Nonlinear Signal Processing ELEG 833

Gonzalo R. Arce

Department of Electrical and Computer Engineering

University of Delaware

arce@ee.udel.edu

May 5, 2005

8 Myriad Smoothers

8.1 FLOM Smoothers

Under the framework of Gaussian processes, the sample mean $(\overline{\beta})$ minimizes the second moment of the shifted variable $X - \beta$ over all possible shifts.

$$E(X) = \bar{\beta} = \arg\min_{\beta} E(X - \beta)^2.$$
(1)

(2)

Second-order moments do not exist with stable processes, but fractional-order moments do. The second moment in (1) can be replaced by fractional lower-order moments (FLOMs) to obtain the following measure of location

$$\beta_p = \operatorname*{arg\,min}_{\beta} \mathrm{E}(|X - \beta|^p), \quad p < 2.$$

FLOM smoothers follow from (2) where FLOM estimates are computed in the running window $\mathbf{X}(n) = [X_1(n), X_2(n), \dots, X_N(n)]^T$ as

$$Y(n) = \operatorname*{arg\,min}_{\beta} \sum_{i=1}^{N} |X_i(n) - \beta|^p \tag{3}$$

with p < 2.

The behavior of FLOM smoothers is markedly dependent on the choice of p. As $p \rightarrow 2$, FLOM smoothers resemble the running mean. As p is reduced in value, FLOM smoothers become more robust and its output can tract discontinuities more effectively.



- FLOM smoothers arise from the location estimation problem under the generalized Gaussian distribution.
- For p < 1, FLOM smoothers are selection type.
- For p > 1, the cost function is convex and the output is not necessarily equal in value to one of the input samples.
- FLOM smoother computation is in general nontrivial.
- A method to overcome this limitation is to force the output of the smoother to be identical in value to one of the input samples.
- Selection type FLOM smoothers are suboptimal and are referred to as gamma filters.





Figure 3: FLOM smoothing of an image for different values of p. Outputs of the FLOM smoother for: (*c*) p = 0.1 (PSNR=31.86dB), (*d*) p = 1 (median smoother, PSNR=37.49dB).



Figure 4: FLOM smoothing of an image for different values of p (continued). (a) p = 2 (mean smoother, PSNR=33.53dB), (b) p = 10 (PSNR=31.15).





Figure 6: Gamma smoothing of an image for different values of p and different window sizes. Output of the 3×3 gamma smoother for (*c*) p = 10 (PSNR=32.32dB), and the 5×5 gamma smoother for (*d*) p = 0.1 (PSNR=28.84dB).



8.2 Running Myriad Smoothers

Given an observation vector $\mathbf{X}(n) = [X_1(n), X_2(n), \dots, X_N(n)]$ and a fixed positive (tunable) value of K, the running myriad smoother output at time n is computed as

$$Y_{K}(n) = \operatorname{MYRIAD}[K; X_{1}(n), X_{2}(n), \dots, X_{N}(n)]$$

=
$$\operatorname{arg\,min}_{\beta} \prod_{i=1}^{N} \left[K^{2} + (X_{i}(n) - \beta)^{2} \right].$$
(4)

$$= \arg \min_{\beta} \sum_{i=1}^{\infty} \log \left[K^2 + (X_i(n) - \beta)^2 \right].$$
 (5)

The myriad $Y_K(n)$ is thus the value of β that minimizes the above cost function.

The definition of the sample myriad involves the free-tunable parameter K. This parameter will be shown to play a critical role in characterizing the behavior of the myriad.



Figure 8: Myriad cost functions for different values of k

Property 8.1 (Linear Property) Given a set of samples, X_1, X_2, \ldots, X_N , the sample myriad $\hat{\beta}_K$ converges to the sample average as $K \to \infty$. This is,

$$\lim_{K \to \infty} \hat{\beta}_{K} = \lim_{K \to \infty} \operatorname{MYRIAD}(K; X_{1}, \dots, X_{N})$$
$$= \frac{1}{N} \sum_{i=1}^{N} X_{i}.$$
(6)

Note that
$$\hat{\beta}_K \leq X_{(N)}$$
 by checking that for any i , and for
 $\beta > X_{(N)}, K^2 + (X_i - \beta)^2 > K^2 + (X_i - X_{(N)})^2$. In the same way, $\hat{\beta}_K \geq X_{(1)}$. Hence,

$$\hat{\beta}_{K} = \arg \min_{X_{(1)} \le \beta \le X_{(N)}} \prod_{i=1}^{N} [K^{2} + (X_{i} - \beta)^{2}]$$

$$= \arg \min_{X_{(1)} \le \beta \le X_{(N)}} \left\{ K^{2N} + K^{2N-2} \sum_{i=1}^{N} (X_{i} - \beta)^{2} + f(K) \right\}$$
(8)

where $f(K) = O(K^{2N-4})$ and O denotes the *asymptotic order* as $K \to \infty$.

