On the Statistical Optimality of Locally Monotonic Regression

Alfredo Restrepo (Palacios) and Alan C. Bovik

Abstract—Locally monotonic regression is a recently proposed technique for the deterministic smoothing of finite-length discrete signals under the smoothing criterion of local monotonicity. Locally monotonic regression falls within a general framework for the processing of signals that may be characterized in three ways: regressions are given by projections that are determined by semimetrics, the processed signals meet shape constraints that are defined at the local level, and the projections are optimal statistical estimates in the maximum likelihood sense. Here, we explore the relationship between the geometric and deterministic concept of projection onto (generally nonconvex) sets and the statistical optimal statistical estimates in the maximum likelihood sense. Here, we explore the relationship between the geometric and deterministic concept of projection onto (generally nonconvex) sets and the statistical concept of likelihood, with the object of characterizing projections under the family of the semi-metrics as maximum likelihood estimates of signals contaminated with noise from a well-known family of exponential densities.

I. INTRODUCTION

We discuss here a statistical aspect of the concept of projection onto (generally nonconvex) sets of signals, or regression, as it was recently proposed in [1], [2], for the processing of finite-length discrete signals. One may argue, particularly with optical images, that some signals carry their information explicitly as shape. A shape constraint is a property defined in the natural domain of the signal, e.g., time or space rather than frequency; we consider shape constraints that are defined at the local level rather than at the global level, e.g., local monotonicity versus (global) monotonicity.

The projections of a signal on a set of signals are defined as the signals in the set that are closest to the signal being projected, under a semi-metric for the space.

Local monotonicity [3] is shape constraint for one-dimensional signals that provides a measure of the smoothness of a signal; it sets a limit on the roughness of a signal by limiting how often the signal may have a change of trend (increasing to decreasing or vice versa). In a sense of the word, it limits the oscillations that a signal may have without making restrictions on the magnitude of the changes of the signal from each coordinate to the following one. Piecewise constancy, piecewise linearity, and local convex/concavity are other, similar shape constraints that have been explored [4]. Algorithms for the computation of the locally monotonic regression of a finite-length signal of length $n$ have a complexity that is exponential [1]; however, much faster algorithms have been developed that employ regression only over a moving window (hence, computation is linear with signal length) [5], and that compute a fuzzy approximation to the true locally monotonic regression using a generalized deterministic annealing algorithm [6].

Manuscript received May 15, 1992; revised September 16, 1993. The associate editor coordinating the review of this paper and approving it for publication was Prof. Gonzalo Arce.

A. Restrepo (Palacios) was with the Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712. He is now with the Departamento de Ingeniería Eléctrica, Universidad de los Andes, A. A. 4976, Bogotá, Colombia.

A. C. Bovik is with the Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712 USA.

IEEE Log Number 9400387.

II. SHAPE CONSTRAINTS

An integer interval $[a, b]$, where $a$ and $b$ are integer numbers, is defined as the subset of the integers that are larger than or equal to $a$ and smaller than or equal to $b$. An $n$-point signal (or a discrete signal of length $n$) is a real function $x$ having as domain a nonempty integer interval $[a, b]$, where $b - a = n - 1$. An $n$-point signal is a point of $R^n$ and may be expressed as the $n$-tuple $[x_1, \ldots, x_n]$ of the values it takes. The origin $[0, 0, \ldots, 0]$ of $R^n$ is denoted $\theta$.

The slope skeleton of an $n$-point signal $x = [x_1, \ldots, x_n]$ is the $(n - 1)$-point signal $s = [s_1, \ldots, s_{n-1}]$ having components

$$s_i = \text{sgn}(x_{i+1} - x_i)$$

where $\text{sgn}$ is the signum function that is, respectively, one, zero, and minus one, when its argument is positive, zero, or negative.

The segments of length $r$ of a signal $w: [1, n] \rightarrow R^r$, $r \leq n$ are the restrictions of $w$ to integer intervals of length $r$. For example, $[2, 3, 4]$ is a segment of length $3$ of the signal $[1, 2, 3, 4]$.

Constancy and linearity are well-known shape constraints. Less commonly used shape constraints are monotonicity, convexity, concavity, piecewise constancy, piecewise linearity, local monotonicity, and local convex/concavity [4], [5].

