The selection of optimal break points in piecewise linear function analysis
by Haifie Loo Lai

A thesis submitted to the Graduate Faculty in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE in Industrial Engineering
Montana State University

Abstract:
Methods are presented which optimally select the break points of a piecewise linear function used to approximate a nonlinear functional.

On one hand, given the analytic functions which are to be approximated by piecewise linear functions, and the number of segments of the linear function, break points are selected so as to minimize the total absolute difference between the analytic function and the approximating piecewise linear functions. On the other hand, given a table of raw data values presented in ordered array, which are to be fitted by piecewise linear functions, break points are selected at the largest value of the independent variable such that no infeasibility is incurred. Infeasibility occurs whenever the deviation of the observed data from the fitted linear segment exceeds a prescribed value K, the maximum allowable deviation specified. These break-point search methods are compatible with and incorporated into the usual method of separable programming to obtain more precise approximating solutions.

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at Montana State University, I agree that the Library shall make it freely available for inspection. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by my major professor, or, in his absence, by the Director of Libraries. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature *Maifie Loo Lai*

Date *Dec. 3. 1973*

THE SELECTION OF OPTIMAL BREAK POINTS IN
PIECEWISE LINEAR FUNCTION ANALYSIS

by

HAIFIE LOO LAI

A thesis submitted to the Graduate Faculty in partial
fulfillment of the requirements for the degree

of

MASTER OF SCIENCE

in

Industrial Engineering

Approved:

William [signature]
Head, Major Department

Donald W. Boyd
Chairman, Examining Committee

K Goering
Graduate Dean

MONTANA STATE UNIVERSITY
Bozeman, Montana

December, 1973

ACKNOWLEDGMENTS

The writer gratefully acknowledges Dr. D. W. Boyd for his invaluable advice in structuring this research. Special thanks to Dr. D. F. Gibson and Dr. W. R. Taylor for their contribution in examining this research.

Finally, to my husband, Ching-E, for his patience and encouragement in fulfilling this assignment.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Methods are presented which optimally select the break points of a piecewise linear function used to approximate a nonlinear functional. On one hand, given the analytic functions which are to be approximated by piecewise linear functions, and the number of segments of the linear function, break points are selected so as to minimize the total absolute difference between the analytic function and the approximating piecewise linear functions. On the other hand, given a table of raw data values presented in ordered array, which are to be fitted by piecewise linear functions, break points are selected at the largest value of the independent variable such that no infeasibility is incurred. Infeasibility occurs whenever the deviation of the observed data from the fitted linear segment exceeds a prescribed value K, the maximum allowable deviation specified. These break-point search methods are compatible with and incorporated into the usual method of separable programming to obtain more precise approximating solutions.

CHAPTER I

INTRODUCTION

The Problem Class

Mathematical programming covers an extensive area composed of a great variety of problems for which numerous methods have been developed for solving particular classes of problems.

One important class of problems occurs in nonlinear programming in which the objective function and each of the constraints consist of separable functions. That is, minimize (maximize)

$$F = \sum_{j=1}^{n} f_j(x_j)$$

subject to $\sum_{j=1}^{n} g_{ij}(x_j)\{\leq, =, \geq\}b_i$

$$x_j \geq 0$$

where $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$.

Such problems can be solved by replacing the functions $f_j(x_j)$ and $g_{ij}(x_j)$ with polygonal approximations. That is, each function is approximated by straight line segments between selected points. Hence, the original nonlinear programming problem is transformed into a linear programming problem so that an approximate solution can be obtained by a slight modification of the simplex method. Thus, for the problem stated above, the linear programming counterpart is, minimize (maximize) $F = \sum_{j=1}^{n} \hat{f}_j(x_j)$

subject to $\sum_{j=1}^{n} \hat{g}_{ij}(x_j)\{\leq, =, \geq\}b_i$

$x_j \geq 0$

where $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$;

$\hat{f}_j(x_j)$ is the polygonal approximation to the function $f_j(x_j)$; and $\hat{g}_{ij}(x_j)$ is the polygonal approximation to the function $g_{ij}(x_j)$.

The key concept of this special technique is that a separable concave (convex) function of a single variable can be approximated as closely as desired by piecewise linear functions. In fact, second order conditions for the existence of an optimum require that all of the $f_j(x_j)$ be concave (convex) functions when the objective is to maximize (minimize) $\sum_{j=1}^{n} f_j(x_j)$.

Summary of Past Work

Previous studies have resulted in the development of algorithms for the separable convex programming problem. For example, Charnes and Lemke (6) and Miller (18) have developed linear approximation methods. The main feature of these methods is to linearize the nonlinear functions on a grid of points spanning a suitable portion of the space of the problem. Let $X^1, X^2, \ldots, X^T$ be a collection of n-vectors ($X^T = (x_1^T, x_2^T, \ldots, x_n^T)$). Any point x_i of this collection may be written as

$x_i = \sum_{t=1}^{T} \lambda^t x_i^t$ where $\sum_t \lambda^t = 1$ for all t. Then, given any function f(x), the linearization of f(x) on the grid $x^1, x^2, \ldots, x^T$ is attained through the approximation $f(x) = \sum_t \lambda^t f(x^t)$. Thus, any nonlinear programming problem becomes a linear problem in the new variable λ^t if f(x) is replaced throughout by its representation above. The nonlinear programming problem can now be stated in the approximate form:

Maximize (minimize) $\sum_t \lambda^t f(x^t)$

subject to $\sum_t \lambda^t = 1, \quad \lambda^t \geq 0$

$$\sum_t \lambda^t g_i(x^t)\{\leq, =, \geq\}0 \text{ for all i.}$$

The transformed problem can be solved by the usual simplex method.