Since adding or multiplying by constants does not affect the $\arg \min$ operator, Equation (8) can be rewritten as

$$\hat{\beta}_{K} = \underset{X_{(1)} \le \beta \le X_{(N)}}{\arg\min} \left\{ \sum_{i=1}^{N} (X_{i} - \beta)^{2} + \frac{O(K^{2N-4})}{K^{2N-2}} \right\}.$$
 (9)

Letting $K \to \infty,$ the term $O(K^{2N-4})/K^{2N-2}$ becomes negligible, and

$$\hat{\beta}_K \to \underset{X_{(1)} \le \beta \le X_{(N)}}{\operatorname{arg min}} \left\{ \sum_{i=1}^N (X_i - \beta)^2 \right\} = \frac{1}{N} \sum_{i=1}^N X_i.$$

Definition 8.1 (Sample mode-myriad) Given a set of samples $X_1, X_2, \ldots, \hat{}$

 X_N , the mode-myriad estimator, \hateta_0 , is defined as

$$\hat{\beta}_0 = \lim_{K \to 0} \hat{\beta}_K,\tag{10}$$

where $\hat{\beta}_K = MYRIAD(K; X_1, X_2, \dots, X_N).$

Property 8.2 (Mode Property) The mode-myriad $\hat{\beta}_0$ is always equal to one of the most repeated values in the sample. Furthermore,

$$\hat{\beta}_0 = \underset{X_j \in \mathcal{M}}{\operatorname{arg\,min}} \prod_{i=1, X_i \neq X_j}^N |X_i - X_j|, \tag{11}$$

where \mathcal{M} is the set of most repeated values.









Figure 12: (*a*) The sample myriad, $\hat{\beta}$, minimizes the product of distances from point A to all samples. Any other value, such as $x = \beta'$, produces a higher product of distances; (*b*) the myriad as *K* is reduced.





generating density becomes highly localized, and the data are visual-

ized as very impulsive (the optimal estimator is a cluster locator).





Property 8.3 (Scale Invariance) Let $\hat{\beta}_K(\mathbf{X})$ denote the myriad of order K of the data in the vector \mathbf{X} . Then, for c > 0,

$$\hat{\beta}_K(c\mathbf{X}) = c\hat{\beta}_{K/c}(\mathbf{X}). \tag{12}$$

Proof : Let X_1, X_2, \ldots, X_N denote the data in \mathbf{X} . Then,

$$\hat{\beta}_{K}(c\mathbf{X}) = \arg\min_{\beta} \prod_{i=1}^{N} \left[K^{2} + (cX_{i} - \beta)^{2} \right]$$

$$= \arg\min_{\beta} \prod_{i=1}^{N} \left[\left(\frac{K}{c} \right)^{2} + \left(X_{i} - \frac{\beta}{c} \right)^{2} \right] \qquad (13)$$

$$= c \left(\arg\min_{\beta} \prod_{i=1}^{N} \left[\left(\frac{K}{c} \right)^{2} + (X_{i} - \beta)^{2} \right] \right).$$

8.3 Optimality of the Sample Myriad

Optimality In The α -Stable Model

Proposition 8.1 Let $T_{\alpha,\gamma}(X_1, X_2, \ldots, X_N)$ denote the maximum likelihood location estimator derived from a symmetric α -stable distribution with characteristic exponent α and dispersion γ . Then,

$$\lim_{\alpha \to 0} T_{\alpha,\gamma}(X_1, X_2, \dots, X_N) = \mathsf{MYRIAD} \{0; X_1, X_2, \dots, X_N\}.$$
(14)

The α -stable triplet of optimality points satisfied by the myriad:

•
$$\alpha = 2 \leftrightarrow K = \infty$$

•
$$\alpha = 1 \leftrightarrow K = \gamma$$

•
$$\alpha = 0 \leftrightarrow K = 0$$

Proposition 8.2 Let α and γ denote the characteristic exponent and dispersion parameter of a symmetric α -stable distribution. Let $K_o(\alpha, \gamma)$ denote the optimal tuning value of K in the sense that $\hat{\beta}_{K_o}$ minimizes a given performance criterion (usually the variance) among the class of sample myriads with non negative linearity parameter. Then,

$$K_o(\alpha, \gamma) = K_o(\alpha, 1)\gamma.$$
(15)

