ECSE 6520: Estimation and Detection Theory

What is Statistical Signal Processing?

Class Notes - 1

January 26, 2014

Contents

1 What is Statistical Signal Processing?

- Signal - A function of one or more parameters, $f(x)$, $x \in \mathbb{R}^N$, $\mathbb{C}^N$, $\mathbb{Z}^N$, etc.
- Digital - Discrete time, quantized, sampled, $f(x)$, $x \in \mathbb{Z}^N$.
- Processing - Analyzing, predicting, decision making, smoothing.
- Statistics - Theory of modeling and decision making under uncertainty.

Digital Signal Processing - Processing signals with computer algorithms.

Statistical Signal Processing - Processing signals with computer algorithms under uncertainty.

Examples of signals -

- Speech signal
- Radar or sonar returns
- Pixelized image
- Number of photons arriving to an EM detector
- Dow-Jones industrial average
- Stream of internet packets
- EEG signals
- Seismic data

Processing - Analyze, infer, predict, identify, smooth, reconstruct etc. These processes can be categorized under

- Detection and Classification

- Estimation
- Filtering, Prediction

Difficulty - Many unknowns and uncertainties. We do not have perfect or complete knowledge of the signal, the medium it interacts with, or the system/process that generates it.

Examples -

- Radar/Sonar
  - The Green's function of the multi-path rich propagation medium for radar/sonar signals.
  - Electronic noise at the receiving antennas.
  - Ambient sound or EM waves.
  - Interference or jamming.
  - Clutter.
- Communications
  - Unknown "channel".
  - Interferences.
  - Ambient or electronic noise.
- Medical Imaging
  - Photon limited imaging systems. X-ray fluoroscopy, PET, SPECT $\Rightarrow$ Low Signal-to-Noise Ratio (SNR).
  - Additive electronic noise.

- Blurred or grainy astronomical images.
- Limited aperture data.

- Dynamically changing medium
  - Multi-target dynamically changing scene.
  - Internet traffic.
  - Dow-Jones industrial average.
  - Functional imaging. MR, optical or PET imaging.

How to process under uncertainty ? - How do we process signals in the presence of unknown variables and uncertainty? Can we model uncertainty and incorporate that to our processing? Statistical signal processing is the study of these questions.

How to model uncertainty ? -

- Most widely accepted and commonly used approach is probability theory.
- Probability theory models uncertainty by specifying the chance of observing certain signals.
- Specify the degree to which we believe a signal reflects the true state of nature.

Examples -

- Electronic noise modeled as additive *Gaussian* process.
- Uncertainty in the phase of a signal modeled as *uniform* random variable over $[0, 2\pi)$.
- Uncertainty in the number of photons striking a PET detector per unit time modeled as *Poisson* process.

- Uncertainty in the number of data packets at a node per unit time modeled as *Poisson* process.

What is statistical inference? -

A statistic is a function of observed data. It can be scalar or vector valued.

Examples - Suppose we observe n scalar values $x_1, x_2, ..., x_N$. The following are statistics:

- Sample mean – $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$.
- Sample variance – $\hat{\sigma^2} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$.
- The data itself – $[x_1, x_2, ..., x_N]^T$.
- Median of the observations –
- Any function of the observations – $f(x_1, x_2, ..., x_N)$.

Probability describes/models the uncertainty we observe in data. Statistics describe the salient features of the data we observe (statistics cannot depend on unknown quantities). It allows us to make inferences as to which probability model reflects the true nature of the signal.

What is Statistical Signal Processing?

- Step 1 – Postulate the probability models that can reasonably capture the uncertainties in the observed data.
- Step 2 – Extract statistics from the data that allows us to make inferences.
- Step 3 – Make an inference with respect to some optimality criteria.
- Step 4 – Validate the "goodness" of your inference by means of mathematical statistics.

2 Detection/Classification

What is Detection? -

- Given two (or more) probability models, which one best explains the data?
- Alternatively, given a single probability model for the data, is it a valid characterization of the data?
- More precisely - Two-class hypothesis testing. Decide between two possible distinct outcomes. Determine the value of a parameter that takes values from the set $A = \{0, 1\}$.
- Decide presence or absence of a radar/sonar target in a scene.
- Detect presence of a tumor in an X-ray, MR or ultrasound image.
- Decode communication symbols into 0s or 1s.
- Segment an image into foreground and background.
- Detect edges in an image.

Example 1 - A modem transmits a constant waveform with amplitude equal to 1 to represent the symbol 1 and transmits another constant waveform with amplitude equal to -1 to represent the symbol 0. The modem is equally likely to transmit either 1 or -1. Suppose the received signal is corrupted with additive noise with variance σ^2. Design a detector/test at the receiver to decode the symbols.