Another similar method for the solution of the separable convex programming problem is presented by Hillier and Lieberman (13), and also makes use of a linear programming approximation of the original nonlinear programming problem. However, the approximation function $\hat{f}_j(x_j)$, to the true curve $f_j(x_j)$, is constructed as $\hat{f}_j(x_j) = s_{j1}x_{j1} + \ldots + s_{jk}x_{jk} + \ldots + s_{jmj}x_{jmj}$; where s_{jk} is the slope of the piecewise linear approximating function. The original variable, x_j, is broken into m_j segments at the breakpoints, b_{jk}. The new variable, x_{jk}, is defined as the k^{th} segment of x_j. The problem is then transformed to:

$$\text{Maximize (Minimize)} \quad \sum_j \left(\sum_{k=1}^{m_j} s_{jk} x_{jk} \right)$$

$$\text{subject to} \quad \sum_j g_{ij}(x_{jm_j}) \{\leq, =, \geq\}\ 0 \text{ for all } i$$

$$\left. \begin{array}{l} x_{jk} \leq b_{jk} \\ x_{jk} \geq 0 \end{array} \right\} \text{for all } j, k.$$

In both cases, the closeness to which the linearized solution approximates the optimum is determined by the fineness of the grid. If the chosen grid is fine enough in the neighborhood of the solution, answers of suitable accuracy can be obtained.

After the linear programming approximation is solved, if a more precise answer is needed, a grid refinement process can be used. The algorithm for grid refinement described by Dantzig (8) starts with the computing of coefficients for each term of the approximating functions, for instance at $x = 0, 0.5, 1$, with grid size 0.5. Upon solving the first linear programming approximation, halve the grid size and compute coefficients for only those new values which are adjacent to the current solution for x. Thus, on the second piecewise approximation, if the previous solution were $x = 0$, compute the coefficient at $x = 0.25$, discarding the value at $x = 1$; if $x = 0.5$ compute coefficients at 0.25 and 0.75, and discard the values at 0 and 1; if the solution x is a weighted average of two grid points 0 and 0.5, then include a grid value of x at .25 and discard the value at 1, etc. In this way, for each successive piecewise approximation problem, coefficients are computed for at most three values of each x and such that the range of x is halved each time. It was pointed out that successive values of x obtained in this manner will stay within the range established by previous cycles.

Another grid refinement procedure is described by Alloin (2). Suppose that a grid $x^1, x^2, \ldots, x^T$ is given, and the associated linear programming approximation is solved, yielding $\lambda^1, \lambda^2, \ldots, \lambda^T$ and the dual solution $\bar{u}_0, \bar{u}_1, \ldots, \bar{u}_m$. When refinement of the grid is needed, of several possible points that might be adjoined to the given grid as a further refinement to obtain a better solution, which point would the simplex method choose as contributing the most to the solution? He suggested that the decomposition procedure can be applied. The procedure is to evaluate the reduced cost for each new column of the linear programming problem and to choose the one with largest reduced cost. Or equivalently, x^{T+1} is the maximizing value of $f_j(x_j) - \sum_i \bar{u}_i g_{ij}(x_j)$, ($x_j$ unconstrained). Thus, a new column which tightens the grid is constructed, a new variable is added, and the simplex method again employed to find a new solution to the expanded linear programming problem.

Purpose of This Study

The survey of past work reveals that the approximation techniques used in all the existing algorithms are simply based upon judicious selection of the break points of the piecewise linear function rather than selection via a stated criterion. Hence, development of a solution technique to include "non-subjective" selection of break points is believed to be useful. Therefore, the main purpose of this study is to develop some criterion for the optimal segmentation of $f_j(x_j)$ for piecewise linear approximation. The solution technique is also to be adaptable to the usual separable convex programming approach.

Closely related to this problem is the analysis of raw data which exhibit no interactions between independent variables. A method is sought by which to analyze the objective function and constraints, which may show strong nonlinearities, while still in raw-data form.

The development of solution techniques for separable convex functions and for the raw-data form of the problem is presented in Chapter II and Chapter III, respectively. Chapter IV contains two illustrative examples, one for the separable convex programming problem, and another for the analysis of raw data. A summary discussion and suggestions are presented in Chapter V.

CHAPTER II

DEVELOPMENT OF BREAK POINT SEARCH METHOD FOR THE SEPARABLE CONVEX FUNCTION

The Approximation of a Separable Function

Consider the case where the given function can be written in the form $f(x_1, x_2, \ldots, x_n) = f_1(x_1) + f_2(x_2) + \ldots + f_n(x_n)$, where $f_j(x_j)$ is a specified function of x_j only, for $j = 1, 2, \ldots, n$. Thus, $f(x_1, x_2, \ldots, x_n)$ is called a "separable" function, since it can be written as a finite sum of separate terms, each of which involves only a single variable.

An approximation technique is available for obtaining a solution to the problem having such a separable objective function. This technique involves reducing the problem to a linear programming problem by approximating each $f_j(x_j)$ by several piecewise linear functions. Consider, for example, the function sketched in Figure 2-1. The piecewise linear approximation is given by $\hat{f}(x)$ where the s's represent the slopes of the approximating lines and the b's represent the break points selected to give the desired degree of approximation. When $f(x)$ is differentiable, the value $df(x)/dx$ can be plotted against x as in Figure 2-2. The area under the $df(x)/dx$ curve may be used to represent $f(x)$ and the histogram used to represent $d\hat{f}(x)/dx$. By choosing appropriate b_i's ($i = 1, 2, \ldots, n$), the absolute difference $|f(x) - \hat{f}(x)|$ can be made as small as desired; thus effecting a suitable approximation. That is, an appropriate criterion to determine the suitability of a specific