A simple empirical formula is

$$K(\alpha) = \sqrt{\frac{\alpha}{2 - \alpha}},\tag{16}$$



8.4 Weighted Myriad Smoothers

Given N observations $\{X_i\}_{i=1}^N$ and nonnegative weights $\{W_i \ge 0\}_{i=1}^N$, let the input and weight vectors be defined as $\mathbf{X} \triangleq [X_1, X_2, \dots, X_N]^T$ and $\mathbf{W} \triangleq [W_1, W_2, \dots, W_N]^T$, respectively. For a given *nominal* scale factor K, the underlying random variables are assumed to be independent and Cauchy distributed with a common location parameter β , but varying scale factors $\{S_i\}_{i=1}^N$: $X_i \sim \text{Cauchy}(\beta, S_i)$:

$$f_{X_i}(X_i;\beta,S_i) = \frac{1}{\pi} \frac{S_i}{S_i^2 + (X_i - \beta)^2}, \quad -\infty < X_i < \infty, \quad (17)$$

and where

$$S_i \stackrel{\triangle}{=} \frac{K}{\sqrt{W_i}} > 0, \ i = 1, 2, \dots, N.$$
 (18)

The weighted myriad smoother output $\hat{eta}_K(\mathbf{W},\mathbf{X})$ is defined as follows

$$\hat{\beta}_{K}(\mathbf{W}, \mathbf{X}) = \arg \max_{\beta} \prod_{i=1}^{N} \frac{S_{i}}{S_{i}^{2} + (X_{i} - \beta)^{2}},$$

$$= \arg \min_{\beta} \prod_{i=1}^{N} \left[K^{2} + W_{i} (X_{i} - \beta)^{2} \right] \quad (19)$$

$$\stackrel{\Delta}{=} \arg \min_{\beta} P(\beta); \qquad (20)$$

Alternatively, we can write $\hat{eta}_K(\mathbf{W},\mathbf{X}) \stackrel{ riangle}{=} \hat{eta}_K$ as

$$\hat{\beta}_{K} = \underset{\beta}{\operatorname{arg\,min}} Q(\beta) \stackrel{\Delta}{=} \underset{\beta}{\operatorname{arg\,min}} \sum_{i=1}^{N} \log \left[K^{2} + W_{i} \left(X_{i} - \beta \right)^{2} \right];$$

$$\hat{\beta}_{K} \text{ is the global minimizer of } P(\beta) \text{ and } Q(\beta) \stackrel{\Delta}{=} \log(P(\beta)).$$
(21)

Definition 8.2 (Weighted myriad) Let $\mathbf{W} = [W_1, W_2, \dots, W_N]$ be a vector of nonnegative weights. Given K > 0, the weighted myriad of order K for the data X_1, X_2, \dots, X_N is defined as

$$\hat{\beta}_{K} = \operatorname{MYRIAD} \left\{ K; W_{1} \circ X_{1}, \dots, W_{N} \circ X_{N} \right\}$$
$$= \operatorname{arg\,min}_{\beta} \sum_{i=1}^{N} \log \left[K^{2} + W_{i} (X_{i} - \beta)^{2} \right], \quad (22)$$

where $W_i \circ X_i$ represents the weighting operation in (22). In some situations, the following equivalent expression can be computationally more convenient

$$\hat{\beta}_{K} = \arg\min_{\beta} \prod_{i=1}^{N} \left[K^{2} + W_{i} (X_{i} - \beta)^{2} \right].$$
 (23)

Note that the weighted myriad has only ${\cal N}$ independent parameters

$$\hat{\beta}_K(\mathbf{W}, \mathbf{X}) = \hat{\beta}_1\left(\frac{\mathbf{W}}{K^2}, \mathbf{X}\right)$$
 (24)

Equivalently:

$$\hat{\beta}_{K_1}(\mathbf{W}_1, \mathbf{X}) = \hat{\beta}_{K_2}(\mathbf{W}_2, \mathbf{X}) \quad \text{iff} \quad \frac{\mathbf{W}_1}{K_1^2} = \frac{\mathbf{W}_2}{K_2^2}.$$
 (25)

Hence, the output depends only on $\frac{\mathbf{W}}{K^2}$.

The objective function $P(\beta)$ is a polynomial in β of degree 2N, with well-defined derivatives of all orders. Therefore, it can have at most (2N-1) local extremes, one of which is the output:



Figure 18: Sketch of a typical weighted myriad objective function $Q(\beta)$ for the weights [1, 2, 3, 2, 1] (solid line), and [1, 100, 3, 2, 1] (dashed line), and the sample set [-1, 10, 3, 5, -3].