- Step 1 - Assume the received signal is in the following form - $x_n = s_n + w_n$ where s_n denotes the constant waveform transmitted at the nth time instance. s_n takes the value of either 1 or -1 with equal probability. w_n denotes the additive electronic noise. Assume that the additive noise can be modeled by a Gaussian distribution with zero mean and a known variance equal to σ^2. See that given s_n, x_n is Gaussian distributed with mean s_n and variance equal to σ^2.

- Step 2 - Use the data itself as a statistic.
- Step 3 - Make a decision with respect to an optimality criteria. Heuristically, we could decide that when x_n is less than 0, the symbol is -1, otherwise 1. Of course, in this course we will developed principled and mathematically rigorous approaches to this problem using the framework of probability and statistics.
- Step 4 - Determine what is the expected error rates at the receiver side. What is the probability that the detector decides 1 when the true transmitted signal is -1?

Example 2 - Suppose you want to decide whether a coin is a fair coin or not.

- Step 1 - Assume the outcome of a coin toss can be modeled with as Bernoulli random variable with parameter p. If $p = 1/2$, then the coin is fair.
- Step 2 - Toss the coin 100 times. Let X be the number of heads out of 100 tosses.
- Step 3 - Decide that the coin is fair if $45 \leq X \leq 55$.
- Step 4 - Use the probability model in Step 1 to determine the probability of making an incorrect decision?

What is Classification? -

- Two or more class hypothesis testing.
- Decide between N number of distinct outcomes. Determine the value of a parameter that takes values from the set $A = \{0, 1, 2, 3, ...N - 1\}$.
- Determine the type of radar target from high range resolution (HRR) radar data. Tank, pick-up truck, school bus, etc.

- Determine the stage of a malignant tumor.
- Determine the characters from an optically scanned document.
- Automated finger print identification.
- Determine layers of Earth's crust.

3 Estimation

What is Estimation? -

- Identify unknown, information bearing parameters based on physics or mathematical models.
- Determine a parameter or a function that takes values NOT necessarily from a finite set. Determine the value of a function or parameter that takes values from the set $\mathbb{R}$, $\mathbb{R}^N$, $\mathbb{C}$, $[0,1]$, etc.
- Detection/classification theory involves making a choice over some countable (usually finite) set of options, while estimation involves making a choice over a continuum of options.
- Reconstruct a radar/sonar image from fast- and slow-time radar/sonar signals.
- Reconstruct a positron emission tomography (PET) image.
- Estimate the range and Doppler of a radar/sonar target.
- Estimate the impulse response function of a communication channel.
- Estimate, extract features from an image (size of a heterogeneity, location, contrast level etc.)

Example 1 - Suppose you have a noisy and blurred image x and would like to estimate the true image s from the observed image x.

- Step 1 - Let the observed image x have the following probability model: $x(n,m) = h * s(n,m) + w(n,m),\ n = 1,...,N,\ m = 1,...,M$ and * denotes convolution and w denotes additive noise. Furthermore, assume that the model for the blur h is a three-by-three symmetric low pass filter with known coefficient and w is a Gaussian process with known mean and variance.

- Step 2 and 3 - Determine a function or a transformation of the observed image that will produce an estimate of the true image s.

- Step 4 - Use the probability model in Step 1 and the estimator developed in S tep 2 and 3 to determine the level of noise or blurring left in the estimated true image.

Example 2 - Suppose we measure the voltage A of a battery with a voltmeter. Because the voltmeter tends to pick-up noise, we make N measurements in hopes of gaining some accuracy.

- Step 1 - Assume the measured signal x_n obeys the following model: $x_n = A + \omega_n$, $n = 1,...,N$ where ω_n is a Gaussian random process with zero mean and σ^2 variance (known).

- Step 2 and 3 - We gather data and estimate A by sample mean $\hat{A} = \frac{1}{N}\sum_{n=1}^{N} x_n$.

- Step 4 - Use probability theory to determine the number of measurements needed to assure a certain predetermined accuracy in your estimate.

Example 3 - If $s(n) = Acos(2\pi f_n + \phi)$, estimate $A,\ f_n,\ \phi$ from the noisy measurement $x(n) = s(n) + \omega(n)$.

What is Filtering, Prediction? -

- In statistics, it goes under the name time series analysis
- Control the trajectory of a moving aircraft.
- Track the trajectory of a moving radar/sonar target.
- Predict future Dow-Jones averages.
- Predict potential faults or future failures in a system (X-ray system, power generators, high voltage electrical machinery).
- Predict internet traffic load.

4 Statistics versus Engineering

Differences in terminology and culture

Statistical inference	Estimation, detection theory
Time series analysis	Filtering, prediction
Hypothesis testing	Detection and classification theory
Point/parameter estimation	Estimation theory
Normal distribution	Gaussian distribution
Type I error, false positive	False alarm
Type II error, false negative	Miss
Test	Detector
Size, significance level	False alarm rate
Power	Detection rate