Tools for the PWL approximation of continuous functions

P. JULIAN *AND O. AGAMENNONI†
Departamento de Ingenieria Eléctrica
Universidad Nacional del Sur
Av. Alem 1253, (8000) Bahia Blanca
ARGENTINA
e-mail: pjulian@criba.edu.ar

Abstract: In this paper, a continuous function approximation methodology is proposed using the High Level (HL) Canonical Piecewise Linear (CPWL) expression introduced in [1]. We consider the case where a finite set of measured data is available, and we use a min-max optimality criterion. Also, an application to the characterization of uncertain nonlinear functions is shown. A salient feature of the methodology is that the approximation problem is reduced to a linear programming problem.

Keywords: Piecewise-linear approximations

1 Introduction

Continuous piecewise linear (PWL) functions have been widely used for function approximation, specially since the introduction of the canonical expression by Chua (see [2] and [3]). The use of a canonical expression is of fundamental importance from the numerical efficiency standpoint (it possesses the minimum number of parameters required). Several approximation techniques have been developed using the CPWL expression introduced by Chua. In [4] an algorithm was proposed to alternately adjust the parameters of the function. In [5] a gradient algorithm was proposed in the context of digital filters. Finally, in [6] a Newton-Gauss algorithm was proposed, where, in contrast to [4], all the parameters were adjusted together.

One limitation of the canonical representation introduced in [3] is that it only permits to represent a particular set of all the PWL mappings with domain in $\mathbf{R}^n$. More specifically, those satisfying the *consistent variation* property [7]. In [1] and [8], a HL CPWL representation was proposed for the family of all the continuous PWL mappings defined over a simplicial partition of a domain in $\mathbf{R}^n$. This representation is able to uniformly approximate any Lipschitz continuous function defined on a compact domain. Moreover, in contrast to neural networks or fuzzy approximations, if the Lipschitz constant of the nonlinear function is known, it is possible to calculate the number of terms required to obtain a given error.

In this paper, an approximation methodology for the approximation of continuous functions is proposed using the HL CPWL expression introduced in [1]. We consider the case where a finite set of measured data is available, and use an optimality criterion based on the minimization of the maximum error (over the data set), as it is explained in section 3. After this, in section 4 an application to the characterization of uncertain nonlinear functions is shown. In this case an "upper" and a "lower" PWL function are used to optimally describe the set of all the possible values of the uncertain function.

A salient feature of the methodology is that the approximation problem is reduced to a linear programming problem, for which efficient solution algorithms exist (see [9][10]).

1.1 Definitions

The notation of [1] is adopted for this paper.

Definition 1 (Intersection) *In a partition of the domain $\mathbf{S} \subset \mathbf{R}^n$, an $n-1$ dimensional hyperplane (boundary) is said to be a first order intersection,*

*P. Julián is with CONICET (Comité Nacional de Investigaciones Científicas y Técnicas)

†O. Agamennoni is with CIC (Comité de Investigaciones de la Pcia. de Bs. As.)

denoted by $S^{(1)}$. A linear manifold of dimension $n-k$ in $\mathbf{S}$ is a k-*th order intersection $S^{(k)}$ if it is the intersection of two or more linear manifolds of type $S^{(k-1)}$, i.e.,*

$$S^{(k)} := \bigcap_{i=1}^{\geq 2} S_i^{(k-1)}$$

2 Overview of HL CPWL functions

Here some basic results on HL CPWL are described. For further details, the reader is referred to [1].

Let $\mathbf{S}$ be a rectangular compact set of the form

$$\mathbf{S} := \{(x_1, ..., x_n) : 0 \leq x_i \leq m_i\delta, i \in \{1, ..., n\}\}, \tag{2.1}$$

where δ is the grid size and $m_i \in \mathbf{Z}_+$, and consider a simplicial partition ([11][1]) produced by a boundary configuration H (The set of vertices or $S^{(n)}$ intersections associated to $\mathbf{S}$ will be noted as V_S).