Property 8.4 (Linear Property) In the limit as $K \to \infty$, the weighted myriad reduces to the normalized linear estimate

$$\lim_{K \to \infty} \hat{\beta}_{K} = \frac{\sum_{i=1}^{N} W_{i} X_{i}}{\sum_{i=1}^{N} W_{i}}.$$
(26)

Property 8.5 (No undershoot/overshoot) The output of a weighted myriad smoother is always bracketed by

$$X_{(1)} \le \hat{\beta}_K(\mathbf{W}; X_1, X_2, \dots, X_N) \le X_{(N)},$$
 (27)

where $X_{(1)}$ and $X_{(N)}$ denote the minimum and maximum samples in the input window.

Property 8.6 (Mode Property) Given a vector of positive weights, $\mathbf{W} = [W_1, \dots, W_N]$, the weighted mode-myriad $\hat{\beta}_0$ is always equal to one of the most repeated values in the sample. Furthermore,

$$\hat{\beta}_0 = \operatorname*{arg\,min}_{X_j \in \mathcal{M}} \left(\frac{1}{W_j}\right)^{\frac{r}{2}} \prod_{i=1, X_i \neq X_j}^N |X_i - X_j|, \quad (28)$$

where \mathcal{M} is the set of most repeated values, and r is the number of times a member of \mathcal{M} is repeated in the sample set.

Property 8.7 (Outlier Rejection Property) Let $K < \infty$, and let W denote a vector of positive and finite weights. The outlier rejection property states that:

$$\lim_{X_N \to \pm \infty} \hat{\beta}_K(\mathbf{W}; X_1, X_2, \dots, X_N) = \hat{\beta}_K(\mathbf{W}; X_1, X_2, \dots, X_{N-1}).$$
(29)

Property 8.8 (Unbiasedness) Let X_1, X_2, \ldots, X_N be all independent and symmetrically distributed around the point of symmetry c. Then, $\hat{\beta}_K$ is also symmetrically distributed around c. In particular, if $E\hat{\beta}_K$ exists, then $E\hat{\beta}_K = c$.



Figure 19: (a) The sample myriad, $\hat{\beta}_K$, indicates the position of a moving bar such that the product of distances from point A to the sample points X_1, X_2, \ldots, X_N is minimum. (b) If the weight $W_4 > 1$ is introduced, the product of distances is more sensitive to the variations of the segment $\overline{X_4A_4}$, very likely resulting in a weighted myriad $\hat{\beta}_K$ closer to X_4 .

8.5 Fast Weighted Myriad Computation

- No explicit formulation is available to compute the weighted myriad.
- But, some characteristics of the objective function can be exploited.
- Turns out certain numerical iterative technique can be applied.

Recall that the weighted myriad is given by

$$\hat{\beta}_{K} = \underset{\beta}{\operatorname{arg\,min}} \log(P(\beta)) \stackrel{\Delta}{=} \underset{\beta}{\operatorname{arg\,min}} Q(\beta)$$
$$= \underset{\beta}{\operatorname{arg\,min}} \sum_{i=1}^{N} \log\left[1 + \left(\frac{x_{i} - \beta}{S_{i}}\right)^{2}\right], \quad (30)$$

it's easi to show that

$$Q'(\beta) = 2 \sum_{i=1}^{N} \frac{W_i (\beta - X_i)}{K^2 + W_i (X_i - \beta)^2}.$$
 (31)

$$Q'(\beta) = 2 \sum_{i=1}^{N} \frac{\left(\frac{\beta - X_i}{S_i^2}\right)}{1 + \left(\frac{X_i - \beta}{S_i}\right)^2}.$$
 (32)
Defining
$$\psi(v) \stackrel{\triangle}{=} \frac{2v}{1 + v^2},$$
 (33)
and referring to (32) the following equation is obtained for the local
extremes of $Q(\beta)$:
$$Q'(\beta) = -\sum_{i=1}^{N} \frac{1}{S_i} \cdot \psi\left(\frac{X_i - \beta}{S_i}\right) = 0.$$
 (34)

By introducing the *positive* functions

$$h_i(\beta) \stackrel{\triangle}{=} \frac{1}{S_i^2} \cdot \varphi\left(\frac{X_i - \beta}{S_i}\right) > 0, \tag{35}$$

for $i=1,2,\ldots,N,$ where

$$\varphi(v) \stackrel{\triangle}{=} \frac{\psi(v)}{v} = \frac{2}{1+v^2},$$
(36)

the local extremes of $Q(\beta)$ in (34) can be formulated as

$$Q'(\beta) = -\sum_{i=1}^{N} h_i(\beta) \cdot (X_i - \beta) = 0.$$
 (37)

