Nonlinear Signal Processing
ELEG 833

Gonzalo R. Arce

Department of Electrical and Computer Engineering
University of Delaware
arce@ee.udel.edu

Fall 2008

Outline

1. Introduction
   - Non-Gaussian Random Processes
     - Generalized Gaussian Distributions and Weighted Medians
     - Stable Distributions and Weighted Myriads
   - Statistical Foundations
   - The Filtering Problem
     - Moment Theory

Signal processing embodies a large set of methods for:

- Representation: Medical imaging
- Analysis: Wavelet, spectral analysis
- Transmission: Wireless, Ethernet, ADSL
- Restoration of information-bearing signals: Digital imaging, digital video.

Linear signal processing enjoys the rich theory of linear systems. Linear filters are also simple to implement.

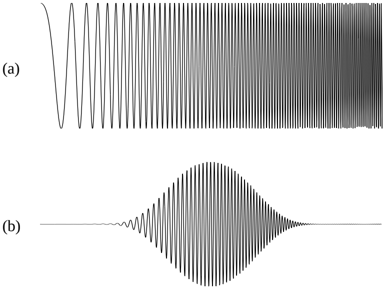


Figure: Frequency selective filtering: (a) chirp signal, (b) linear FIR filter output.

Why Nonlinear Signal Processing

- The underlying processes of interest may tend to produce outliers not predicted by a Gaussian model.
- The signal density functions may have tails that decay at rates lower than those of a Gaussian distribution.
- Linear methods obeying the superposition principle suffer serious degradation upon the arrival of samples corrupted with high-amplitude noise.
- Nonlinear methods exploit the statistical characteristics of the noise to overcome the limitations of traditional signal processing.

Consider again the bandpass filtering example using a chirp signal degraded by non-Gaussian noise. The linear FIR filter output is severely degraded.

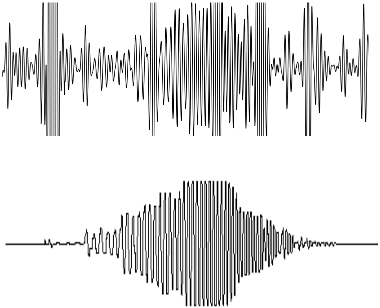


Figure: Frequency selective filtering in non-Gaussian noise: (a) linear FIR filter output, (b) nonlinear filter

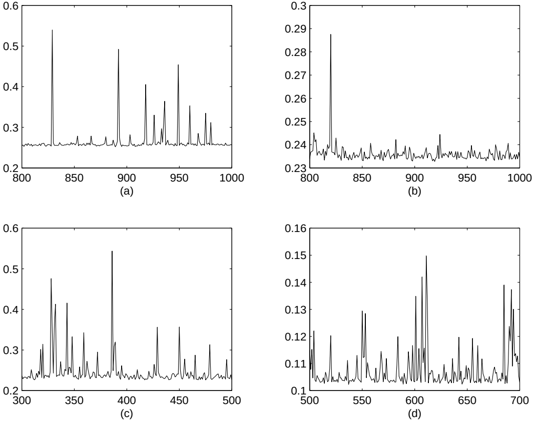


Figure: RTT time series measured in seconds between a host at the University of Delaware and hosts in (a) Australia (12:18 AM - 3:53 AM); (b) Sydney, Australia (12:30 AM - 4:03 AM); (c) Japan (2:52 PM - 6:33 PM); (d) London, UK (10:00 AM - 1:35 PM). All plots shown in 1 minute interval samples.

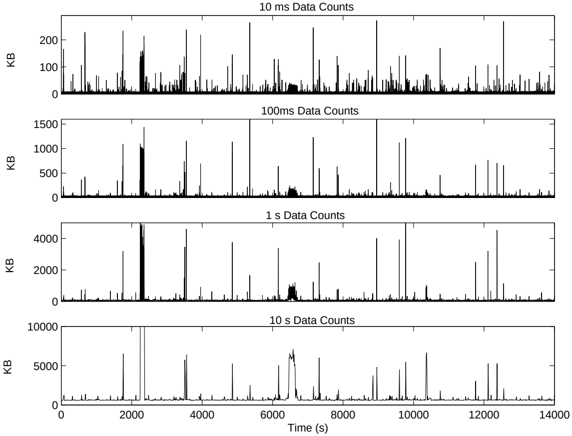


Figure: Byte counts measured over 14000 seconds in a web server of the ECE Department at the University of Delaware viewed through different aggregation intervals: from top to bottom, 10ms, 100ms 1s, 10s.

