SIGNAL STATISTICS AND DETECTION

Professor Andrew E. Yagle, EECS 206 Instructor, Fall 2005

Dept. of EECS, The University of Michigan, Ann Arbor, MI 48109-2122

I. Abstract

The purpose of this document is to introduce EECS 206 students to signal statistics and their application to problems in signal detection and time delay estimation. It will be useful for the first couple of laboratories.

A. Table of contents by sections:

II. Definitions

This section defines various statistical properties of signals, and gives simple examples of how to compute them directly from the signal itself.

A. Basic Properties of Signals

First, a quick review of **intervals**. We define the following four cases:

- The *closed* interval $[a, b] = \{t : a \leq t \leq b\}$; The *half-open* interval $(a, b] = \{t : a < t \leq b\}$;
- The *half-open* interval $[a, b) = \{t : a \leq t < b\}$; The *open* interval $(a, b) = \{t : a < t < b\}$.

If a or b is $\pm\infty$, use a half-open or open interval. Never write (even implicitly) "$t = \infty$"–this is uncouth! Now let $x(t)$ be a continuous-time signal. We define the following *basic properties* of $x(t)$:

- The **support** of $x(t)$ is the interval $[a, b] = \{t : a \leq t \leq b\}$ such that $x(t) = 0$ for all $t < a$ and for all $t > b$.
- $x(t)$ can be zero *somewhere* within the support $[a, b]$. $x(t)$ is zero *everywhere* outside the support $[a, b]$.
- The **duration** of $x(t)$ having **support** $[a, b]$ is $b - a$. The duration of $x(t)$ is "how long it lasts";
- A unit step $u(t)$ has support $[0, \infty)$. A sinusoid $A\cos(\omega t + \theta)$ has support $(-\infty, \infty)$ and duration$\rightarrow \infty$.
- The **maximum** of $x(t)$ is its largest value. The **minimum** of $x(t)$ is its smallest value.

Special case: Consider $x(t) = \begin{cases} \frac{2}{t+1} - 1 & \text{for } 0 < t < \infty \\ 0 & \text{for otherwise} \end{cases}$ The *support* of $x(t)$ is $(0, \infty)$ and its *duration*$\rightarrow \infty$, but $x(t)$ does not have a *minimum* or *maximum* value. We say that its **infimum**=-1 and its **supremum**=1, even though there is no value of t for which $x(t)$ actually attains these values (remember $x(0) = 0$).

B. Statistical Properties of Signals

Let $x(t)$ be a real-valued signal with *support* $[a, b]$. We define the following *statistical properties* of $x(t)$:

- **Mean:** $M(x) = \bar{x}(t) = \frac{1}{b-a}\int_a^b x(t)dt$=average value of $x(t)$.
- **Mean Square:** $MS(x) = M(x^2) = \overline{x^2}(t) = \frac{1}{b-a}\int_a^b x^2(t)dt \geq 0$.
- **rms:** $rms(x) = \sqrt{MS(x)} = \sqrt{\frac{1}{b-a}\int_a^b x^2(t)dt} \geq 0$. "rms" stands for Root-Mean-Square.
- **Variance:** $\sigma_x^2 = MS(x - M(x)) = \frac{1}{b-a}\int_a^b [x(t) - M(x)]^2 dt \geq 0$. Remember $M(x)$ is a *number*.
- Easier to compute: $\sigma_x^2 = M[(x - M(x))^2)] = M(x^2) + [M(x)]^2 - 2M(x)M(x) = MS(x) - (M(x))^2 \geq 0$.
- **Standard Deviation:** $\sigma_x = \sqrt{\sigma_x^2}$; we can always compute this since $\sigma_x^2 \geq 0$.
- **Energy:** $E(x) = \int_a^b x^2(t)dt \geq 0$. Note that **energy=duration$\times$MS(x)**.
- **Average power**=MS(x) as defined above. **Instantaneous power**=$x^2(t)$.
- Note the units: Power=$\frac{\textbf{ENERGY}}{\textbf{TIME}}$. EX: Watts=$\frac{\textbf{JOULES}}{\textbf{SECONDS}}$.

C. Periodic Signals

If $x(t)$ is a **periodic** signal, then $x(t) = x(t+T) = x(t+2T) = \ldots$ for *all time* t (NOT just for $t \geq 0$). T is the **period** of $x(t)$. The official EECS 206 lecture notes call T the *fundamental period* and let any integer multiple of T be "a period" of $x(t)$. The idea is that if a signal is periodic with period T, it can also be regarded as periodic with period kT for any integer k. While this is true, don't try it during a job interview!

