**MULTI-DIMENSIONAL
PIECEWISE POLYNOMIAL
CURVE FITTING**

by

Philip B. Zwart

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

This research was supported in part by the Atomic Energy Commission under Research Contract No. A(111-1)-1493 and by the Department of Defense under Themis Grant No. F44620-F44620-69-C-0116.

**Report No. COO-1493-34
(CSSE-705)**

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

ABSTRACT

Any set of hyperplanes partitions R^n into polygonal pieces. Linearly partitioned piecewise polynomials (LPPP) result when polynomials on each of these pieces are put together in a smooth manner. There is presented an LPPP canonical form which leads to computationally feasible multi-dimensional piecewise polynomial curve fitting.

I. INTRODUCTION

The problem of explicitly describing an n -dimensional surface $z = F(x)$, x an n -vector, which approximately fits some data points (x_k, z_k) arises in mathematical modeling and data analysis with computers. The data points may be the results of a computer simulation (or experimental results, surveys, economic data, etc.) of a function whose values must play a role (where smoothness is required) in a more involved simulation. One way of attempting to describe such a surface is to piece together polynomial functions. This paper is concerned with the process of selecting a piecewise polynomial function which approximates the data.

Section II first describes what the pieces of R^n are to be like. Any set of hyperplanes naturally partitions R^n into polygons. Then the concept of a linearly partitioned piecewise polynomial of degree s is defined. This corresponds to having polynomials of degree $\leq s$ defined on each of the polygons in such a way that the resulting function in R^n possesses continuous derivatives of order $s-1$. Non-degeneracy of a partition is also defined. Section III shows that, for non-degenerate partitions, the piecewise polynomial function must take on an especially simple canonical form. Section IV contains a few comments on how the canonical form is useful in finding the desired surface $z = F(x)$.

II. DEFINITIONS:

A linear partition of R^n is a set of vectors $(a_1, b_1), \dots, (a_m, b_m)$, where the a_i 's are n -component vectors and the b_i 's are real numbers. Associated with a linear partition are the closed polygonal regions, R_I , I a subset of $1, \dots, m$, where R_I is defined by

$$R_I = \{x | a_i \cdot x \geq b_i, i \in I, a_i \cdot x \leq b_i, i \notin I\}$$

and is non-empty.

A linear partition is non-degenerate if for every linearly dependent $a_{i_1}, \dots, a_{i_k}$ we have

$$\text{rank} \begin{pmatrix} a_{i_1} \\ \vdots \\ a_{i_k} \end{pmatrix} < \text{rank} \begin{pmatrix} a_{i_1}, b_{i_1} \\ \vdots \\ a_{i_k}, b_{i_k} \end{pmatrix} .$$

Nondegeneracy says the intersection of any k of the hyperplanes $a_i \cdot x = b_i$, which partition R^n , must have dimension $n-k$. In particular, the intersection of any $n+1$ of these hyperplanes is empty.

Two regions R_I and R_J are called k -neighbors if either $I=J \cup \{k\}$ or $J=I \cup \{k\}$.

Linear partitions possess some properties which will be of use in the following section. We present them here.

Proposition 1: For any R_I and R_J , there exist a sequence $R_{I_1}, \dots, R_{I_\ell}$ such that R_{I_j} and $R_{I_{j+1}}$ are k_j -neighbors, $j=1, \dots, \ell-1$ and $R_{I_1} = R_I, R_{I_\ell} = R_J$.

Proof: Let U be the union of all R_K 's which can be connected to R_I by such chains. U is the union of a finite number of closed sets, so it is closed. We need only show that U is an open set.