If addition and multiplication by a scalar $r \in \mathbf{R}^1$ are defined as

$$\begin{array}{l} a)\ \ (f+g)(z) = f(z) + g(z), \forall z \in \mathbf{S} \\ b)\ \ (r \cdot f)(z) = r \cdot f(z), \forall z \in \mathbf{S} \end{array} \tag{2.2}$$

then the space $PWL_H[\mathbf{S}]$ of all continuous PWL mappings defined on $\mathbf{S}$ with the boundary configuration H is a *linear vector space.*

In [8] a set of HL CPWL functions which are a basis of $PWL_H[\mathbf{S}]$ was found. The building block of the basis is a generating function

$$\begin{array}{r} \gamma(x_1, x_2) = ||-x_1| + x_2| - |-x_1 + |x_2|| + \\ + |-x_1| + |x_2| - |-x_1 + x_2| \end{array} \tag{2.3}$$

Based on (2.3), a set of functions $\gamma^0(x) = x$, $\gamma^1(x) = \gamma(x, x)$, $\gamma^2(x_1, x_2) = \gamma(x_1, x_2)$ and in general

$$\gamma^k(x_1, ..., x_k) = \gamma\left(x_1, \gamma^{k-1}(x_2, ..., x_k)\right), \tag{2.4}$$

is introduced. A distinctive property of (2.4) is that it possesses k nestings of absolute value functions, and accordingly it is said to have *nesting level* (*n.l.*) equal to k.

The elements of the basis follow from the composition of the functions $\gamma^k(\cdot, ..., \cdot)$, $k = 0, 1, ..., n$ with the set of functions

$$\Im = \{1, x_k - j_k\delta\},$$

where $k \in \{1, ..., n\}$ and $j_k \in \{0, ..., m_k - 1\}$. In addition, they can be expressed in vector form as

$$\Lambda = \left[\Lambda^{0^T}, ..., \Lambda^{n^T}\right]^T, \tag{2.5}$$

ordered according to its *n.l.*, where Λ^i is the vector containing the *n.l.* $= i$ functions. Accordingly, any $f \in PWL_H[\mathbf{S}]$ can be written as

$$f(x) = c^T\Lambda(x),$$

where $c = \left[c_0^T, c_1^T, ..., c_n^T\right]^T$, and every vector c_i is a parameter vector associated to the *n.l.* $= i$ vector function Λ^i.

3 Optimal Approximation

When an experimental system is under study, the information which describes its behavior is generally a finite set of input and output data measurements. If we want to find a function that fits the collected data, first it is necessary to specify the structure of the function and then a criterion must be chosen to select the best approximation. With this idea in mind, the approximation problem is stated as follows.

We assume that a set of points

$$X = \{x_1, ..., x_m\} \tag{3.1}$$

is available, where $x_i \in \mathbf{S}$, $\forall i = 1, ..., m$ and $\mathbf{S}$ is a compact set of the form (2.1). In addition, $\mathbf{S}$ is partitioned with a simplicial partition with grid size δ. Associated to X, there is also a set of function values

$$F = \{f_1, ..., f_m\} \tag{3.2}$$

which can be thought as the values of a function $f : \mathbf{S} \mapsto \mathbf{R}^1$ whose explicit expression is unknown, but it can be measured, and accordingly $f_i = f(x_i)$, $i = 1, ..., m$.

A HL CPWL function $f_p \in PWL_H[\mathbf{S}]$ is proposed to interpolate the values (3.2). The criterion to be used is the minimization of the maximum error

$$\max_{x_i \in X} |f_p(x_i) - f_i| \tag{3.3}$$

over all the points of the set X. This leads us to state the following optimization (min-max) problem:

$$\min_{f_p \in PWL_H[\mathbf{S}]} e_M(f_p, X) \tag{3.4}$$

subject to

$$e_M(f_p, X) = \max_{x_i \in X} |f_i - f_p(x_i)|$$

Note that any $f_p \in PWL_H[\mathbf{S}]$ can be written as $f_p(x) = c^T\Lambda(x)$, $c \in \mathbf{R}^q$. Then, (3.4) can be formulated in the equivalent form

$$\min_{c \in \mathbf{R}^q} \left\{ \max_{x_i \in X} \left| f_i - c^T\Lambda(x_i) \right| \right\} \tag{3.5}$$

Next lemma shows the main result of this section.