Periodic signals have the following properties:

- **Support**=$(-\infty, \infty)$ and **duration** $\rightarrow \infty$;   **Energy** $\rightarrow \infty$ (unless $x(t) = 0$ for all t);
- All of the above definitions (except energy) still apply: just use **[a,b]=one period**.

D. Simple Examples

Problem #1: $x(t) = \begin{cases} 2t & \text{for } 0 \leq t \leq 3 \\ 0 & \text{otherwise} \end{cases}$. Compute the basic and statistical properties of $x(t)$.

Solution #1: Plug:

- **Support**=[0,3]. **Duration**=3-0=3. **Min**=0. **Max**=6. All by inspection.
- **Mean**=$M(x) = \frac{1}{3-0}\int_0^3 (2t)\, dt = 3$. **Mean square**=$MS(x) = \frac{1}{3}\int_0^3 (2t)^2 dt = 12$.
- **Variance**=$\sigma_x^2 = MS(x - M(x)) = \frac{1}{3}\int_0^3 (2t-3)^2 dt = 3$ (hard way–that integral requires too much work).
- **Variance**=$\sigma_x^2 = MS(x) - (M(x))^2 = 12 - (3)^2 = 3$ (easy–both $MS(x)$ and $M(x)$ are much easier integrals).
- $\frac{\textbf{STANDARD}}{\textbf{DEVIATION}} = \sigma_x = \sqrt{3}$.   **rms**=Root-Mean-Square=$\sqrt{MS(x)} = \sqrt{12}$.
- **Energy=duration$\times$MS(x)**=(3)(12) = 36.

Use of the *Cauchy Principal Value* (anyone who spells this "principle" will be sent to the "principal"!):

- <u>Problem #2</u>: Compute the mean $M(x)$ of $x(t) = 1 + e^{-|t|}$.
- <u>Solution #2</u>: The signal has **support**=$(-\infty, \infty)$ and **duration** $\to \infty$.
- So $M(x) = \frac{1}{\infty}\int_{-\infty}^{\infty}(1+e^{-|t|})dt$. Uh-oh! NOW what do you do?
- Use the *Cauchy Principal Value*: compute for support=$[-T, T]$ and let $T \to \infty$:
- $M(x) = \underset{T\to\infty}{\text{LIM}} \frac{1}{2T}\int_{-T}^{T}(1+e^{-|t|})dt = \underset{T\to\infty}{\text{LIM}} \frac{2}{2T}\int_{0}^{T}(1+e^{-t})dt = \underset{T\to\infty}{\text{LIM}} \frac{1}{T}(T+1-e^{-T}) = 1$
- <u>We used:</u> Symmetry to get rid of $|t|$; the limit should be evident since e^{-T} is bounded by one.

We actually won't be doing this in EECS 206; toy problems like this don't happen in the real world.

Discrete-time signals: Just use sums instead of integrals, but there is one major change in **duration**.

- <u>Problem #3</u>: $x[n] = \{\underline{3}, 1, 4, 2, 5\}$. Compute the basic and statistical properties of $x[n]$.
- The underline denote time $n = 0$: $\underline{3}$ means that $x[0] = 3$. Then $x[1] = 1, x[2] = 4$, etc.
- **Note: Duration** of signal with **support** $[a, b]$ is duration=$(b - a + 1)$, **not** $(b - a)$!
- **Why?** The number of integers between a and b *inclusive* is $(b - a + 1)$, not $(b - a)$.
- <u>Solution #3</u>: Just use sums instead of integrals in above definitions–makes things easier!
- **Support**=[0,4]. **Duration**=$4 - 0 + 1 = 5$. **Min**=1. **Max**=5. All by inspection.
- **Mean**=$M(x) = \frac{1}{5}(3 + 1 + 4 + 2 + 5) = 3$. **Mean square**=$MS(x) = \frac{1}{5}(3^2 + 1^2 + 4^2 + 2^2 + 5^2) = 11$.
- **Variance**=$\sigma_x^2 = MS[x - M(x)] = \frac{1}{5}[(3-3)^2 + (1-3)^2 + (4-3)^2 + (2-3)^2 + (5-3)^2] = 2$ (hard way).
- **Variance**=$\sigma_x^2 = MS(x) - (M(x))^2 = 11 - (3)^2 = 2$ (the easy way).
- **STANDARD DEVIATION** $= \sigma_x = \sqrt{2}$. **rms**=Root-Mean-Square=$\sqrt{MS(x)} = \sqrt{11}$.
- **Energy**=$(3^2 + 1^2 + 4^2 + 2^2 + 5^2) = 55$. **Average power**=MS(x)=11.