Pick $x \in U$. Set U_x equal to the union of all R_I 's which contain x . x is not contained in the union of the R_I 's not containing x , which is a closed set containing $R^n - U_x$. So x is an interior point of U_x . We will show that $U_x \subset U$. Suppose $x \in R_K \subset U$. Pick $R_L \subset U_x$. Since $x \in R_K \cap R_L$, we must have $a_i \cdot x = b_i, \forall i \in (K-L) \cup (L-K)$. Suppose $K-L = \{i_1, \dots, i_r\}$, $L-K = \{j_1, \dots, j_s\}$. Then set $I_j = K - \{i_1, \dots, i_j\}$, $j=0, \dots, r$, and $I_{r+j} = (K - \{i_1, \dots, i_r\}) \cup \{j_1, \dots, j_j\}$, $j = 1, \dots, s$. The R_{I_j} , $j = 0, \dots, r+s$ are non-empty (because each contains x) and $R_{I_0}, \dots, R_{I_{r+s}}$ is a chain of neighbors connecting R_K to R_L . Thus, $R_L \subset U, \forall R_L \subset U_x$. And x being an interior point of $U_x \subset U$, is an interior point of U . U being open and closed must equal R^n . So $R_J \subset U$.

Proposition 2: If the partition is non-degenerate, then

$R_I \neq \emptyset$ implies R_I contains interior points i.e. R_I has dimension n .

Proof: Suppose $R_I \neq \emptyset$. Pick $\hat{x} \in R_I$. Let $J = \{i | a_i \cdot \hat{x} = b_i\}$.

Set $x_J = \hat{x}$. If, for some $k \in J$, there exists $x \in R_I$ such that $a_i \cdot x = b_i$ iff $i \in J - \{k\}$, then replace J by $J - \{k\}$ and x_J by this new x . If this procedure eventually exhausts J , then the last x is an interior point of R_I .

Suppose this procedure does not exhaust J. Then it eventually occurs that

a) $a_j \cdot x_J = b_j, j \in J$ and

b) for $k \in J \setminus I$ $\min a_k \cdot x$

subject to $a_j \cdot x = b_j, j \in J - \{k\}$

is equal to $a_k \cdot x_J = b_k,$

c) for $k \in J - I$

$\max a_k \cdot x$

subject to $a_j \cdot x = b_j, j \in J - \{k\}$

is equal to $a_k \cdot x_J = b_j.$

Since at least one of $J \setminus I$ and $J - I$ is nonempty, we must have that for some $k \in J$, there exist multipliers of $\alpha_j, j \in J$ such that $\sum_{\substack{j \in J \\ j \neq k}} \alpha_j a_j = a_k$. Thus, the set of a_j 's, $j \in J$ is linearly dependent and $a_j \cdot x_J = b_j, j \in J$. This contradicts the non-degeneracy of the partition. Thus, J must be exhausted and the proposition is proved.

Proposition 3: $a_j \cdot x = b_j$

for some $x \in R_I$ implies that

$\dim \{x \mid a_j \cdot x = b_j\} \cap R_I = n - 1$

Proof: Use an argument like that used in the proof of proposition 2 with the additional requirement that j never be removed from J.

A function F is said to be a linearly partitioned piecewise polynomial (LPPP) of degrees iff there is a linear

partition of R^n such that

- 1) F is continuously differentiable of order $s-1$.
- 2) F_I (F restricted to R_I) is a polynomial of degree $\leq s, \forall R_I$.
- 3) F_I is a polynomial of degree s for some R_I .
- 4) If R_I and R_J are k -neighbors then

$$F_I = F_J \text{ in } R_I \cap R_J.$$

F is a basic LPPP of degree s iff

$$F(x) = \text{sgn}(a_1 \cdot x - b_1) \alpha_1 (a_1 \cdot x - b_1)^s$$

for some nonzero α_1 , some n -vector a_1 , and some real number b_1 .