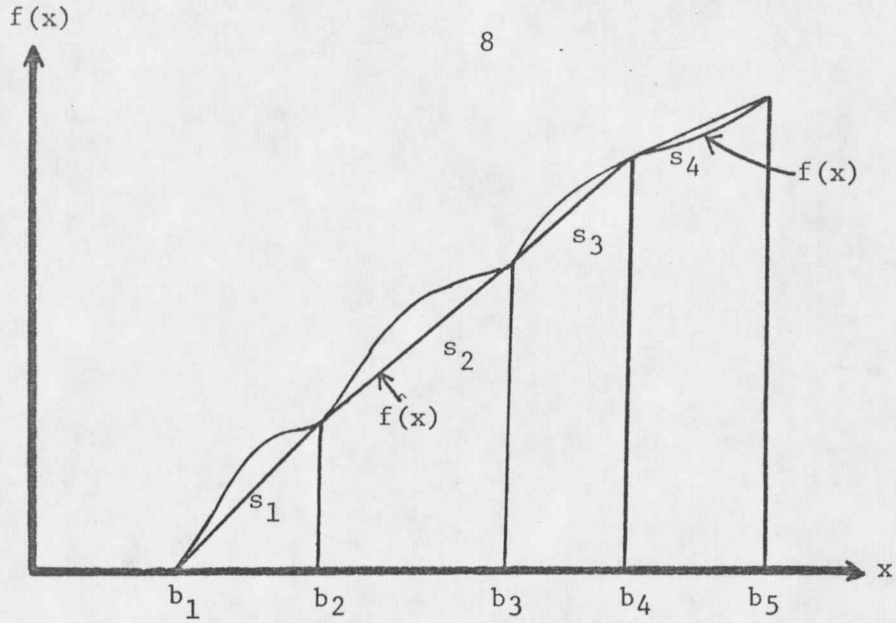

Figure 2-1. Piecewise Linear Approximation of an Arbitrary Concave-Convex Function

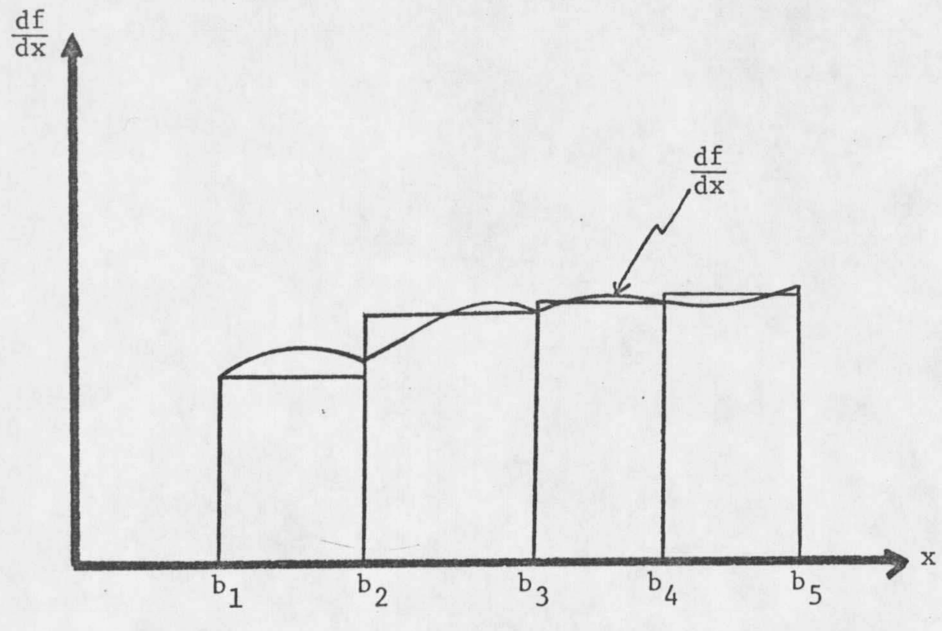

Figure 2-2. Histogram Approximation of df/dx

approximation is to solve for the break points by minimizing the absolute value of the total difference between the linear approximating lines and the true curve.

For example, given $f(x) = x^2$. Let x be bounded by $0 \leq x \leq 10$ and let the curve be approximated by two linear functions. Then, the break point can be determined by minimizing the following function (see Figure 2-3):

$$\text{Minimize} \quad z^{(1)} = \underbrace{\int_0^{x_1} |s_1 x - x^2| dx}_{z_1} + \underbrace{\int_{x_1}^{x_2=10} |s_1 x_1 + s_2(x - x_1) - x^2| dx}_{z_2}$$

where x_1 is the break point, and s_1 and s_2 are the slopes of the corresponding linear functions.

By definition, $s_1 = f(x_1)/x_1 = x_1^2/x_1 = x_1$. Then by convexity, $z_1 = \int_0^{x_1} |s_1 x - x^2| dx = \int_0^{x_1} (x_1 x - x^2) dx = x_1^3/6$. Therefore, minimize

$$z^{(1)} = x_1^3/6 + \int_{x_1}^{x_2=10} (s_1 x_1 + s_2 x - s_2 x_1 - x^2) dx$$

$$= -\frac{1}{2} x_1^3 + 10 x_1^2 + \frac{1}{2} s_2 x_1^2 - 10 s_2 x_1 + 50 s_2 - 1000/3 \quad \text{-------------(2-1)}$$

Since $s_2 = \dfrac{f(x_2) - f(x_1)}{x_2 - x_1} = \dfrac{100 - x_1^2}{10 - x_1} = 10 + x_1$, substitute into equation 2-1 to obtain the following equation:

$$z^{(1)} = 5x_1^2 - 50x_1 + 500/3 \quad \text{--------------------------------------(2-2)}$$