III. Properties

Why bother with even more math? Because we will use ALL of these in the applications below.

A. Correlation

Three more definitions: the **correlation** between two real-valued signals is:

$$C(x,y) = \int_{-\infty}^{\infty} x(t)y(t)dt. \quad \text{If period = T, then :} \quad C(x,y) = \int_{0}^{T} x(t)y(t)dt$$
$$C(x,y) = \sum_{n=-\infty}^{\infty} x[n]y[n]. \quad \text{If period = N, then :} \quad C(x,y) = \sum_{n=0}^{N-1} x[n]y[n] \tag{1}$$

and the **normalized correlation** or **correlation coefficient** is:

$$C_N(x,y) = \rho(x,y) = \frac{C(x,y)}{\sqrt{C(x,x)C(y,y)}} = \frac{C(x,y)}{\sqrt{E(x)E(y)}} \tag{2}$$

Note that if *either* x or y has finite support, then the range of integration or summation in the definition of $C(x,y)$ becomes a finite interval, which is the intersection of the supports of x and y.

Two signals x and y are **uncorrelated** or **orthogonal** if their correlation $C(x,y) = 0$.

B. Properties

All of these assume that x and y are real-valued and have the same support.

- Means: $M(x+y) = M(x) + M(y)$: Mean of the sum is the sum of the means.
- Proof: $M(x+y) = \frac{1}{b-a}\int_a^b [x(t)+y(t)]\,dt = \frac{1}{b-a}\int_a^b x(t)\,dt + \frac{1}{b-a}\int_a^b y(t)\,dt = M(x) + M(y)$.
- Mean-Square: $MS(x+y) = MS(x) + MS(y)$ **if** x and y are **uncorrelated** (i.e., $C(x,y) = 0$).
- Proof: $MS(x+y) = M[(x+y)^2] = M(x^2) + M(y^2) + 2M(xy) = MS(x) + MS(y) + 2\frac{C(x,y)}{\textbf{duration}}$.
- Correlation: $C(x,x) = E(x); C(x,y) = C(y,x); C(x+y,z) = C(x,z) + C(y,z)$. Proof: $C(x,y)$ definition.
- Correlation: $|C_N(x,y)| \leq 1$ with equality if $y = ax$ for some constant a. Proof:

$$0 \leq \int \left(\frac{x}{\sqrt{E(x)}} - \frac{y}{\sqrt{E(y)}}\right)^2 dt = \frac{\int x(t)^2 dt}{E(x)} + \frac{\int y(t)^2 dt}{E(y)} - 2\frac{\int x(t)y(t)dt}{\sqrt{E(x)}\sqrt{E(y)}} = 2 - 2\frac{C(x,y)}{\sqrt{E(x)}\sqrt{E(y)}} \rightarrow C_N(x,y) \leq 1.$$

$$0 \leq \int \left(\frac{x}{\sqrt{E(x)}} + \frac{y}{\sqrt{E(y)}}\right)^2 dt = \frac{\int x(t)^2 dt}{E(x)} + \frac{\int y(t)^2 dt}{E(y)} + 2\frac{\int x(t)y(t)dt}{\sqrt{E(x)}\sqrt{E(y)}} = 2 + 2\frac{C(x,y)}{\sqrt{E(x)}\sqrt{E(y)}} \rightarrow C_N(x,y) \geq -1.$$

C. Interpretation

The property $|C_N(x,y)| \leq 1$ means that $C_N(x,y)$ can be regarded as a measure of similarity of x and y: To determine whether two signals x and y are similar, compute $C_N(x,y)$. The closer this is to unity (one), the more alike are the signals x and y. In fact, consider two finite-duration real-valued discrete-time signals $x[n]$ and $y[n]$, each having support $[1,N]$. Let $\underline{x} = [x[1], x[2] \ldots x[N]\}$ and $\underline{y} = [y[1], y[2] \ldots y[N]\}$. Then:

$$C(x,y) = \sum_{n=1}^{N} x[n]y[n] = \underline{x} \cdot \underline{y} = ||\underline{x}|| \cdot ||\underline{y}|| \cos\theta = \sqrt{||\underline{x}||^2}\sqrt{||y||^2}\cos\theta = \sqrt{C(x,x)C(y,y)}\cos\theta \tag{3}$$

which implies that

$$C_N(x,y) = \frac{C(x,y)}{\sqrt{C(x,x)C(y,y)}} = \cos\theta \tag{4}$$

- $C_N(x,y)$ is the cosine of the "angle" θ between the two signals x and y;
- If $y = ax$, then x and y are in the same direction, and $\theta = 0 \rightarrow C_N(x,y) = \cos(0) = 1$;
- If x and y are uncorrelated, then x and y are perpendicular and $\theta = 90^o \rightarrow C_N(x,y) = \cos(90^o) = 0$.
- This is why **uncorrelated** signals are also called **orthogonal** signals!
- A good example of how an abstract concept (the angle between two signals?!) can be useful and insightful.