Notice that the above function is an LPPP because with the linear partition consisting of the single vector (a_1, b_1)

$$1) \frac{\partial^{s-1}}{\partial x_1^{j_1} \dots \partial x_n^{j_n}} (\text{sgn}(a_1 \cdot x - b_1) \alpha_1 (a_1 \cdot x - b_1)^s)$$

$$= s! a_1^{j_1} \dots a_n^{j_n} (a_1 \cdot x - b_1) \alpha_1 (a_1 \cdot x - b_1)$$

$$2)+3) \quad F_\phi(x) = -\alpha_1 (a_1 \cdot x - b_1)^s, \text{ and}$$

$$F_{\{1\}}(x) = \alpha_1 (a_1 \cdot x - b_1)^s.$$

$$4) \quad F_\phi(x) = F_{\{1\}}(x) = 0, \text{ if } x \in R_\phi \cap R_{\{1\}}$$

III. CANONICAL FORM:

An LPPP (with partition $(a_1, b_1), \dots, (a_m, b_m)$) is said to be in canonical form, if it is expressed as

$$F(x) = F_0 + \sum_{i=1}^m F_i$$

where F_0 is a polynomial and F_i is a basic LPPP with partition (a_i, b_i) .

Theorem: If F is an LPPP of degree s , with non-degenerate partition, then F can be expressed in canonical form.

Proof: We first prove a few lemmas.

Lemma 1: If R_I and R_J are k -neighbors, then

$$F_I(x) = F_J(x) \quad \forall x \ni a_k \cdot x = b_k.$$

Proof: By proposition 3, $\dim R_I \cap R_J = n-1$.

Since $F_I(x) = F_J(x) \quad x \in R_I \cap R_J$, an $n-1$ dimensional subset of $\{x | a_k \cdot x = b_k\}$, and F_I and F_J are polynomials, we must have $F_I(x) = F_J(x), \quad \forall x \ni a_k \cdot x = b_k$.

Lemma 2: If R_I and $R_{I-\{k\}}$ are non-empty, then there exists a real number α_k^I such that $F_I(x) - F_{I-\{k\}}(x) = 2\alpha_k^I (a_k \cdot x - b_k)^s$.

Proof: Let $y_1, \dots, y_n$ be a new orthogonal coordinate system which is a linear transformation of the $x_1, \dots, x_n$ coordinate system such that $y_1 = a_k \cdot x - b_k$.

Now,

i) $F_I(0, y_2, \dots, y_n) = F_{I-\{k\}}(0, y_2, \dots, y_n)$, follows from lemma 1, and

$$\text{ii) } \frac{\partial^r}{\partial y_1^{i_1} \dots \partial y_n^{i_n}} F_I(0, y_2, \dots, y_n) = \frac{\partial^r}{\partial y_1^{i_1} \dots \partial y_n^{i_n}} F_{I-\{k\}}(0, y_2, \dots, y_n)$$

for $i_1 + \dots + i_n = r \leq s-1$, follows from the fact that F is continuously differentiable of order $s-1$.

Expansion of F_I and $F_{I-\{k\}}$ in Taylor series about any point in $R_I \cap R_{I-\{k\}}$ and use of i) and ii) gives

$$F_I(y) - F_{I-\{k\}}(y) = \alpha y_1^s.$$

Taking $\alpha_k^I = \frac{1}{2} \alpha$, we get

$$F_I(x) - F_{I-\{k\}}(x) = 2\alpha_k^I (a_k \cdot x - b_k)^s.$$

Lemma 3: If $k, \ell \in I$ and R_I and $R_{I-\{\ell\}}$ satisfy $R_I \cap R_{I-\{\ell\}} \cap \{x | a_k \cdot x = b_k\} \neq \emptyset$, then $\alpha_k^I = \alpha_k^{I-\{\ell\}}$.

Proof: $a_\ell \cdot x = b_\ell$ implies that

$$F_I(x) = F_{I-\{\ell\}}(x) \text{ and } F_{I-\{k\}}(x) = F_{I-\{k, \ell\}}(x)$$

by lemma 1. Thus,

$$F_I(x) - F_{I-\{k\}}(x) = F_{I-\{\ell\}}(x) - F_{I-\{k, \ell\}}(x) \quad \forall x \text{ s.t. } a_\ell \cdot x = b_\ell.$$

That is, $2\alpha_k^I(a_k \cdot x - b_k) = 2\alpha_k^{I-\{l\}}(a_k \cdot x - b_k), \forall x \ni a_l \cdot x = b_l$.