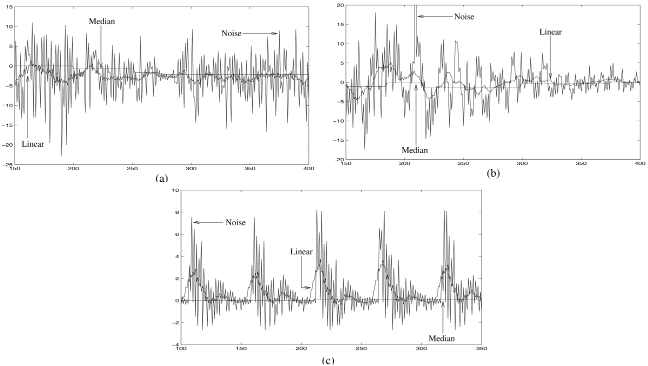


Figure: (a-c) Different noise and interference characteristics in ADSL lines. A linear and a nonlinear filter (recursive median filter) are used to overcome the channel limitations, both with the same window size.

Figure: JPEG compressed image

Figure: Output of unsharp masking using FIR filters

Figure: Output of median sharpeners.

Outline

1. Introduction
   - Non-Gaussian Random Processes
     - Generalized Gaussian Distributions and Weighted Medians
     - Stable Distributions and Weighted Myriads
   - Statistical Foundations
   - The Filtering Problem
     - Moment Theory

Non-Gaussian Random Processes

The Gaussian model appears naturally in many applications as a result of the Central Limit Theorem.

Theorem (Central Limit Theorem)

Let $X_1, X_2, \cdots$, be a sequence of i.i.d. random variables with zero mean and variance σ^2. Then as $N \rightarrow \infty$, the normalized sum

$$S_N = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} X_i \qquad (1)$$

converges almost surely to a zero-mean Gaussian variable with the same variance as X_i.

- In a wide range of applications, the Gaussian model does not produce a good fit which, at first, may seem to contradict the principles behind the central limit theorem.
- In order for the CLT to be valid, the variance of the superimposed random variables must be finite.
- If the random variables possess infinite variance, it can be shown that the series in the central limit theorem converges to a non-Gaussian impulsive variable.

  The generalized central limit theorem explains the presence of non-Gaussian, infinite variance processes, in practical problems.

In this course we will considers two model families that encompass a large class of random processes with different tail characteristics:

- *generalized Gaussian* distribution
- *stable* distributions

The tail of a distribution can be measured by the mass of the tail (*order*), defined as $\mathrm{P_r}(X > x)$ as $x \to \infty$.

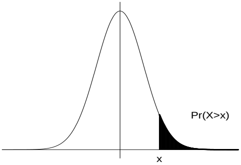


Figure: Mass of the tail of a Gaussian distribution

Generalized Gaussian Distributions and Weighted Medians

A first approach starts with a Gaussian distribution and allow the exponential rate of tail decay to be a free parameter. The result is the generalized Gaussian density function.

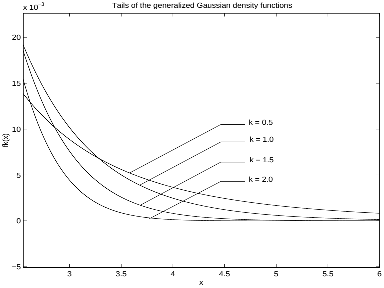


Figure: Tails of the Generalized Gaussian density functions for different k (Rate of tail decay).

- Of special interest is the Laplacian distribution.
- Weighted median filters are optimal for samples obeying Laplacian statistics.

  In general, weighted median filters are more efficient than linear filters in impulsive environments.

Stable Distributions and Weighted Myriads

A wide variety of processes exhibit statistics more impulsive than exponential that are characterized with algebraic tailed distributions:

- Radar clutter is the sum of signal reflections from irregular surfaces.
- Transmitters in a multiuser communication system generate small independent signals, the sum of which reaches a user's receiver.
- Rotating electric machinery generates impulses caused by contact between distinct parts of the machine.
- Standard atmospheric noise is the superposition of electrical discharges caused by lightning activity around the Earth.