Lemma 1 *Let X, F, H and $\mathbf{S}$ as described above. The problem (3.5) can be stated as the linear programming problem*

$$\min \lambda$$

subject to

$$\begin{cases} -c^T\Lambda(x_i) - \lambda \leq -f_i, \forall x_i \in X \\ \ \ c^T\Lambda(x_i) - \lambda \leq \ \ f_i, \forall x_i \in X \\ \lambda \geq 0 \end{cases}$$

on the parameters c and λ.

If we define

$$\lambda = \max_{x_i \in X} \left| f_i - c^T\Lambda(x_i) \right|$$

it is direct to see that

$$\left| f_i - c^T\Lambda(x_i) \right| \leq \lambda, \forall x_i \in X.$$

This equation can be written as

$$\begin{array}{l} +\left(f_i - c^T\Lambda(x_i) \right) \leq \lambda, \forall x_i \in X \\ -\left(f_i - c^T\Lambda(x_i) \right) \leq \lambda, \forall x_i \in X \end{array} \tag{3.6}$$

Then, the problem (3.5) can be stated as the minimization of λ subject to (3.6) and $\lambda \geq 0$, which is the condition stated by Lemma 1.

Remark 1: The PWL function f is uniquely determined by the function values on the vertices of the domain $\mathbf{S}$. In other words, there is a one to one relationship between the elements of the coefficient vector c and the q $(= \prod_{i=1}^{n}(1+m_i))$ values of f over the vertices of $\mathbf{S}$, i.e. the values $\{f(x_i), i = 1, 2, ..., q : x_i \in V_S\}$. With this idea in mind, the stated approximation problem has the following alternative interpretation. Given a set of points X arbitrarily distributed over a domain $\mathbf{S}$ and its associated set of measurements F, find the "best"

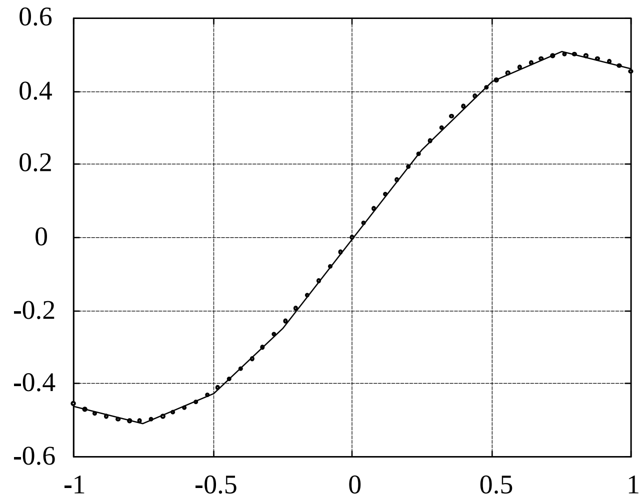

Figure 1: Data and approximation

or the more representative set of function values corresponding to a set of equidistant points (which are the vertices of the partition). This is a problem of data reduction which usually appears in identification problems when large amounts of measurements are stored from an experiment.

Example 1: Consider the function $f(x) = \sin(x)\cos(x)$, defined over the domain

$$\mathbf{S} = \left\{ x \in \mathbf{R}^1 : -1 \leq x \leq 1 \right\},$$

and a set of points

$$X = \{x_i \in \mathbf{S} : x_i = k \cdot 0.04\}$$

with $k \in Z$.

Using the criterion (3.5) a PWL function $g : \mathbf{S} \mapsto \mathbf{R}^1$ of the form $c^T\Lambda(x)$ with a grid step $\delta = 0.25$ was obtained.