IV. Histograms

In the real world, we don't know what a noise signal is. But often (see below), we don't *need* to know it; we only need to know its above statistical properties. We can obtain good estimates of them from a **histogram** of the signal, as we now demonstrate.

A. What is a Histogram?

Procedure for constructing the histogram of a known real-valued finite-duration discrete-time signal $x[n]$:

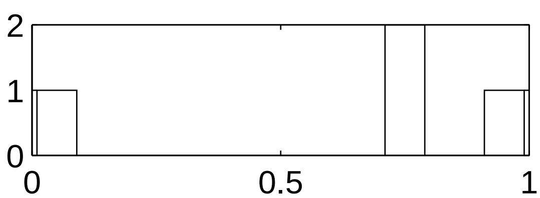
1. Determine $L = min(x)$ (L for Lower) and $M = max(x)$ (M for Maximum);
2. Divide up the interval $[L, M]$ into N subintervals, each of length $\frac{M-L}{N}$;
3. These N subintervals are called **bins**. Typical values of N: 20,50,100.
4. Count the number of times the value of $x[n]$ lies in a given bin. Repeat for each bin;
5. Plot: Vertical axis=#times value of $x[n]$ lies in a bin. Horizontal axis: bins.

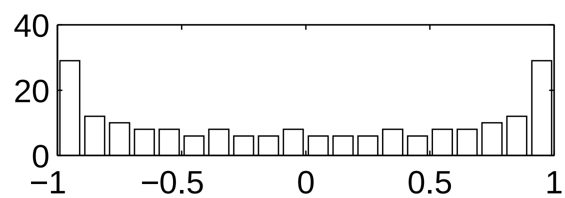
A toy example, from the official lecture notes, of computing a histogram:

- Problem: Plot the histogram of $x[n] = |\cos(\frac{\pi}{4}n)|$ using 10 bins.
- $x[n] = |\cos(\frac{\pi}{4}n)| = \{\ldots \underline{1}, \frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}, 1, \frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}, 1 \ldots\}$;
- So $x[n]$ is periodic with period=4 (not 8, due to the absolute value);
- In one period: $x[n]$ has value 0 once, value $\frac{\sqrt{2}}{2}$ twice, and value 1 once;
- $L = min(x) = 0$; $M = max(x) = 1$; each bin has width $\frac{1-0}{10} = 0.1$;
- Plot histogram: $x[n]$ value lies in $[0, 0.1]$ bin once, $[0.7, 0.8]$ bin twice, and $[0.9, 1]$ bin once (below left).
- Matlab: `X=abs(cos(pi/4*[0:3]));subplot(421),hist(X,10)`. Doesn't tell us anything useful.

A more realistic example (although we still have to know the signal):

- Problem: Plot the histogram of $y[n] = \cos(\frac{\pi}{100}n)$ using 20 bins;
- $y[n]$ has period=200→an average of ten values in each bin;
- $L = min(y) = -1$; $M = max(y) = 1$; each bin has width $\frac{1-(-1)}{20} = 0.1$;
- Plot histogram: See below right. Matlab: `X=cos(pi/100*[0:199]);subplot(422),hist(X,20)`
- Note that $y[n]$ spends most of its time near its peak values of ± 1. So what? See below.





B. What are Histograms for?

In real-world applications, our signal processing task (whatever it is) is complicated by the presence of *noise* added to our data. **Noise** can be defined (in EECS 206) as an undesired and unknown signal whose effects we wish to eliminate from our data.

If we knew what the noise signal was, we could just subtract it from our data and get rid of it. But since we don't know what it is, we can't do that! So what do we do?

We will see in the sections below that we can perform our signal processing tasks fairly effectively if we know only the **statistical properties** of the noise, not the noise signal itself. But how do we compute these statistical properties if we don't know the signal?

The answer is to compute the histogram of the noise during some time when we are not performing any tasks, and then use the histogram to compute the statistical properties of the noise. We will see how to do this below. The point is that even though the noise *signal* will be different when we are performing a task, the noise *statistics* will be the same (or almost the same).

Another reason to use histograms is for **coding** a signal. For example, the second histogram above revealed that $y[n]$ spent most of its time near its peaks. That means it is more important to represent the peak values accurately than the values near zero, since the peak values occur more often.