Since $a_k \neq a_l$ (because $a_k \cdot x = b_k$ and $a_l \cdot x = b_l$ intersect and the partition is non-degenerate), we must have $\alpha_k^I = \alpha_k^{I-\{l\}}$.

Lemma 4: For fixed k , all α_k^I 's are the same.

Proof: α_k^I 's are only defined for $I \in \mathcal{L}$ the set of those I 's for which $k \in I$ and $R_I \cap \{x | a_k \cdot x = b_k\} \neq \emptyset$. Lemma 3 shows that if any two of these R_I 's are l -neighbors then the α_k^I 's are equal. It follows that if two of these, say R_{I_1} and R_{I_2} , are connected by a chain $R_{I_1}, \dots, R_{I_h}, I_j \in \mathcal{L} \ j=1, \dots, h$ where R_{I_j} and $R_{I_{j+1}}$ are l_j -neighbors, $j=1, \dots, h-1$, then

$\alpha_k^{I_1} = \alpha_k^{I_h}$. An argument like that used in the proof of proposition 1 (with $\{x | a_k \cdot x = b_k\}$ replacing R^n) shows that such a chain exists for any $I, J \in \mathcal{L}$. Thus, $\alpha_k^I = \alpha_k^J$ for any $I, J \in \mathcal{L}$. Henceforth, the superscript will be dropped and we will use α_k .

We now turn to the proof of our theorem. Pick a non-empty R_I . Set $F_0(x) = F_I(x) - \sum_{i \in I} \alpha_i(a_i \cdot x - b_i)^S + \sum_{i \notin I} \alpha_i(a_i \cdot x - b_i)^S$

and $F_i(x) = \sum_{i=1}^m \text{sgn}(a_i \cdot x - b_i) \alpha_i(a_i \cdot x - b_i)^S, i = 1, \dots, m$. We claim that $F = F_0 + \sum_{i=1}^m F_i$. Clearly, $F_I(x) = F_0(x) + \sum_{i=1}^m F_i(x) \ \forall x \in R_I$.

Pick any other R_J . By proposition 1 there exist $R_{I_1}, \dots, R_{I_\ell}$ such that $R_{I_1} = R_I, R_{I_\ell} = R_J$ and R_{I_j} and $R_{I_{j+1}}$ are k -neighbors, $j = 1, \dots, \ell$. Using lemmas 2 and 4 we get

$$F_J(x) = F_I(x) + \sum_{k_j \in L^+} 2\alpha_{k_j}(a_{k_j} \cdot x - b_{k_j})^S - \sum_{k_j \in L^-} 2\alpha_{k_j}(a_{k_j} \cdot x - b_{k_j})^S$$

where $L^+ = \{k_j | I_{j+1} = I_j + \{k_j\}\}$ and $L^- = \{k_j | I_{j+1} = I_j - \{k_j\}\}$.

Notice that

- i) if $i \in I \cap J$, then i appears in L^+ the same number of times as it appears in L^- ,
- ii) if $i \in I - J$, then i appears in L^+ one time less often than it appears in L^- , and
- iii) if $i \in J - I$, then i appears in L^+ one more time than it appears in L^- .

i), ii) and iii) lead to

$$(1) \quad F_J(x) = F_I(x) + \sum_{i \in J-I} 2\alpha_i (a_i \cdot x - b_i)^S - \sum_{i \in I-J} 2\alpha_i (a_i \cdot x - b_i)^S.$$

As noted above

$$F_I(x) = F_0(x) + \sum_{i \in I} \alpha_i (a_i \cdot x - b_i)^S - \sum_{i \notin I} \alpha_i (a_i \cdot x - b_i)^S.$$

Substituting in (1) gives

$$\begin{aligned} F_J(x) &= F_0(x) + \sum_{i \in J} \alpha_i (a_i \cdot x - b_i)^S - \sum_{i \notin J} \alpha_i (a_i \cdot x - b_i)^S, \\ &= F_0(x) + \sum_{i=1}^m F_i(x), \quad \forall x \in R_J. \end{aligned}$$