The resulting parameter vector is:

$$\begin{array}{c} c = [-0.4617, -0.1863, 0.5148, 0.3854, 0.2591, \\ 0.0101, -0.2439, -0.4106, -0.5148]^T; \end{array}$$

and the value of the maximum error, $\lambda = 7 \times 10^{-3}$. Fig. 1 shows the function values belonging to the set F (dots) and the PWL function g. Fig. 2 shows the error between both functions and the upper and lower bounds given by λ and $-\lambda$.

apro1 erro1

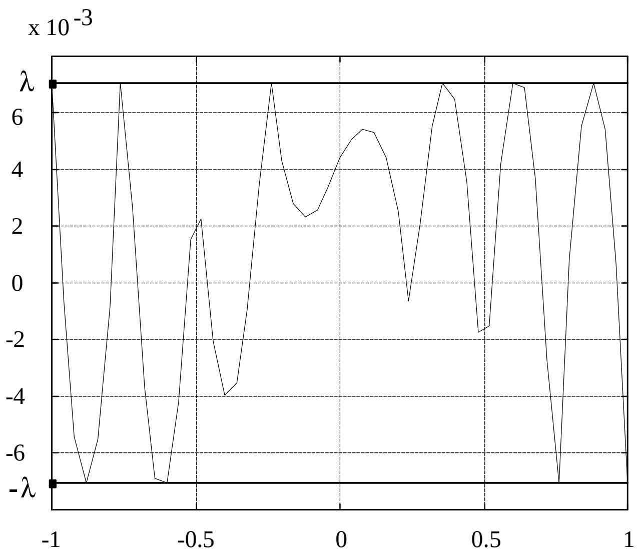

Figure 2: Approximation errors

4 Application to Robust Approximation

In the previous section a method was obtained to interpolate a function using a single PWL function $f \in PWL_H[\mathbf{S}]$. In this section, the function measured is assumed to be uncertain. To clarify this point, consider a set of data points X as (3.1) and an uncertain function defined as a member of the family of functions

$$\Im = \left\{ f : \mathbf{S} \mapsto \mathbf{R}^1 : f(x) = f_N(x) + \Delta(x) \right\}, \quad (4.1)$$

where f_N is a nominal function and Δ satisfies $\sup_{x \in \mathbf{S}} \|\Delta(x)\| \leq K$. In addition, let us consider that a set $F = \{f_1, ..., f_m\}$ of measured values of members of $\Im$ over the points of X is available, i.e., $f_i = f(x_i)$, $f \in \Im$, $x_i \in \mathbf{S}$.
We search for an "upper" function $f_2 \in PWL_H[\mathbf{S}]$, and a "lower" function $f_1 \in PWL_H[\mathbf{S}]$, satisfying

$$f_1(x_i) \leq f(x_i) \leq f_2(x_i), \forall x_i \in \mathbf{S},$$

to characterize the uncertain function, in the sense that

$$f(x_i) = \alpha f_1(x_i) + (1 - \alpha) f_2(x_i)$$

$\forall x_i \in X$, $f \in \Im$, where $0 \leq \alpha \leq 1$. In addition, it is also desirable that the "band" defined by these two functions is as narrow as possible. This is equivalent to find the two functions f_1 and f_2 that solve the following optimization problems:

Problem 1:

$$\min_{f_1 \in PWL_H[\mathbf{S}]} \left\{ \max_{x_i \in X} |f_i - f_1(x_i)| \right\} \quad (4.2)$$

subject to

$$f_i - f_1(x_i) \geq 0$$

Problem 2:

$$\min_{f_2 \in PWL_H[\mathbf{S}]} \left\{ \max_{x_i \in X} |f_i - f_2(x_i)| \right\} \quad (4.3)$$

subject to

$$f_2(x_i) - f_i \geq 0$$

The solutions to these problems can be presented as linear programming problems as it is stated in the following lemma. Before that, consider that the upper and lower functions have the expressions $f_1(x) = c_1^T \Lambda(x)$ and $f_2(x) = c_2^T \Lambda(x)$.