Detailed derivations of all of the algorithms below are given in EECS 401 and EECS 455.

C. Using Histograms to Compute Statistical Properties

Again, the simplest way to illustrate this is to use a simple example. Consider again the signal $x[n] = |\cos(\frac{\pi}{4}n)|$. We can estimate its mean directly from its histogram by assuming that the two values in the $[0.7, 0.8]$ bin are both in the middle of the bin at 0.75. Of course this is not true, but since a value is just as likely to lie above the bin midpoint as below it, the errors will tend to cancel out. We obtain:

$$\begin{aligned}
M(x): \text{ from histogram} &= 0.612 = & \frac{1}{4}[1(0.05) + 2(0.75) + 1(0.95)] \\
M(x): \text{ its actual value} &= 0.604 = & \frac{1}{4}[1 + \sqrt{2}/2 + 0 + \sqrt{2}/2] \\
MS(x): \text{ from histogram} &= 0.507 = & \frac{1}{4}[1(0.05)^2 + 2(0.75)^2 + 1(0.95)^2] \\
MS(x): \text{ its actual value} &= 0.500 = & \frac{1}{4}[1 + (\sqrt{2}/2)^2 + 0 + (\sqrt{2}/2)^2]
\end{aligned} \tag{5}$$

For more realistic problems, the results are even better than they are for this small toy example.

V. Application: Reading your CD's Pits

A. Introduction

When a CD is burned (and some of them should be! :)), what happens physically is that a strong laser carves pits (or not) into its surface, along a track. The presence or absence of a pit designates a bit value of 0 or 1. Groups of 16 bits each represent a single 16-bit binary number elvis[n], which in turn come from sampling the continuous-time audio signal elvis(t) 44,100 times per second. This will be discussed in much more detail later in EECS 206 when we discuss sampling of signals.

When a CD is read, a weaker laser illuminates the track on the CD, either being reflected (or not) back to the CD reader. Again, the presence or absence of a pit designates a bit value of 0 or 1. Groups of 16 bits each represent a single 16-bit binary number elvis[n], etc.

This omits error-correcting codes, which recover bits missing because you spilled coffee on your CD.

However, the basic problem is to ensure that the sequence of bits (0 or 1) read by your CD player is the same as the sequence Elvis recorded in the studio. Reading laser reflections from pits or absence of pits is complicated by an additive and unknown noise signal $v(t)$. What can we do?

B. Algorithm

Let $y(t)$ be the strength of the laser reflection back to the CD reader, as a function of time t. 44,100 16-bit numbers per second results in a bit every $T = \frac{1}{16}\frac{1}{44{,}100} = 1.4\mu sec$. Then we can proceed as follows:

- Partition $y(t)$ into blocks. Each block is T seconds long and treated separately.
- Within each block, the **ideal** reflected signal $x(t)$ would be either $x(t) = 0$ or $x(t) = 1$ (after scaling).
- But due to the additive noise $v(t)$, the **actual** signal is $y(t) = x(t) + v(t)$, so $y(t) = v(t)$ or $y(t) = v(t) + 1$.
- **So:** within the k^{th} block, compute $\hat{M}(y) = \frac{1}{T}\int_{kT}^{(k+1)T} y(t)\,dt$= mean value of data in k^{th} block.
- If $\hat{M}(y) > \frac{1}{2}$, decide $x(t) = 1$ within the k^{th} block, i.e., the k^{th} bit=1.
- If $\hat{M}(y) < \frac{1}{2}$, decide $x(t) = 0$ within the k^{th} block, i.e., the k^{th} bit=0.
- Repeat for each block, generating a sequence of estimated bits that is close to the actual sequence.

C. Performance

How well does this work? We make the following assumptions (all of which are very good):

- Each bit is equally likely to be 0 or 1: Pr[x(t)=0]=Pr[x(t)=1]=$\frac{1}{2}$.
- Knowing the value of a bit tells you nothing about the value of any other bit.
- The noise signal $v(t)$ is zero-mean: $M(v) = 0$. Then $M(y) = M(x) + M(v)$=1 or 0.
- The noise signal $v(t)$ is uncorrelated with delayed versions of itself: $C(v(t)v(t-D)) = 0$ for $D \neq 0$.
- The statistical properties of the noise signal $v(t)$ do not change over time.

