{\phi}_p)]^{-p}}. \tag{20}$$

Theorem 3.2. If the relations in Theorem 3.1 are satisfied at a point $\check{\phi}_\infty$ and the $e_{ui}(\phi)$ are convex for $i \in I_u$ and the $e_{li}(\phi)$ are concave for $i \in I_l$, then $\check{\phi}_\infty$ is optimal.

Proof for Case 1. Assuming the $e_{ui}(\phi)$ are convex and the $-e_{li}(\phi)$ are convex, we have

$$\begin{aligned} e_{ui}((1-\lambda)\,\phi^1+\lambda\phi^2) &\leqslant (1-\lambda)\,e_{ui}(\phi^1)+\lambda e_{ui}(\phi^2),\\ -e_{li}((1-\lambda)\,\phi^1+\lambda\phi^2) &\leqslant -(1-\lambda)\,e_{li}(\phi^1)-\lambda e_{li}(\phi^2), \end{aligned} \tag{21}$$

for $0 \leqslant \lambda \leqslant 1$. Therefore, for $p \geqslant 1$,

$$\begin{aligned} &\Big(\sum_{i\in J_u}[e_{ui}((1-\lambda)\,\phi^1+\lambda\phi^2)]^p+\sum_{i\in J_l}[-e_{li}((1-\lambda)\,\phi^1+\lambda\phi^2)]^p\Big)^{1/p}\\ &\quad\leqslant\Big(\sum_{i\in J_u}[(1-\lambda)\,e_{ui}(\phi^1)+\lambda e_{ui}(\phi^2)]^p\\ &\qquad+\sum_{i\in J_l}[-(1-\lambda)\,e_{li}(\phi^1)-\lambda e_{li}(\phi^2)]^p\Big)^{1/p}\\ &\quad\leqslant(1-\lambda)\Big(\sum_{i\in J_u}[e_{ui}(\phi^1)]^p+\sum_{i\in J_l}[-e_{li}(\phi^1)]^p\Big)^{1/p}\\ &\qquad+\lambda\Big(\sum_{i\in J_u}[e_{ui}(\phi^2)]^p+\sum_{i\in J_l}[-e_{li}(\phi^2)]^p\Big)^{1/p}, \end{aligned} \tag{22}$$

from Minkowski's inequality (Ref. 7). Hence,

$$U((1-\lambda)\,\phi^1+\lambda\phi^2) \leqslant (1-\lambda)\,U(\phi^1)+\lambda U(\phi^2). \tag{23}$$

Therefore, $U(\phi)$ is convex, and Theorem 3.2 follows.

Proof for Case 2. Using Eq. (21) we have, for $p \geqslant 1$,

$$\begin{aligned} &\Big(\sum_{i\in I_u}[-e_{ui}((1-\lambda)\,\phi^1+\lambda\phi^2)]^{-p}+\sum_{i\in I_l}[e_{li}((1-\lambda)\,\phi^1+\lambda\phi^2)]^{-p}\Big)^{-1/p}\\ &\quad\geqslant\Big(\sum_{i\in I_u}[-(1-\lambda)\,e_{ui}(\phi^1)-\lambda e_{ui}(\phi^2)]^{-p}\\ &\qquad+\sum_{i\in I_l}[(1-\lambda)\,e_{li}(\phi^1)+\lambda e_{li}(\phi^2)]^{-p}\Big)^{-1/p}\\ &\quad\geqslant(1-\lambda)\Big(\sum_{i\in I_u}[-e_{ui}(\phi^1)]^{-p}+\sum_{i\in I_l}[e_{li}(\phi^1)]^{-p}\Big)^{-1/p}\\ &\qquad+\lambda\Big(\sum_{i\in I_u}[-e_{ui}(\phi^2)]^{-p}+\sum_{i\in I_l}[e_{li}(\phi^2)]^{-p}\Big)^{-1/p}. \end{aligned} \tag{24}$$

This inequality (Ref. 7) is a counterpart to Minkowski's inequality. Hence,

$$-U((1-\lambda)\,\phi^1 + \lambda\phi^2) \geqslant -(1-\lambda)\,U(\phi^1) - \lambda U(\phi^2). \tag{25}$$

Therefore, $U(\phi)$ is convex, and Theorem 3.2 follows.