Lemma 2 *Let X, F, H and $\mathbf{S}$ as described above. The problems (4.2) and (4.3) can be stated as the linear programming problems*

$$i) \min \lambda_1$$

subject to

$$\begin{cases} -c_1^T \Lambda(x_i) - \lambda_1 \leq -f_i, \forall x_i \in X \\ -c_1^T \Lambda(x_i) \geq -f_i, \forall x_i \in X \\ \lambda_1 \geq 0 \end{cases}$$

and

$$ii) \min \lambda_2$$

subject to

$$\begin{cases} c_2^T \Lambda(x_i) - \lambda_2 \leq f_i, \forall x_i \in X \\ c_2^T \Lambda(x_i) \geq f_i, \forall x_i \in X \\ \lambda_2 \geq 0 \end{cases}$$

on the parameters c_1, c_2, λ_1 and λ_2.

It can be directly inferred from the proof of Lemma 1
Example 2: In this case a function of the form (4.1) is considered. Intentionally the same function of the preceding example is chosen for f_N ($f_N = \cos(x)\sin(x)$). The uncertainty term is $\Delta(x) = \beta\cos(8x)$, with $0 \leq \beta \leq 0.2$. The domain is

$$\mathbf{S} = \left\{ x \in \mathbf{R}^1 : -1 \leq x \leq 1 \right\},$$

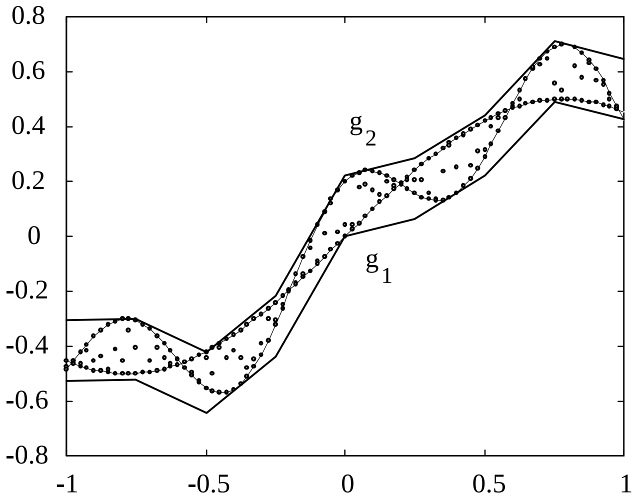

Figure 3: Data, lower and upper HL CPWL functions

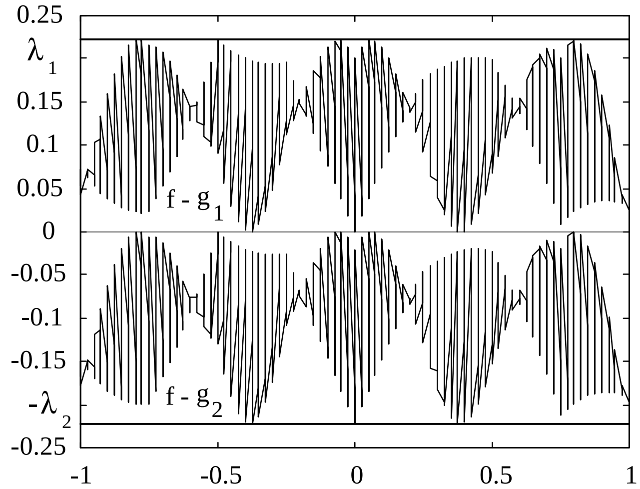

Figure 4: Approximation errors

and the set of points is

$$X = \{x_i \in \mathbf{S} : x_i = k \cdot 0.025\}$$

with $k \in Z$. The function set F was obtained using randomly selected functions of $\Im$ evaluated over the points of X.