IV. CURVE FITTING:

Suppose we have data points (x_i, y_i) , $i = 1, \dots, r$, where $x_i = (x_{i1}, \dots, x_{in})$, and we wish to determine a function

$F(x)$ such that $F(x_i) \approx y_i$, $i = 1, \dots, r$. We are interested in synthesizing F as a piecewise polynomial function of degree s possessing continuous derivatives of order $s-1$. We could concentrate on functions of the form

$$F(x) = F_0(x) + \sum_{i=1}^m F_i(x),$$

where $F_0(x)$ is a polynomial of degree s and

$F_i(x) = \text{sgn}(a_i \cdot x - b_i) \alpha_i (a_i \cdot x - b_i)^s$. The planes $a_i \cdot x = b_i$ partition R^n into regions over which the pieces of F are defined.

According to section III, if the partition is non-degenerate then $F_0 + \sum_{i=1}^m F_i$ encompasses all possible piecewise polynomials of degree s possessing continuous derivatives of order $s-1$.

We wish to determine the coefficients in $F(x) = F_0(x) + \sum_{i=1}^m F_i(x)$ so that $F(x)$ gives the best fit to the data points. In the discussion below we take "best" to mean that the coefficients are chosen so that $\sum_{i=1}^r (F(x_i) - y_i)^2$ is minimized. (The comments in part 1 are also pertinent in the case when $\max_{i=1, \dots, r} |F(x_i) - y_i|$ is minimized). There are two levels of complication.

1. Fixed Partition. We assume that the partition (the (a_i, b_i) 's) are predetermined. The parameters to be determined are the coefficients in $F_0(x)$ and the α_i 's. These parameters occur linearly in $F(x)$. The problem is merely a linear regression problem.

The number of parameters is equal to $N(s,n)+m$, where $N(s,n)$ is the no. of coeff. in a polynomial of degree s in n -variables.

For example, any piecewise cubic in 3 dimensions which has a non-degenerate linear partition of m planes, involves $14+m$ parameters. For $m = 4$, such a piecewise cubic could have 15 pieces.

2. Free Partition. We assume that the number of planes is pre-determined but that their positions are to be chosen. The parameters to be determined are the coefficients in $F_0(x)$ and the $\alpha_i, a_{ij}, b_i, i = 1, \dots, m, j = 1, \dots, n$. This is a problem in non-linear regression. Notice that the $F_i(x)$ possess continuous derivatives of order $(s-1)$ with respect to the α_i, a_{ij} , and b_i . All higher order derivatives exist except for derivatives of $F_i(x_k)$ of order s with respect to the a_{ij}, b_i when the data point x_k satisfies $a_i \cdot x_k = b_i$. In such cases, it is reasonable to set the derivative equal to zero. (This last comment may be of interest when $s = 0$ or 1 .)

Nonlinear regression techniques, such as appear in [2], can be used on this type problem. The fact that second derivatives of F with respect to the parameters are mostly zero may be of some use.

Once the (a_i, b_i) 's are fixed, the other coefficients are uniquely determined by linear regression. Thus, the sum of squares of the deviations is a well-defined function of the (a_i, b_i) 's. An unconstrained minimization technique,

such as appears in [1], can be used to find a_{ij} , b_i , $i = 1, \dots, m$, $j = 1, \dots, n$ which minimize this objective function. If this involves too many parameters it may be useful to fix the a_{ij} 's and only regress on the b_i 's.

Suppose $s = 3$, $n = 3$, and $m = 4$. The nonlinear regression problem would involve 30 parameters. The minimization problem would involve 16 parameters. The minimization problem in which the partition planes are predetermined except for translation, would involve 4 parameters.

BIBLIOGRAPHY

- [1] Fletcher, R. and Powell, M.J.D. - "A Rapidly Convergent Descent Method for Minimization", The Computer Journal, 6, 163-168, (1963).
- [2] Marquardt, D.W. - "An Algorithm for Least-Squares Estimation of Nonlinear Parameters", SIAM Journal, 11, 2, 431-441, (1963).