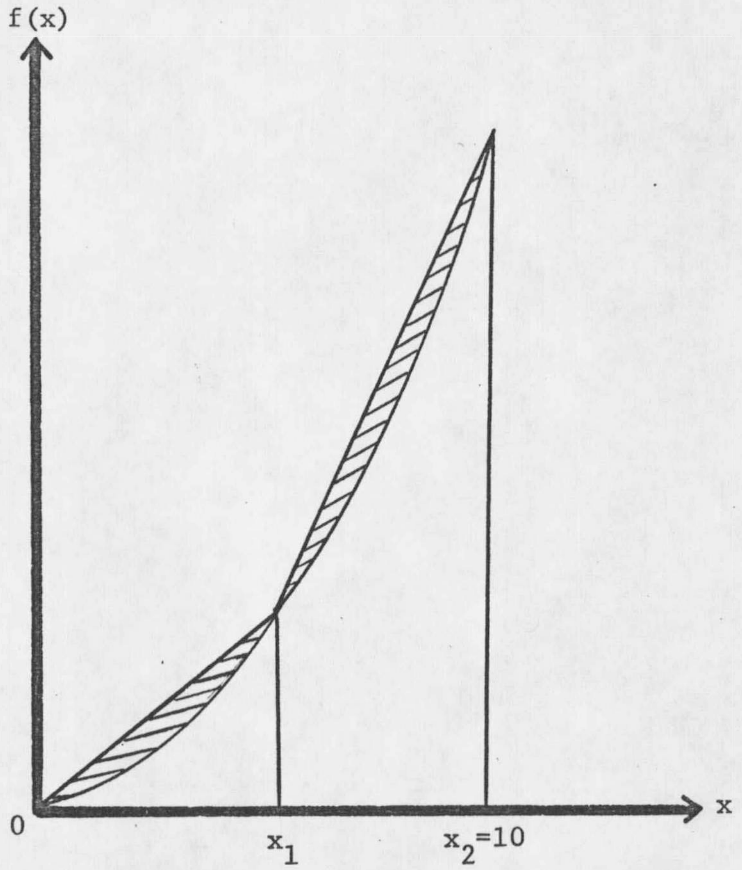

Figure 2-3. Piecewise Linear Approximation of x^2 Using One Break Point

Take the derivative with respect to x_1 and set it equal to zero to obtain the minimum. Thus,

$$10x_1 - 50 = 0 \quad \text{------------------------} (2\text{-}3)$$

$\therefore x_1 = 5$, $s_2 = 10 + x_1 = 15$. Check the second derivative: $\frac{\partial^2 z^{(1)}}{\partial x_1^2} = 10 > 0$. Therefore, $x_1 = 5$ is the desired minimum. The break point is at $x = 5$, and the two approximating linear functions are $5x$ for the first segment and $-50 + 15x$ for the second segment.

For the same function above, when three linear functions are used to approximate the curve, the objective becomes to minimize the following function (see Figure 2-4):

$$\text{Minimize } Z^{(2)} = \underbrace{\int_0^{x_1} |s_1x - x^2|dx}_{z_1} + \underbrace{\int_{x_1}^{x_2} |s_1x_1 + s_2(x - x_1) - x^2|dx}_{z_2}$$

$$+ \underbrace{\int_{x_2}^{x_3=10} |s_1x_1 + s_2(x_2 - x_1) + s_3(x - x_2) - x^2|dx}_{z_3}$$

A straightforward way to minimize the above function is simply to take the derivatives with respect to the unknowns, x_1 and x_2, and set them equal to zero. Thus, the following simultaneous nonlinear equations are obtained.

$$-x_2^2 + 2x_1x_2 = 0 \quad \text{------------------------} (2\text{-}4)$$

$$-x_1x_2 + 0.5x_1^2 - 50 + 10x_2 = 0 \quad \text{------------------------} (2\text{-}5)$$

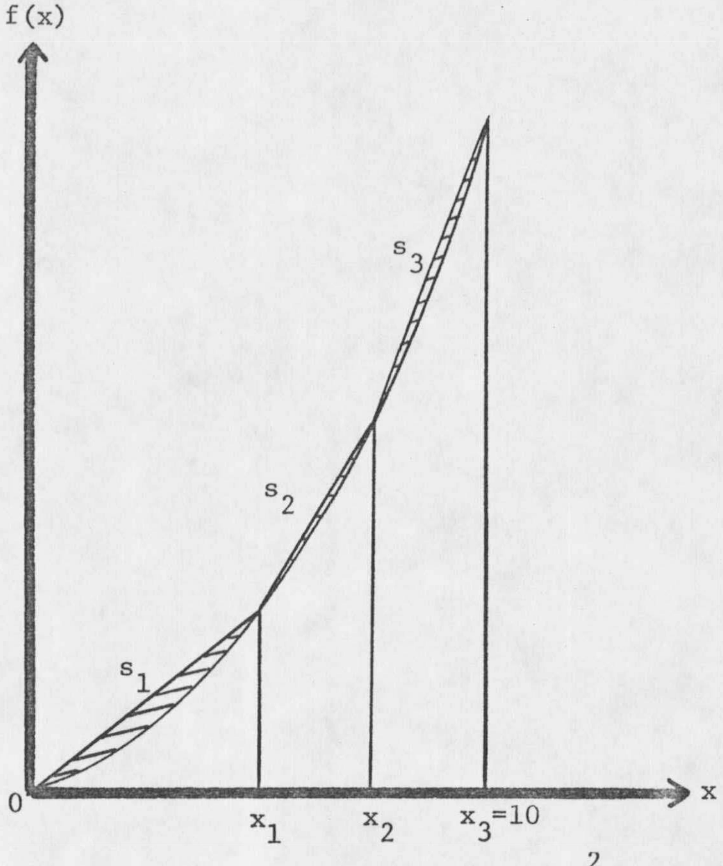

Figure 2-4. Piecewise Linear Approximation of x^2 Using Two Break Points

Solve the above equations to obtain $x_1 = 0.5x_2$, $x_2 = 20$, or 6.66. $X_2 =$ 20 exceeds 10, the bounding value of x. Hence, 6.66 is chosen and $x_1 =$ 3.33.

It may be noted that when the given function is more complicated than the parabola x^2, simultaneous nonlinear equations which are difficult to solve will be obtained. An alternative way to find the minimum is therefore desirable.