Using the criteria (4.2) and (4.3) two PWL functions $g_1, g_2 : \mathbf{S} \mapsto \mathbf{R}^1$ of the form $c_1^T \Lambda(x)$ and $c_2^T \Lambda(x)$, with a grid step $\delta = 0.25$ were obtained.

The resulting parameter vectors are :

$$c_1 = [-0.5271, 0.0227, -0.5055, 1.3094, 0.9155, \\ -1.4874, 0.3786, 0.4384, -1.3291]^T,$$

$$c_2 = [-0.3057, 0.0227, -0.5055, 1.3094, 0.9155, \\ -1.4874, 0.3786, 0.4384, -1.3291]^T.$$

The upper and lower bounds are $\lambda_1 = \lambda_2 = 0.2214$. Fig. 3 shows the function values belonging to the set F (dots), the PWL functions g_1 and g_2 (solid lines) and also the extreme functions $f + 0 \cdot \Delta$ and $f + 0.2 \cdot \Delta$ (fine lines). Fig. 4 shows the errors $f(x_i) - g_1(x_i)$, and $f(x_i) - g_2(x_i)$, $\forall x_i \in X$, together with the bounds λ_1 and λ_2.apro3 erro3

5 Final Comments

A methodology has been proposed to approximate continuous functions using HL PWL functions, when a finite set of input and output measurements is available. The approach can also be used to compress information into a grid of equispaced points when large amounts of data are collected from experimental setups, as discussed in Remark 1.

An optimality criterion is used which leads to a linear programming problem. This is an important property of the method, provided that this type of optimization can be solved by efficient numerical algorithms.

An application to the characterization of uncertain functions has been proposed. The scheme produces a pair of optimal lower and an upper HL CPWL functions which define a "band" containing all the measurements values. As exposed in [12], this can be useful in topics like robust system identification .

References

[1] P. Julián, A. Desages, and O. Agamennoni, "On the high level representation of canonical piecewise linear functions," in *Proc. of ISCAS 98*, (Monterey, California), 1998.

[2] L. O. Chua and S. Kang, "Section-wise piecewise-linear functions: canonical representation, properties and applications," *Proc. IEEE*, vol. 65, pp. 915–929, 1977.

[3] S. Kang and L. O. Chua, "A global representation of multidimensional piecewise-linear functions with linear partitions," *IEEE Trans. Circuits Syst.*, vol. 25, pp. 938–940, 1978.

[4] L. O. Chua and A. Deng, "Canonical piecewise-linear modeling," *IEEE Trans. Circuits Syst.*, vol. CAS-33, pp. 511–525, 1986.

[5] J. N. Lin and R. Unbehahuen, "Adaptive nonlinear digital filter with canonical piecewise-linear function structure," *IEEE Trans. Circuits Syst.*, vol. 37, pp. 347–353, 1990.

[6] P. Julián, M. Jordán, and A. Desages, "Canonical piecewise linear approximation of smooth functions," *IEEE Trans. on Circ. and Syst.*, vol. 45, pp. 567–571, 1998.

[7] L. O. Chua and A. Deng, "Canonical piecewise-linear representation," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 511–525, 1988.

[8] P. Julián, A. Desages, and O. Agamennoni, "High level canonical piecewise linear representation using a simplicial partition." accepted in IEEE Trans. on Circ. and Syst., to appear, 1998.

[9] Nemhauser, Kan, and Todd, *Optimization*. North-Holland, 1996.

[10] R. Vanderbei, *Linear Programming: Foundations and Extensions*. Kluwer Academic Publishers, 1996.

[11] M. Chien and E. Kuh, "Solving nonlinear resistive networks using piecewise-linear analysis and simplicial subdivision," *IEEE Trans. Circuits Syst.*, vol. CAS-24, pp. 305–317, 1977.

[12] P. Julián, O. Agamennoni, and E. N. Sanchez, "On-line nonlinear identification using pwl mappings with applications to wiener modelling." submitted to ECC 99, 1999.