Fixed Point Formulation

Equation (37) can be written as

 $\beta = \frac{\sum_{i=1}^{N} h_i(\beta) \cdot X_i}{\sum_{i=1}^{N} h_i(\beta)}$ (38)

By defining the mapping

$$T(\beta) \stackrel{\triangle}{=} \frac{\sum_{i=1}^{N} h_i(\beta) \cdot X_i}{\sum_{i=1}^{N} h_i(\beta)},$$
(39)

the local extremes of $Q(\beta)$, or the roots of $Q'(\beta)$, are seen to be the *fixed* points of $T(\cdot)$:

$$\beta^* = T(\beta^*). \tag{40}$$

The following *fixed point iteration* results in an efficient algorithm to compute these fixed points:

$$\beta_{m+1} \stackrel{\Delta}{=} T(\beta_m) = \frac{\sum_{i=1}^{N} h_i(\beta_m) \cdot X_i}{\sum_{i=1}^{N} h_i(\beta_m)}.$$
 (41)

In the classical literature, this is also called the *method of successive* approximation for the solution of the equation $\beta = T(\beta)$.

It has been proven that the iterative method of (41) converges to a fixed point of $T(\cdot)$; thus,

$$\lim_{m \to \infty} \beta_m = \beta^* = T(\beta^*).$$
(42)

Fixed Point Weighted Myriad Search

Step 1: Select the initial point $\hat{\beta}_0$ among the values of the input samples: $\hat{\beta}_0 = \operatorname*{arg\,min}_{X_i} P(X_i).$

Step 2: Using $\hat{\beta}_0$ as the initial value, perform L iterations of the fixed point recursion $\beta_{m+1} = T(\beta_m)$ of (41). The final value of these iterations is then chosen as the weighted myriad: $\hat{\beta}_{\text{FP}} = T^{(L)}(\hat{\beta}_0)$.

This algorithm can be compactly written as

$$\hat{\beta}_{\mathsf{FP}} = T^{(L)} \left(\arg\min_{X_i} P(X_i) \right). \tag{43}$$

8.6 Weighted Myriad Smoother Design

8.6.1 Center-Weighted Myriads for Image Denoising

The notion of center weighting can be applied to the myriad structure, leading to:

$$Y = \text{MYRIAD} \{K; X_1, \dots, W_c \circ X_c, \dots, X_N\}.$$
(44)

The cost function in (21) is now modified to

$$Q(\beta) = \log \left[K^2 + W_c (X_c - \beta)^2 \right] + \sum_{X_i \neq X_c} \log \left[K^2 + (X_i - \beta)^2 \right].$$
(45)

- In addition to the center weight W_c , the CWMy has the free parameter (K) that controls the impulsiveness rejection.
- The center weight in the CWMy smoother is data dependent.
- For different applications, the center weight should be adjusted based on their data ranges.
- For grayscale image denoising (values normalized between 0 and 1), the parameters are:
 - (1) Choose $K = (X_{(U)} + X_{(L)})/2$, where $1 \le L < U \le N$, with $X_{(U)}$ being the Uth smallest sample in the window and $X_{(L)}$ the Lth smallest sample.

(2) Set $W_c = 10,000$.

- When there is "salt" noise in the window (outliers having large values), the myriad structure assures that they are deemphasized because of the outlier rejection property of K.
- For a single "pepper" outlier sample, the cost function (45) evaluated at $\beta = K$ will always be smaller than that at $\beta = 0$. Thus, "pepper" noise will never go through the smoother.

A 2-pass CWMy smoother can be defined as follows:

$$\mathbf{Y} = 1 - \mathrm{CWMy}(1 - \mathrm{CWMy}(\mathbf{X})). \tag{46}$$





with $W_c = 10,000$ and $K = (X_{(21)} + X_{(5)})/2$ (PSNR=39.98dB)







Figure 24: Output of the Center weighted myriad smoother for different values of the center weight W_c (a) Original image, (b) 100 (PSNR=36.74dB)



ues of the center weight W_c (*c*) 10,000 (PSNR=39.98dB), (*d*) 1,000,000 (PSNR=38.15dB).





8.6.2 Myriadization

- First, design a constrained linear smoother for Gaussian or noiseless environments using FIR filter (smoother) design techniques.
- Then, plug in these smoother coefficients into weighted myriad structure.
- Choose the suitable K according to the impulsiveness of the environment.
- Note that the smoother coefficients W_i must be non-negative and satisfy the normalization condition $\sum_{i=1}^{N} W_i = 1$