A signal is constant if its slope skeleton $s$ is null: $s = \theta$. A signal is said to be nondecreasing if the components of its slope skeleton are nonnegative and nonincreasing if they are nonpositive. A signal is monotonic if it is either nonincreasing or nondecreasing. A signal is said to be convex if its slope skeleton is nondecreasing and concave if its slope skeleton is nonincreasing. These are global shape constraints.

A signal is locally monotonic of degree $a$ (lomo-$a$) if each of its segments of length $a$ is monotonic. A signal is locally convex/concave of degree $a$ (lconv-$a$) if each of its segments of length $a$ has a monotonic slope skeleton. These are shape constraints defined at the local level.

A signal is linear if its slope skeleton is constant. The algebraic span of the constant signal $[1, 1, \ldots, 1]$ and the linear signal $[1, 2, \ldots, n]$ is the collection of the linear signals of length $n$.

A signal may be segmented into longest constant segments in a unique way. For example, the longest constant segments in $[1, 3, 3, 2, 2, 4, 5, 5]$ are $[1, 3, 3, 2, 2, 4, 5]$, and $[3, 2, 2, 4, 5]$ [4], and $[5, 5]$. A signal is said to be piecewise constant of degree $a$ (pico-$a$) if, besides the first and last segments, the shortest segments of its segmentation into constant segments have length at least $a$.

Each coordinate point $i$ of a signal $s$ such that $s_{i-1} - 2s_i + s_{i+1}$ is nonzero is a point where the slope of $s$ changes and is called a hinge of $s$. In addition, the first and last coordinates of a signal are called hinges; accordingly, a linear signal (of length larger than one) has exactly two hinges. A signal is said to be piecewise linear of degree $a$ (pil-$a$) if the difference between each two consecutive hinges is at least $a$; any signal is pil-$1$. An $n$-point signal may have as few as two hinges and as many as $n$ hinges.

For example, the signal in Fig. 1(a) is lomo-$4$, loco-$7$, pico-$3$, and pil-$1$. The signal in Fig. 1(b) is lomo-$2$, loco-$5$, pico-$1$, and pil-$3$. All of these shape constraints are defined locally and are novel measures of the smoothness of a signal; the larger the degree $a$, the smoother the signal.

III. SEMI-METRICS AND LIKELIHOOD FUNCTIONS

A semi-metric differs from a metric in that it may lack the triangle inequality property. A semi-metric measures, in a coordinatewise way, the similarity of a pair of signals. A semi-metric for $R^n$ is a function $d: R^n \times R^n \rightarrow [0, \infty)$ that is positive definite and symmetric,
A Family of projections have been characterized in [1]. The distance between nondecreasing, and continuous.

Assume that a shape constraint has been specified. Let \( A \) be the set of signals meeting the constraint and \( \mathcal{S} \) be a set of signals having the required shape.

Given a signal \( x \) that is not in \( A \), if \( A \) is a nonempty proper subset of \( \mathbb{R}^n \) and \( x \in (\mathbb{R}^n - A) \), the set

\[
P_A(x) = \{ a \in A : d(x,a) = D_p(x,A) \}
\]

gives the set of the projections of \( x \) on \( A \) (or of regressions of \( x \) with respect to \( A \)) under the semi-metric \( d \).

The existence and multiplicity of projections have been characterized in [1]. The distance between \( x \) and \( A \) is given by the smallest radius \( \rho \) for which the boundary of \( A \setminus B(\theta,\rho) \) contains \( x \) (\( \setminus \) stands for the operator of Minkowski set addition [7]).

Algorithms for the computation of linear regression are well known. Algorithms that compute locally monotonic projections under the \( p \)-semi-metrics \( d_p \), \( p \in (0, \infty) \) are given in [1]. Algorithms for the computation of projections under the Euclidean metric for other shape constraints such as piecewise linearity and local convex/concavity may be found in [4], [6].

B. A Family of Densities

We make use of a family of generalized exponential probability density functions that have been used extensively in robust statistics [8] and in the design of order statistic filters [9]-[11]. It is defined as follows:

\[
f_p(x) = \gamma \exp \left( -\zeta |x|^p \right) \quad p \in (0, \infty), x \in \mathbb{R}^n
\]

where \( \gamma \) and \( \zeta \) are positive constants (that depend on \( p \)) that determine the variance of the random variable and ensure that each density \( f_p \) integrates to one. The Gaussian and Laplacian densities are in the family.