Within each block, there are two ways we can be wrong:

- Pr[choose x(t)=1 when x(t)=0]=$Pr[\hat{M}(y) > \frac{1}{2}]$=$Pr[\hat{M}(v) + 0 > \frac{1}{2}]$=$Pr[\hat{M}(v) > +\frac{1}{2}]$ (false alarm)
- Pr[choose x(t)=0 when x(t)=1]=$Pr[\hat{M}(y) < \frac{1}{2}]$=$Pr[\hat{M}(v) + 1 < \frac{1}{2}]$=$Pr[\hat{M}(v) < -\frac{1}{2}]$ (undetected)

Then the *probability of error* is (using Pr[x(t)=0]=Pr[x(t)=1]=$\frac{1}{2}$)

- Pr[error]=$\frac{1}{2}$Pr[choose $x(t)$=1 when $x(t)$=0]+$\frac{1}{2}$Pr[choose $x(t)$=0 when $x(t)$=1]
- Note that without the factor of $\frac{1}{2}$, we could get Pr[error]> 1!
- Simplifies to: $Pr[\text{error}] = \frac{1}{2}Pr[|\hat{M}(v)| > \frac{1}{2}]$, where $\hat{M}(v) = \frac{1}{T}\int_0^T v(t)\,dt$ and $v(t)$ is the *actual* noise.
- Don't confuse $M(v)$ (=0 by assumption) and $\hat{M}(v)$ (computed from the *actual* noise $v(t)$)
- They *should* be close to each other (due to *ergodicity*) but there is a small chance that $|\hat{M}(v)| > \frac{1}{2}$.

The point is that we can estimate $Pr[|\hat{M}(v)| > \frac{1}{2}]$ from the histogram of v *without knowing* v. And we can obtain the histogram of $v(t)$ by simply recording $y(t) = v(t)$ when $x(t) = 0$ (no CD in the player at all), computing the histogram of $y(t)$, and assuming the statistical properties of $v(t)$ don't change over time.

VI. Application: Are the whales mating?

A. Introduction

Suppose a marine biologist wants to know if the whales are mating. We know the sound $w(t)$ whales make when they are mating (don't ask how). But when we lower a hydrophone into the water, we hear $y(t) = \begin{cases} w(t)+v(t) & \text{if whales are mating;} \\ v(t) & \text{if whales not mating} \end{cases}$, where $v(t)$ is additive noise that drowns out the desired $w(t)$. The problem is to determine whether $w(t)$ is present or absent in our heard signal $y(t)$.

We could just compute and threshold $\hat{M}(y)$ as above. This would work (poorly) if $M(w) > 0$ and $M(v) = 0$; in fact, this *is* what we are doing when $w(t) = 1$ (which it isn't). But we should be able to take advantage of our knowledge of $w(t)$ to do better than that.

B. Algorithm

We make the following assumptions:

- The noise $v(t)$ is zero-mean: $M(v) = 0$;
- $v(t)$ is uncorrelated with delayed versions of itself: $C(v(t)v(t-D)) = 0$ for $D \neq 0$.

We already know that $C(w,w) = E(w) \geq 0$. Then we can:

- Compute $\hat{C}(w,y) = \int w(t)y(t)\,dt$ from the data $y(t)$ and the *known* signal $w(t)$;
- $\hat{C}(w,y) = \int w(t)y(t)\,dt = \begin{cases} C(w,w)+C(w,v) \approx E(w) & \text{if whales are mating;} \\ C(w,v) \approx 0 & \text{if whales not mating} \end{cases}$;
- If $\hat{C}(w,y) > \frac{1}{2}E(w)$, decide the whales are mating.
- If $\hat{C}(w,y) < \frac{1}{2}E(w)$, decide the whales aren't mating.
- The threshold $\frac{1}{2}E(w)$ assumes that it is equally likely that the whales are/aren't mating.
- We can adjust this threshold to bias this algorithm either way, if desired.

C. Performance

Again, there are two ways we can be wrong:

- Pr[choose whales mating when they are not]$=Pr[\hat{C}(w,v) > \frac{1}{2}E(w)] = Pr[\hat{C}(w,v) > \frac{1}{2}E(w)]$
- Pr[choose whales not mating when they are]$=Pr[C(w,w)+\hat{C}(w,v) < \frac{1}{2}E(w)] = Pr[\hat{C}(w,v) < -\frac{1}{2}E(w)]$
- $Pr[\text{error}] = \frac{1}{2}Pr[|\hat{C}(w,v)| > \frac{1}{2}E(w)]$ where $\hat{C}(w,v) = \int w(t)v(t)\,dt$.
- We can compute the probabilities for $C(w,v)$ from the histogram of v:
- Compute a derived histogram for $C(w,v)$ by changing bins for v to bins for $C(w,v)$, a function of v.

This is called a *matched filter* since the data are correlated with the signal to be detected. We will see later in the term that this can be computed by filtering the data through a system with impulse response $w(-t)$.