The justification for using stable distribution models lies in the generalized central limit theorem which includes the well known “traditional” CLT as a special case.

> A random variable X is stable if it can be the limit of a normalized sum of i.i.d. random variables.

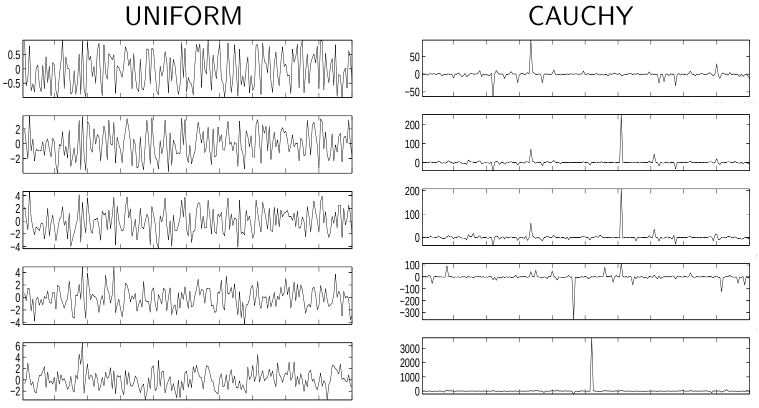


Figure: Traditional Vs. generalized CLT. The plots show the normalized sum of 1, 2, 3, 10 and 30 uniform(-1,1) or Cauchy(0,1) random variables

Outline

1. Introduction
   - Non-Gaussian Random Processes
     - Generalized Gaussian Distributions and Weighted Medians
     - Stable Distributions and Weighted Myriads
   - Statistical Foundations
   - The Filtering Problem
     - Moment Theory

Statistical Foundations

Estimation theory is a branch of statistics.
Goal: Derive information of a random processes from a set of observed samples.

- Given an observation $\{X(n)\}$, the goal is to extract the information embedded within the observed signal.
- In general some parameter β of the signal represents the information of interest. This parameter may be the local mean, variance, the local range, etc.

Location Estimation

Observed signals are random variables described by a probability density function (pdf), $f(x_1, x_2, \cdots, x_N)$.

- The pdf is usually parameterized by an unknown parameter β.
- β defines a class of pdfs where each member is defined by a particular value.

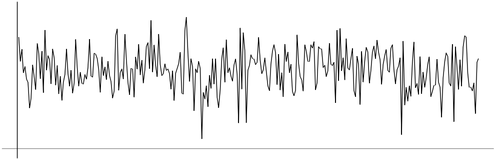


Figure: A Gaussian noise sample.

Example

If our signal consists of a single point ($N = 1$) and β is the mean, the pdf of the data under the Gaussian model is

$$f(x_1;\beta) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{1}{2\sigma^2}(x_1-\beta)^2\right] \tag{2}$$

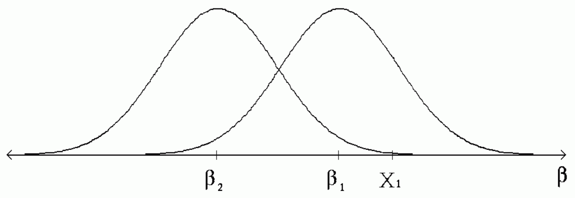
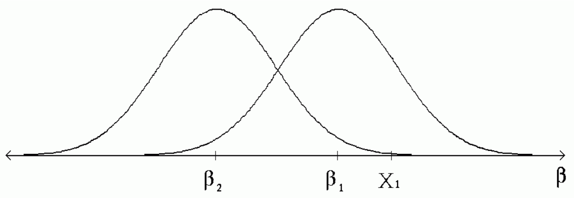


FIGURE: Estimation of parameter β based on the observation X_1.

- The value of β affects the probability of X_1. We should be able to infer the value of β from the observed value of X_1
- Notice that β determines the location of the pdf. As such, β is referred to as the *location parameter*.
- Rules that infer the value of β from sample realizations of the data are known as *location estimators*.



Running Smoothers

The *running mean* is the simplest form of filtering. Given the sequence $\{\cdots, X(n-1), X(n), X(n+1), \cdots\}$, it is defined as

$$Y(n) = \mathrm{MEAN}(X(n-N), X(n-N+1), \cdots, X(n+N)). \tag{3}$$

- The output is the average of the samples within a window centered at n.
- At each point the running mean computes a location estimator: the sample mean
- If the underlying signals are not Gaussian, the mean should be replaced by a more appropriate location estimator such as the running median.