Assume that the minimum value of $z_1 + z_2$ is known for $x_2 = K$, where K is assumed to be the optimal second break point. Then, in order to minimize $z^{(2)}$, (the minimum of $z_1 + z_2$)+ $\int_{x_2=K}^{x_3=10} |s_1x_1 + s_2(x_2 - x_1) + s_3(x - x_2) - x^2|dx$ must be minimized. If in the previous solution, $x_2 =$ K is used instead of 10, equations 2-2 and 2-3 will be of the form:

$$z^{(1)} = \frac{K^3}{6} - \frac{K^2}{2}x_1 + \frac{K}{2}x_1^2 \quad \text{------}(2\text{-}6)$$

$$Kx_1 - K^2/2 = 0 \quad \text{------}(2\text{-}7)$$

Solving for x_1 in terms of K, then $x_1 = 0.5$ K leads to $s_2 = K + x_1 =$ 1.5K. Therefore, the minimum value of z_1 and z_2 in terms of K can be determined from equation 2-6 as the

$$\text{minimum of } z^{(1)} = \frac{K^3}{6} - \frac{K^2}{2}(0.5K) + \frac{K}{2}(0.5K)^2 = 0.042K^3, \text{ and}$$

$$z^{(2)} = 0.042K^3 + \int_{x_2=K}^{x_3=10} |s_1x_1 + s_2(x_2 - x_1) + s_3(x - x_2) - x^2|dx \quad \text{------}(2\text{-}8)$$

Similarly $s_3 = \dfrac{100 - x_2^2}{10 - x_2} = 10 + K$. Equation 2-8 can be written as

$$z^{(2)} = -0.125K^3 + 5K^2 - 50K + 166.7$$

Take the derivative with respect to K and set it equal to zero to obtain the minimum. Hence,

$$-0.375K^2 + 10K - 50 = 0$$

$$K = 6.66 \text{ or } 20$$

K = 20 exceeds 10, the bounding value of K, and hence must be discarded. Since $\dfrac{\partial^2 z^{(2)}}{\partial K^2} = 5 > 0$ at K = 6.66, $z^{(2)}$ has been minimized. Therefore, $x_2 = K = 6.66$, $x_1 = 0.5K = 3.33$. The break points are at $x_1 = 3.33$ and $x_2 = 6.66$, and are the same as the values obtained by using the previous direct method.

In like manner, if four linear approximating lines are desired, the same technique can be applied to x^2 (see Figure 2-5). The objective now is to minimize $z^{(3)}$

$$z^{(3)} = \underbrace{\int_0^{x_1} |s_1 x - x^2|\,dx}_{z_1} + \underbrace{\int_{x_1}^{x_2} |s_1 x_1 + s_2(x - x_1) - x_2|\,dx}_{z_2}$$

$$+ \underbrace{\int_{x_2}^{x_3} |s_1 x_1 + s_2(x_2 - x_1) + s_3(x - x_2) - x^2|\,dx}_{z_3}$$

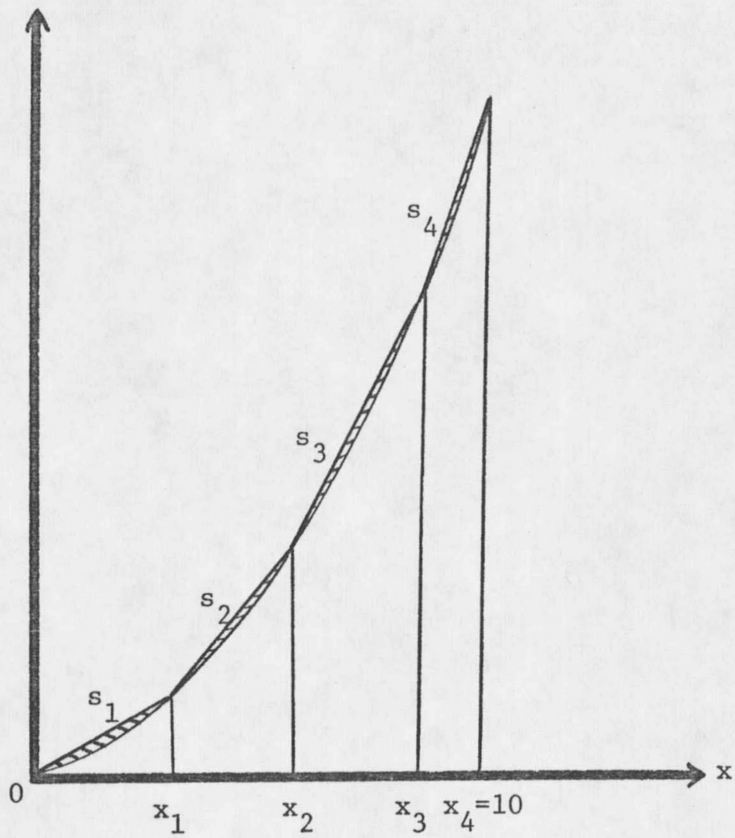

Figure 2-5. Piecewise Linear Approximation of x^2 Using Three Break Points

$$+ \underbrace{\int_{x_3}^{x_4=10} \left| x_1 x_1 + s_2(x_2 - x_1) + s_3(x_3 - x_2) + s_4(x - x_3) - x^2 \right| dx}_{z_4}$$

First, using the direct method by taking the derivatives with respect to the unknowns, x_1, x_2, x_3, and setting them equal to zero yields the following simultaneous nonlinear equations:

$$- \frac{1}{2}x_2^2 + x_1 x_2 = 0 \quad \text{--(2-9)}$$

$$-x_1 x_2 + \frac{1}{2}x_1^2 - \frac{1}{2}x_3^2 + x_2 x_3 = 0 \quad \text{--(2-10)}$$

$$-x_2 x_3 + \frac{1}{2}x_2^2 - 50 + 10x_3 = 0 \quad \text{--(2-11)}$$

Solve the above equations for $x_1 = 0.5x_2$, $x_2 = \frac{2}{3}x_3$, and $x_3 = 15$, or 7.5. 15 exceeds 10, the bounding value of x. Hence, $x_3 = 7.5$ is chosen, $x_2 = 5$ and $x_1 = 2.5$.