VII. Application: Time-Delay Estimation

A. Introduction

The basic idea behind radar and sonar is to transmit a signal $s(t)$, which reflects off something (a bird, a plane, a submarine, two mating whales) and returns to the sender. The signal $s(t)$ is a *pulse* having short *duration*, since the same antenna is used for transmitting and receiving, so that $s(t)$ must end before the reflected and delayed signal $s(t-D)$ arrives.

There is additive noise, so detecting the reflected and delayed $s(t)$ is not trivial. We observe

$$y(t) = s\left(t - \frac{2d}{c}\right) + v(t) \tag{6}$$

where d=distance to target and c=speed of light (radar) or sound in water (sonar) along path. Note that the time delay is the *two*-way travel time, since the pulse $s(t)$ must travel to the target and back again.

The next page shows three plots. The middle plot is the noisy $y(t)$. If this were an air traffic radar, and you had to decide where the plane is, how comfortable would you be? What if your grandmother were on the plane, coming to visit you? The top plot shows pulse $s(t)$ with its actual delay; was that your guess?

B. Algorithm

We make the following assumptions:

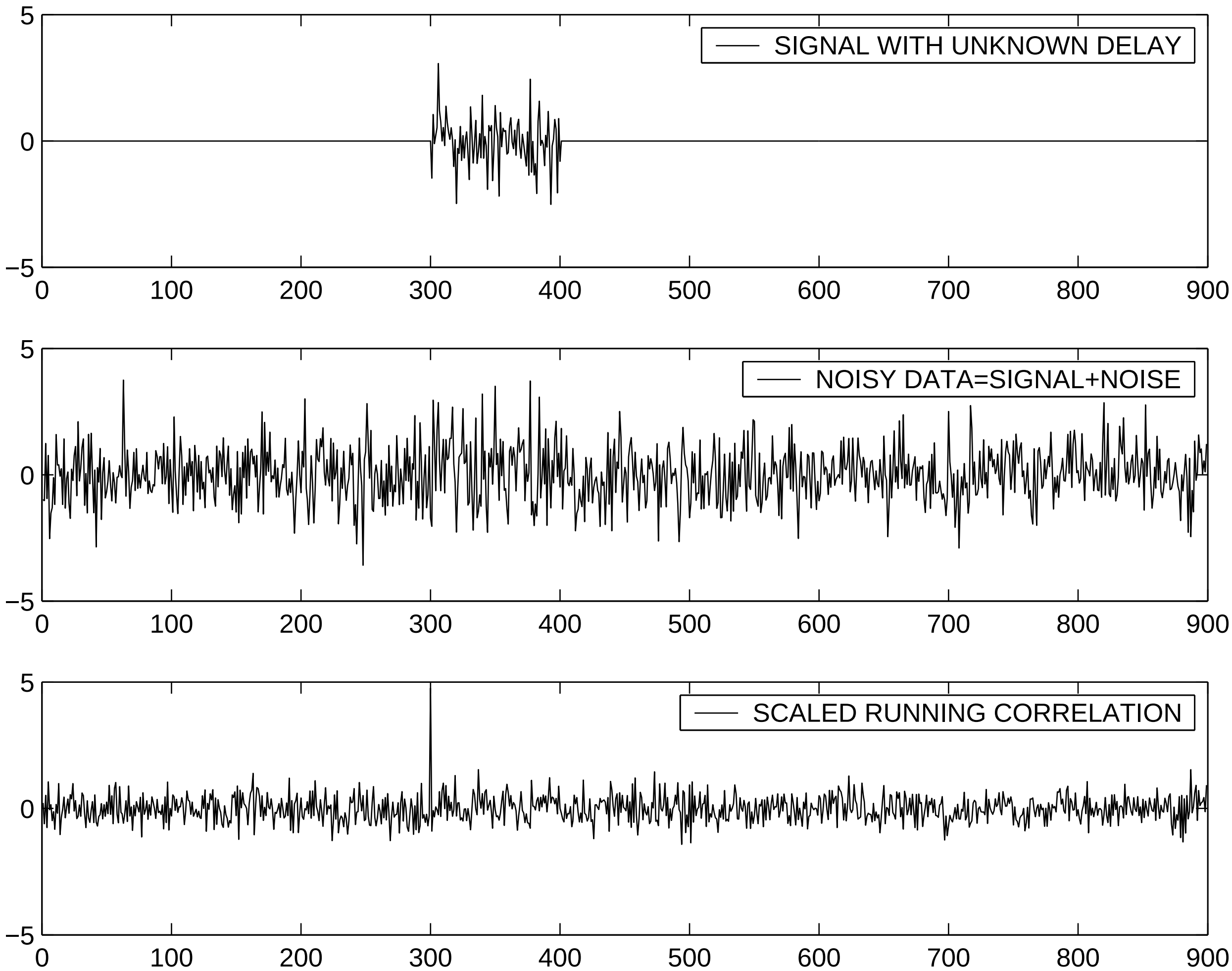
- $s(t)$ is uncorrelated with delayed versions of itself: $\int s(t)s(t-D)\,dt \approx 0$ for $D \neq 0$;
- $v(t)$ is uncorrelated with delayed versions of itself: $\int v(t)v(t-D)\,dt = 0$ for $D \neq 0$.