Outline

1 Introduction
- Non-Gaussian Random Processes
  - Generalized Gaussian Distributions and Weighted Medians
  - Stable Distributions and Weighted Myriads
- Statistical Foundations
- The Filtering Problem
  - Moment Theory

The Filtering Problem

A filter can have arbitrary input and output signals and consequently it is found in a wide range of disciplines including economics, engineering, and biology.
Denote a random sequence as $\{X\}$ and let $\mathbf{X}(n) \subseteq R^N$ be a N-long element observation vector

$$\begin{aligned} \mathbf{X}(n) &= [X(n),\ X(n-1),\ \cdots, X(n-N+1)]^T \\ &= [X_1(n),\ X_2(n),\ \cdots,\ X_N(n)]^T \end{aligned} \quad (4)$$

where $X_i(n) = X(n-i+1)$. If the vector $\mathbf{X}(n)$ is related to a desired signal $D(n)$, the output of the filter is $\hat{D}(n)$.

The optimal filtering problem thus reduces to minimizing the cost function associated with the error e(n) under an error criterion.

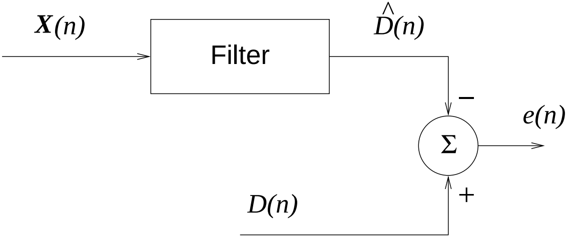


Figure: Filtering as a joint process estimation

Under Gaussian statistics, the estimation becomes linear and the filters reduce to FIR linear filters. The linear filter output is defined as

$$Y(n) = \mathrm{MEAN}(W_1 \cdot X_1(n), W_2 \cdot X_2(n), \cdots, W_N \cdot X_N(n)), \qquad (5)$$

where the $W_i's$ are real-valued weights assigned to each input sample.

Under the Laplacian model, the median becomes the estimate of choice and weighted medians become the filtering structure. The output of a weighted median is defined as

$$Y(n) = \mathrm{MEDIAN}(W_1 \diamond X_1(n), W_2 \diamond X_2(n), \cdots, W_N \diamond X_N(n)), \tag{6}$$

where the operation $W_i \diamond X_i(n)$ replicates the sample $X_i(n)$, W_i times, for example:

$$\mathrm{MEDIAN}(2 \diamond 5, 1 \diamond 0, 2 \diamond -1) =$$
$$\mathrm{MEDIAN}(5, 5, 0, -1, -1) = 0$$

For stable processes, the weighted myriad filter emerges as the ideal structure. In this case the filter output is defined as

$$Y(n) = \mathrm{MYRIAD}\left(k;\ W_1 \circ X_1, W_2 \circ X_2, \cdots, W_N \circ X_N\right), \tag{7}$$

where $W_i \circ X_i(n)$ represents a nonlinear weighting operation to be described later, and k in (7) is a free tunable parameter that will play an important role in weighted myriad filtering.

Moment Theory

Historically, signal processing has relied on second order moments. The first-order moment

$$\mu_X = E\{X(n)\} \tag{8}$$

and the second-order moment characterization provided by the *autocorrelation* of stationary processes

$$R_X(k) = E\left\{X(n)X(n+k)\right\}, \tag{9}$$

are deeply etched into traditional signal processing practice, for example, the PSD of a Gaussian random process is found as:

$$S(f) = \mathcal{F}\left\{R_X\right\},$$

where $\mathcal{F}$ represents the Fourier transform.

Second-order descriptions do not provide adequate information to process non-Gaussian signals. The alternatives are:

- *Higher-order* statistics: Moments of order greater than two.

$$E\{|X(n)|^p\} \qquad \text{for } p > 2 \tag{10}$$

  They provide information that is unaccessible to second-order moments, but become less reliable in impulsive environments.

- *fractional lower-order statistics* (FLOS) consisting of moments of orders less than two.

$$E\{|X(n)|^p\} \qquad \text{for } p < 2 \tag{11}$$