4. Examples

Example 4.1. The first example is to find a second-order model of a fourth-order system when the input to the system is an impulse for different values of p. The transfer function of the fourth-order system is

$$G(s) = (s+4)/(s+1)(s^2+4s+8)(s+5). \tag{26}$$

The transfer function of the second-order model considered is

$$H(s) = c/[(s+\alpha)^2 + \beta^2]. \tag{27}$$

Therefore, in our case, we have

$$S(t) = (3/20)\exp(-t) + (1/52)\exp(-5t) - [\exp(-2t)/65](3\sin 2t + 11\cos 2t), \tag{28}$$

$$F(\phi, t) = (c/\beta)\exp(-\alpha t)\sin\beta t, \tag{29}$$

$$\phi = [\alpha, \beta, c]^T, \tag{30}$$

$$\psi_i = t_i, \qquad 0 \leqslant t_i \leqslant 20, \tag{31}$$

$$e(\phi, t) = F(\phi, \psi) - S(\psi), \tag{32}$$

$$U(\phi) = \left[\sum_{i=1}^{101} |e(\phi, t_i)|^p\right]^{1/p}. \tag{33}$$

The above objective function, which corresponds to the situation $S_u = S_l = S$ and $w_u = w_l = 1$ in Case 1, was set up using 101 uniformly spaced points over the range 0 and 20 sec. The values of p used for optimization[5] were 2, 6, 10, 20, 30, 40, 50, 70, 80, 100, 200, 1000, and 10,000 (see Fig. 1).

[5] Techniques for least pth approximation with extremely large values of p are described elsewhere (Ref. 8).

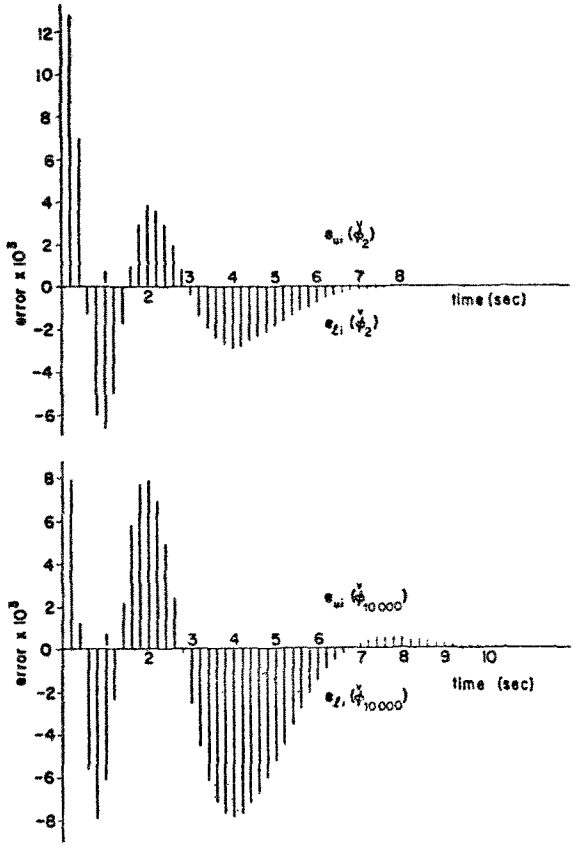

Fig. 1. Optimum error for least-squares and least-10,000th approximation for Example 4.1.

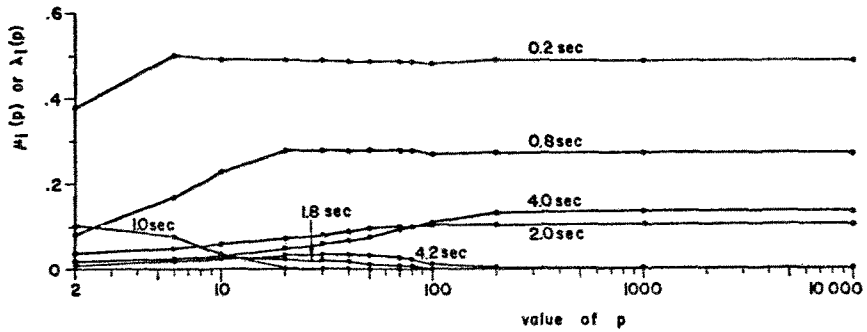

Fig. 2. $\mu_i(p)$ or $\lambda_i(p)$, as appropriate, calculated at specified values of time and for certain values of p for Example 4.1.