C. Likelihood Functions

Assume that a system outputs (deterministic) signals of length \( n \) that are characterized by a shape constraint \( Q \), and that any signal meeting the shape constraint may be an output of the system, with uniform probability on the set \( A = \{ s \in \mathbb{R}^n : Q(s) \} \); signals not in \( A \) are outputs of the system with probability zero. Also, before the output signal \( t \) can be observed, it becomes contaminated with a random signal \( \varepsilon \). Given an observed noisy signal \( z = t + \varepsilon \), it is desired to obtain an estimate \( \hat{s} \) of \( t \). The likelihood \( L(z; a) \) of \( z \) being equal to a signal \( a \) in \( A \) plus a sample \( z \) from \( \varepsilon \) depends both on \( a \) and \( z \). Given the distribution of the components of \( \varepsilon \), the set
A., and an observation x, it is desired to find signals a in A that maximize L(x; a) as a ranges over A (see Fig. 3). For a white random signal g of i.i.d. components having a density f(x), the likelihood function L(x; a) is given by

\[ L(x; a) = \exp \left[ -s(\{ |x_1 - a_1| + \cdots + |x_n - a_n| \} \right] \]

and a monotonic dependence of the likelihood on the distance between x and a appears. Thus, a maximum likelihood estimate is a signal a in A that is closest to x under the semimetric \( d_p \); that is, the maximum likelihood estimates are the p-th projections of \( x \) on A.

In [12], an M estimate is an estimate T that minimizes a function of the form \( \Sigma p(x; T) \), where p is an arbitrary function, and an ML estimate \( \tau \) as one that minimizes \( \Sigma \ln g(\tau) \). Of course, the equivalence between minimizing a loss function and maximizing a likelihood function is well known; here, we point out a particular relationship between maximum likelihood estimation and semi-metrics within the context of projection.

D. Semi-Metrics that A are Obtained from Likelihood Functions

As above, the likelihood function is in many cases an increasing function of a semi-metric; in fact, a large class of likelihood functions determine translation-invariant semi-metrics. For example, let \( g_1, g_2, \ldots, g_n \) be a collection of continuous probability density functions that are unimodal and increasing on \( -\infty, 0 \) and strictly decreasing on \( 0, \infty \), and positive on the real line \( \mathbb{R}^n \). The exponential densities defined above, and many others, e.g., the Cauchy density, meet these requirements. Let \( g : \mathbb{R}^n \rightarrow \mathbb{R}^n \) be given by \( g(v) = g_1(v_1)g_2(v_2)\cdots g_n(v_n) \), where \( v = [v_1, v_2, \ldots, v_n] \). Then, \( \ln(g) \) is a continuous function that is maximal at \( \theta \), and the function \( \delta : \mathbb{R}^n \rightarrow \mathbb{R}^n \) given by

\[ \delta(u, v) = \ln \frac{g(\theta)}{g(u - v)} \]

is a continuous, nondecreasing semi-metric for \( \mathbb{R}^n \).

If A is the set of signals meeting a given shape constraint \( Q, g = [z_1, z_2, \ldots, z_n] \) is a random signal of independent components where each component \( z_i \) has a density \( g_i \), and \( g = a + z \) is a noisy version of a signal a in A, then, given an observed sample signal x, the likelihood that x results from the addition of a signal a and a sample z of g is \( L(x; a) = g(x - a) \). Under the hypothesis that a is in A, L(x; a) is maximized by a signal s that minimizes \( \ell(x; a) : a \in A \), that is, by any of the projections of x on A under the semi-metric \( \delta \).

IV. CONCLUSION

The technique of regression is widely used in statistics, but it is rarely used with the purpose of shaping or filtering signals. Linear regression [13] and isotonic regression [14], [15] are perhaps the most commonly used types of regression in statistics. Conditions of monotonicity arise naturally in certain classes of problems; for example, consider the collection of attention times of the elements of a queue [16]. Projections on the set of monotonic signals are called monotonic regressions. Isotonic regression and related topics have been considered previously for image processing [17].

Here, we have examined a signal-smoothing paradigm, designed under geometric and deterministic concepts of projection, that is optimal for signal estimation in the maximum likelihood sense. The use of semi-metrics in the projection apparatus, although not common in signal processing, provides estimators of signals contaminated with highly impulsive noise. A family of robust estimators, each optimal for a given density, is obtained. The shape constraints considered here provide criteria of smoothness that emphasize different characteristics of signals. We believe that the use of shape constraints that are defined at the local level, together with the concept of projection provide a powerful tool for the shaping, smoothing, and filtering of signals.

REFERENCES