Now, solve for the minimum by the alternative step-by-step method. If x_3 of the previous solution is some fixed value less than 10, then the value of K in terms of x_3 is again obtained from the minimum of $z_1 + z_2 + \int_{x_2=K}^{x_3} |s_1 x_1 + s_2(x_2 - x_1) + s_3(x - x_2) - x^2| dx$. Similarly $s_3 = x_3 + K$, and the above expression can be transformed to:

$$-0.125K^3 + 0.5K^2 x_3 - 0.5Kx_3^2 + 0.167x_3^3 \quad \text{---------------------------(2-12)}$$

Take the derivative with respect to K and set it equal to zero:

$$-0.375K^2 - 0.5x_3^2 + Kx_3 = 0 \quad \text{-----------------------------------} (2\text{-}13)$$

Solve for K in terms of x_3 to obtain $K = 0.666x_3$. The minimum value of (minimum of $z_1 + z_2$)+ $\int_{x_2=K}^{x_3} |s_1x_1 + s_2(x_2 - x_1) + s_3(x - x_2) - x^2| dx$ equals $0.020x_3^3$. Thus, $z^{(3)}$ can be written as follows:

$$z^{(3)} = 0.020x_3^3 + \int_{x_3}^{x_4=10} |s_1x_1 + s_2(x_2 - x_1) + s_3(x_3 - x_2) + s_4(x - x_3) - x^2| dx$$

$$= -0.147x_3^3 + 5x_3^2 - 50x_3 + 500/3$$

with $x_2 = K = 0.666x_3$ and $x_1 = 0.5K = 0.5x_2 = 0.333x_3$. Take the derivative with respect to x_3 and set it equal to zero to obtain the minimum. Hence,

$$-0.441x_3^2 + 10x_3 - 50 = 0 \quad \text{-----------------------------------} (2\text{-}14)$$

$x_3 = 7.48$, $x_2 = 0.666x_2 = 4.98$, $x_1 = 0.5x_2 = 0.333x_3 = 2.49$.

These values are almost the same as those obtained from the direct method. The small difference is due to the round-off error in calculation.

From the preceding example, one may conclude that if an additional approximating linear function is required, one can simply add one more segment and obtain new break points by modifying the previous answer. Thus, the refinement process is suitably carried out by the dynamic programming approach. The original problem of n variables is transformed to n problems of one variable only. And at each stage, a problem of one variable is optimized.

The solution technique as illustrated on the previous page can be generalized to any single-variable function f(x) exhibiting convexity or concavity. An objective function containing such terms as e^x, log x, etc., can be very difficult to handle. However, any single-variable function can always be expanded in a power series (1). That is to say, any single-variable function can always be expressed in the form of a polynomial and hence can be easily integrated and differentiated. Thus, applying the approach of the preceding example, a linear or nonlinear equation of a single variable in each stage can always be obtained and the solution easily identified.

The solution technique presented so far is actually nothing more than a general approach to find the best approximation under the criterion that the total absolute difference between the approximating function and the given function be minimized.

The general approach when n linear functions are used to approximate the given function, f(x), is summarized as follows:

1. Check that the given function is convex within the region of the desired approximation.
2. Minimize $z^{(1)}$ where two linear approximating functions are to be used.

$$z^{(1)} = \underbrace{\int_0^{x_1} |s_1 x - f(x)|dx}_{z_1} + \underbrace{\int_{x_1}^{x_2} |s_1 x_1 + s_2(x - x_1) - f(x)|dx}_{z_2}$$

where x_1 is the break point, s_1 and s_2 are the slopes of the corresponding linear approximating functions, and x_2 is the assumed known optimal second break point. Or, minimize

$$z^{(1)} = \underbrace{\int_0^{x_1} |a_1 + s_1 x - f(x)|\,dx}_{z_1} + \underbrace{\int_{x_1}^{x_2} |a_1 + s_1 x_1 + s_2(x - x_1) - f(x)|\,dx}_{z_2}$$

when $f(x)$ has the vertical intercept a_1.

3. Solve for the value of x_1 in terms of x_2, and express the minimum value of $z_1 + z_2$ in terms of x_2.
4. Formulate a new objective function as the sum of the minimum of $z_1 + z_2$ and the total absolute difference of the third interval of approximation. Then, apply the same minimization technique to solve for the value of the current unknown variable, x_2, in terms of x_3.
5. Formulate another new objective function as the sum of the minimum of $z_1 + z_2 + z_3$ and the total difference of the fourth interval of approximation. Apply the same technique to solve for the value of the current unknown, x_3, in terms of x_4.
6. Repeat the formulation and minimization of the new objective function. Each time include one more interval of approximation. Thus, the relationship between x_1, x_2, ..., etc. is established step-by-step. When the relationship between x_{n-1} and

x_n (x_n = K, the boundary value) is established, the problem is then solved in its entirety.

This solution procedure developed for separable convex functions can also be modified to handle the case of separable nonconvex functions. It is known that linear interpolation always gives overestimation or underestimation when the given function is convex or concave, respectively. If the interval of "convexity" and "concavity" for a given function within the boundary of desired approximation can be identified, the difficulty in determining the absolute value is removed, and the same procedure can be applied to obtain the optimal break points.

The solution procedure suggested then is to identify the interval of "convexity" by setting second derivatives greater-than or equal-to zero to obtain the bound value and similarly to identify the interval of "concavity". Then, the same technique of solving for the optimal break points can be applied to each convex and concave part of the function. This is illustrated via the following example:

Given $f(x) = \frac{1}{6}x^3 - \frac{3}{2}x^2$, where $0 \leq x \leq 10$

$$\frac{\partial f(x)}{\partial x} = \frac{1}{2}x^2 - 3x$$

$$\frac{\partial^2 f(x)}{\partial x^2} = x - 3$$

$$\frac{\partial f^2(x)}{\partial x^2} = x-3 \geq 0 \qquad \text{if } x \geq 3$$

$$\frac{\partial f^2(x)}{\partial x^2} = x-3 \leq 0 \qquad \text{if } x \leq 3$$

Accordingly, the interval where convexity holds is $3 \leq x \leq 10$ and the interval where concavity holds is $0 \leq x \leq 3$.

The problem is then divided into two new problems:

1. Approximate $f(x) = \frac{1}{6}x^3 - \frac{3}{2}x^2$ by linear functions on the interval $0 \leq x \leq 3$.
2. Approximate $f(x) = \frac{1}{6}x^3 - \frac{3}{2}x^2$ by linear functions on the interval $3 \leq x \leq 10$.