The values of $\mu_i(p)$ and $\lambda_i(p)$ were calculated for different values of p and for different values of t. Seven values of t were considered, of which four were the points where the approximately equal extrema occurred. Figure 2 demonstrates the validity of Eqs. (15) and (16).

Example 4.2. Here, we consider the design of 10-ohm to 1-ohm quarter-wave transmission-line transformers taking, for convenience, 0.3 as an upper specification for the magnitude of the reflection coefficient ρ over a specified 100% frequency band. The basic problem has been previously defined and analyzed in the context of optimality by Bandler (Ref. 3, see also Refs. 9 and 10). The optimum two-section transformer violates the specification. The optimum three-section transformer, on the other hand, satisfies the specification. The value of p for both cases in the optimization process was 10,000. 101 uniformly spaced sample points were used.

The maximum values of $e_{ui}(\phi_p)$, where $F = |\rho|$ and $S_u = 0.3$, and the frequencies at which they occur are shown in Tables 1 and 2. Using Eq. (11) and Table 1, nonzero multipliers in the ratio

$$3.001 : 2.001 : 1.000$$

can be found; and, using Eq. (19) and Table 2, nonzero multipliers in the ratio

$$2.999 : 2.807 : 1.759 : 1.000$$

Table 1. Optimum two-section transformer with $p = 10^4$ (specification violated, Case 1).

Maximum of $e_{ui}(\phi_p)$	0.12857552	0.12857031	0.12856139
Frequencies at which maximum occurs, GHz	0.5	1.0	1.5

Table 2. Optimum three-section transformer with $p = 10^4$ (specification satisfied, Case 2).

Maximum of $-e_{ui}(\phi_p)$	0.10270671	0.10270739	0.10271219	0.10271799
Frequencies at which maximum occurs, GHz	0.5	0.77	1.23	1.5

can be found. These results may be compared with those of Bandler (Ref. 3).

5. Conclusions

It has been shown (Ref. 5) that, for violated specifications in generalized least pth approximation, we have a close analogy with minimization of a penalty term designed to bring one closer to the boundary of the feasible region, whereas, for satisfied specifications, we have an analogy with minimization of a penalty term designed to steer a feasible solution deeper into the feasible region. Conventional approximation problems fit into the former category, but conditions of optimality are required for both.

In the limit, the conditions for a minimax approximation are obtained as is to be expected. It is felt that new insight into minimax algorithms is gained, since appropriate algorithms might attempt to force these conditions in an iterative manner by using extremely large values of p.

References

1. DEM'YANOV, V. F., *Sufficient Conditions for Local Minimax*, Zh. Vychisl. Mat. Fiz., Vol. 10, No. 5, 1970.
2. MEDANIC, J., *Solution of the Convex Minimax Problem by the Newton–Raphson Method*, Proceedings of the 8th Annual Allerton Conference on Circuit and System Theory, Urbana, Ill., 1970.
3. BANDLER, J. W., *Conditions for a Minimax Optimum*, IEEE Transactions on Circuit Theory, Vol. CT-18, No. 4, 1971.
4. TEMES, G. C., and ZAI, D. Y. F., *Least pth Approximation*, IEEE Transactions on Circuit Theory, Vol. CT-16, No. 2, 1969.
5. BANDLER, J. W., and CHARALAMBOUS, C., *Theory of Generalized Least pth Approximation*, IEEE Transactions on Circuit Theory, Vol. CT-19, No. 3, 1972.
6. BANDLER, J. W., *Optimization Methods for Computer-Aided Design*, IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-17, No. 8, 1969.
7. HARDY, G. H., LITTLEWOOD, J. E., and PÓLYA, G., *Inequalities*, Cambridge University Press, Cambridge, England, 1934.
8. BANDLER, J. W., and CHARALAMBOUS, C., *Practical Least pth Approximation with Extremely Large Values of* p, Conference Record of the 5th Asilomar Conference on Circuits and Systems, Pacific Grove, Calif., 1971.

Additional Bibliography

9. BANDLER, J. W., and MACDONALD, P. A., *Cascaded Noncommensurate Transmission-Line Networks as Optimization Problems*, IEEE Transactions on Circuit Theory, Vol. CT-16, No. 3, 1969.
10. BANDLER, J. W., and MACDONALD, P. A., *Optimization of Microwave Networks by Razor Search*, IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-17, No. 8, 1969.

Printed in Belgium