Discussion and Adaptation of the Solution Technique to Separable Convex Programming

The development of the solution technique stems from the interative, dynamic programming approach; that is, to start with a small portion of the problem and to find the optimal solution for this smaller problem. It then gradually enlarges the problem, finding the current optimal solution from the previous one, until the original problem is solved in its entirety.

Any single-variable function can always be expanded in a power series of the form $a_0x^n + a_1x^{n-1} + . . . + a_{n-1}x + a_n$, a polynomial approximation. Although the solution approach developed can be applied to any single-variable function, it does have certain limitations.

First, it can be applied only to functions of a single variable. Therefore, when a separable function is encountered, one can apply this technique to each single-variable subfunction, separately. When a nonseparable function is encountered, only in very rare cases can one apply this approach. However, it might be possible to use a suitable transformation of variables to transform the original non-separable function into a separable function of some new variables. For example, let $y = x_1 + x_1x_2 + x_2$. This is a non-separable function since it contains a cross product term. However, if we let $x_1 = u + w$, and $x_2 = u - w$, then $y = u + w + u^2 - w^2 + u - w = 2u + u^2 - w^2$ has been transformed into a separable function of the new variables u and w. In many cases this kind of transformation does not work at all. Therefore, the technique developed is generally limited to use with separable functions.

The second pitfall in this approach is the possible necessity of solving for the roots of a nonlinear equation of polynomial form within the region $0 \leq x \leq K$. But, several methods exist for finding the roots of nonlinear equations and are presented in Appendix II.

Next, consider the application of optimal break points to the separable convex programming problem:

$$\text{Maximize} \quad \sum_{j=1}^{n} f_j(x_j)$$

subject to $\sum_{j=1}^{n} a_{ij}x_j \leq b_i$ for $i = 1, 2, \ldots, m$;

and $x_j \geq 0$ for $j = 1, 2, \ldots, n$;

where $f_j(x_j)$ is a concave function for $j = 1, 2, \ldots, n$. Let m_j be the number of break points for the approximating function $\hat{f}_j(x_j)$ (see Figure 2-6), and $b_{j1}, b_{j1} + b_{j2}, \ldots, \sum_{k=1}^{m_j} b_{jk}$ be the values of x_j at which the break points occur, and where $\sum_{k=1}^{m_j} b_{jk}$ is the upper bound of the value of x_j. Also let s_{jk} $(k = 1, 2, \ldots, m_j)$ be the slope of the piecewise linear function when $\sum_{\ell=1}^{k-1} b_{j\ell} < x_j < \sum_{\ell=1}^{k} b_{j\ell}$. Then,

$$x_{jk} = \begin{cases} 0 & \text{if } x_j \leq \sum_{j=1}^{k-1} b_{j\ell} \\ x_j - \sum_{\ell=1}^{k-1} b_{j\ell} & \text{if } \sum_{\ell=1}^{k-1} b_{j\ell} \leq x_j \leq \sum_{\ell=1}^{k} b_{j\ell} \\ \sum_{\ell=1}^{k} b_{j\ell} - \sum_{\ell=1}^{k-1} b_{j\ell} & \text{if } x_j \geq \sum_{\ell=1}^{k} b_{j\ell} \end{cases}$$

for $k = 1, 2, \ldots, m_j$ and $j = 1, 2, \ldots, n$

It follows that

$$0 \leq x_{jk} \leq b_{jk}$$

and

$$x_j \equiv x_{j1} + x_{j2} + \ldots + x_{jmj}.$$

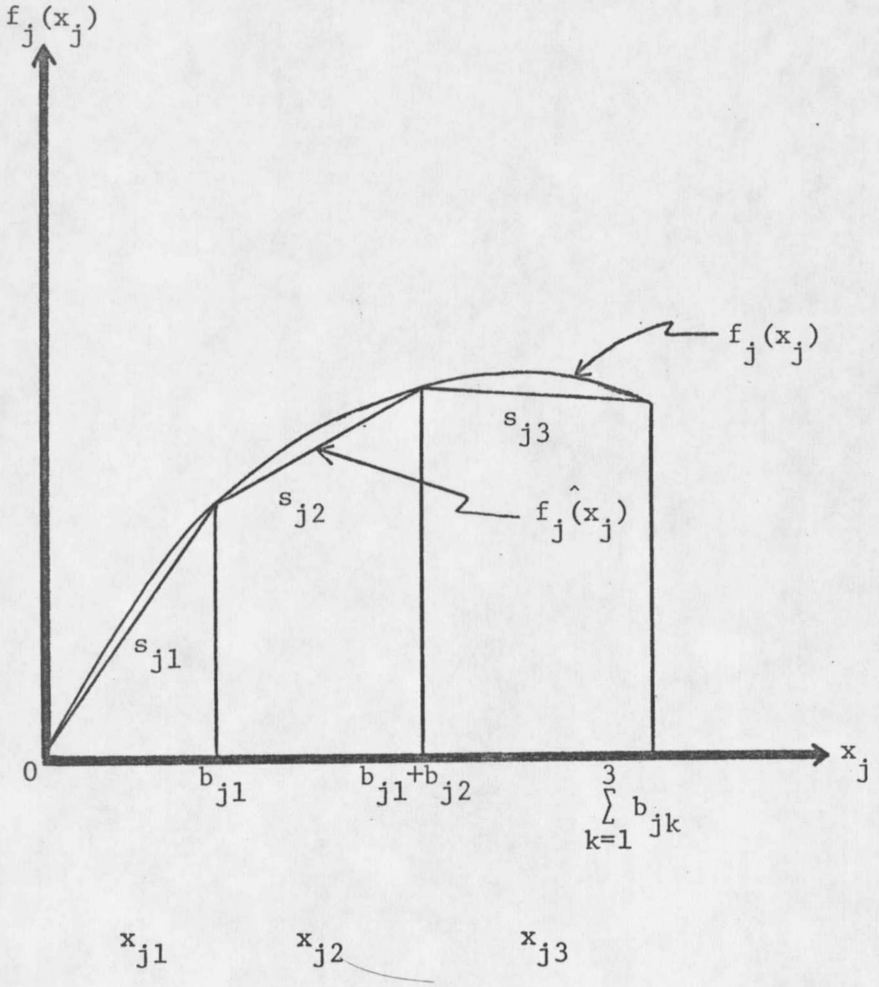

Figure 2-6. Approximation of $f_j(x_j)$ by $\hat{f_j}(x_j)$

The piecewise linear approximating function $\hat{f}_j(x_j)$ can be written as $\hat{f}_j(x_j) = s_{j1}x_{j1} + s_{j2}x_{j2} + \ldots + s_{jmj}x_{jmj}$. Thus, the original problem can be reformulated as:

Maximize $\sum_{j=1}^{n} \left(\sum_{k=1}^{m_j} s_{jk}x_{jk} \right)$

subject to $\sum_{j=1}^{n} a_{ij} \left(\sum_{k=1}^{m_j} x_{jk} \right) \leq b_i$ for $i = 1, 2, \ldots, m$;

and $x_{jk} \leq b_{jk}$ for $j = 1, 2, \ldots, n$; $k = 1, 2, \ldots, m_j$;

and $\sum_{k=1}^{m_j} x_{jk} \geq 0$ for $j = 1, 2, \ldots, n$.

This transformed problem can be solved as a linear programming problem using the usual simplex method. If $(x^*_{11}, x^*_{12}, \ldots, x^*_{nm_n})$ is the optimal solution to this problem, then $x_1 = \sum_{k=1}^{m_1} x^*_{1k}$, $x_2 = \sum_{k=1}^{m_2} x^*_{2k}$, $\ldots$, $x_n = \sum_{p=1}^{m_n} x^*_{nk}$ must be the optimal solution to the approximate form of the original problem (13).

An alternative way to solve the separable programming problem was developed by Miller (18). Given the function sketched in Figure 2-7, let the range of x be $0 \leq x \leq K$ and suppose that m break points x_m have been chosen, where $x_0 = 0$, $x_1 < x_2 < \ldots < x_m = K$. When $x_k \leq x \leq x_{k+1}$,

$$\hat{f}(x) = f(x)_k + \frac{f(x)_{k+1} - f(x)_k}{x_{k+1} - x_k}(x - x_k).$$

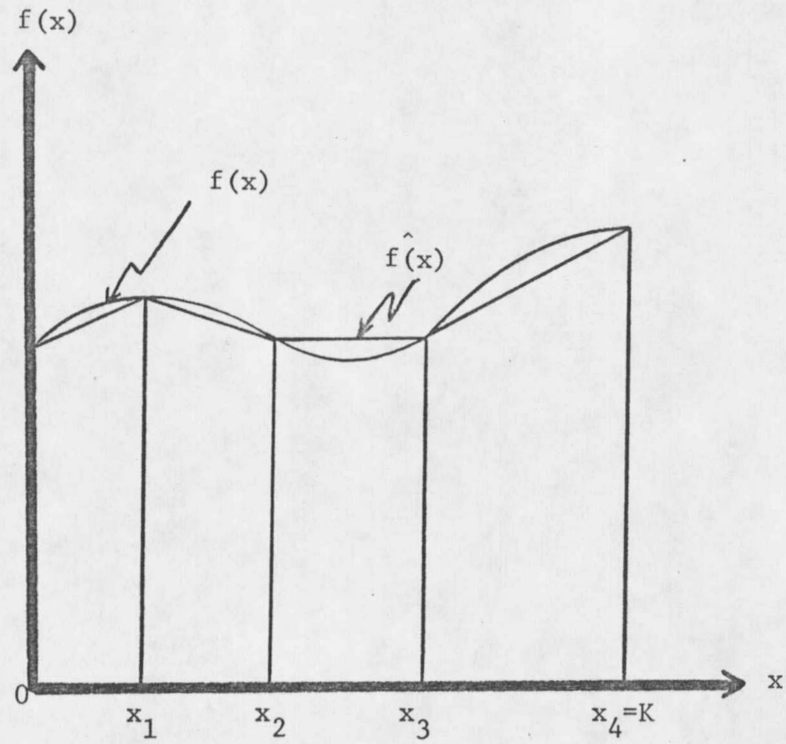

Figure 2-7. Piecewise Linear Approximation of $f(x)$ by $\hat{f}(x)$

Any x in the interval $x_k \leq x \leq x_{k+1}$ can be written as $x = \lambda x_{k+1} + (1 - \lambda)x_k$ for some λ, where $0 \leq \lambda \leq 1$. Subtracting x_k from both sides,

$$x - x_k = \lambda x_{k+1} + x_k - \lambda x_k - x_k = \lambda(x_{k+1} - x_k),$$

so that

$$\hat{f}(x) = f(x)_k + \frac{f(x)_{k+1} - f(x)_k}{x_{k+1} - x_k}\lambda(x_{k+1} - x_k) = \lambda f(x)_{k+1} + (1 - \lambda)f(x)_k.$$

Let $\lambda = \lambda_{k+1}$, $1 - \lambda = \lambda_k$

then, when $x_k \leq x \leq x_{k+1}$ there exists a unique λ_k and λ_{k+1} such that $x = \lambda_k x_k + \lambda_{k+1}x_{k+1}$, and $\hat{f}(x) = \lambda_k f(x)_k + \lambda_{k+1}f(x)_{k+1}$ with $\lambda_k + \lambda_{k+1} = 1$ and $\lambda_k, \lambda_{k+1} \geq 0$. Hence for $0 \leq x \leq K$,

$$x = \sum_{k=0}^{m}\lambda_k x_k, \; \hat{f}(x) = \sum_{k=0}^{m}\lambda_k f(x)_k, \; \sum_{k=0}^{m}\lambda_k = 1, \text{ and } \lambda_k \geq 0.$$

In addition, no more than two of the λ_k may be positive and they must also be adjacent.

Hence, for a separable programming problem, after the linearization solution technique is applied to each single-variable function, the problem can be transformed into a linear programming problem by the procedure illustrated as follows.

Suppose that for each single variable, a sequence of break points has been chosen. Assume for the moment that the same number of break points m has been chosen for each variable. Then, x_j can be expressed