We already know $C(s,s) = E(s) \geq 0$. Then we can do the following:

- Compute the *running correlation* $\hat{R}C(t) = \int y(\tau)s(\tau - t)\,d\tau$ as a function of delay t;
- $\hat{R}C(t) = \int y(\tau)s(\tau - t)\,d\tau = \int [s(\tau - D) + v(\tau)]s(\tau - t)\,d\tau$ (note τ is a dummy variable)
- $= \int s(\tau - D)s(\tau - t)\,d\tau + \int v(\tau)s(\tau - t)\,d\tau \approx \int s(\tau)s(\tau + D - t)\,d\tau$ (neglecting $\int v(\tau)s(\tau - t)\,d\tau$)
- $\hat{R}C(t) \approx E(s)$ at $t = D$; and $\hat{R}C(t) \approx 0$ for $t \neq D$;
- <u>In summary:</u> Compute $\hat{R}C(t)$ and look for a peak. The location t of the peak is an estimate of D;

- Time delay estimation can be interpreted as matched filtering (see above):
- For each possible value of the time delay D, see whether the correlation with delayed $s(t)$ exceeds threshold.
- The "noise" in $\hat{R}C(t)$ is $\int v(\tau)s(\tau - t)\,d\tau$ (compare this to $C(w,v)$ in the previous section);

The third plot is a plot of running correlation $RC(t)$. Note the spike, which indicates very accurately what the delay is, and thus where the plane is! In the real world, the spike is not that sharp, since $s(t)$ is not completely uncorrelated with delayed versions of itself, and there are many other considerations in radar and sonar. But don't you feel better about your grandmother on the plane?



VIII. Application: Unknown Signal in Unknown Noise

A. Introduction

Say what? How the heck are you supposed to do that? Yet you can, with some assumptions.

Suppose we want to know whether the whales are mating, but we *don't know w(t)* (maybe we're prudes). All we know is that $w(t)$ is zero-mean, so we can't even use the algorithm we used for CD pits. We do know the *energies* of $v(t)$ and $w(t)$, and we know that $v(t)$ and $w(t)$ are uncorrelated (*the* assumption).

B. Algorithm

We make the following assumptions:

- The unknown noise $v(t)$ is zero-mean: $M(v) = 0$;
- The unknown signal $w(t)$ is zero-mean: $M(w) = 0$;
- The unknown signal and noise are uncorrelated: $C(v, w) = 0$.

We already know that $MS(v + w) = MS(v) + MS(w)$ since $C(v, w) = 0$. Then we can do this:

- Compute $\hat{MS}(y) = \frac{1}{T}\int_0^T y^2(t)dt$ from the data $\{y(t), 0 \le t \le T\}$. Then:
- $\hat{MS}(y) = \begin{cases} MS(w) + MS(v) & \text{if whales are mating;} \\ MS(v) & \text{if whales not mating} \end{cases}$;
- If $\hat{MS}(y) > \frac{1}{2}MS(w) + MS(v)$, decide the whales are mating;
- If $\hat{MS}(y) < \frac{1}{2}MS(w) + MS(v)$, decide the whales aren't mating;
- Threshold $\frac{1}{2}MS(w) + MS(v)$ assumes that it is equally likely that the whales are/aren't mating;
- You can adjust this threshold to bias this algorithm either way, if desired;
- Can also use $\text{rms}(y) = \sqrt{MS(y)}$ as the signal statistic to be thresholded;
- This actually makes more sense in terms of units, since $\text{rms}(ay) = |a|\text{rms}(y)$ (linear scaling).

IX. Application: Frequency and phase tracking

A. Introduction

Your FM radio has (and requires) the ability to track changes in frequency and phase, since both of these drift over time. Doppler radar also has the ability to track changes in frequency between the transmitted signal $s(t)$ and the received signal (one of the "other considerations" mentioned above). This is how doppler radar can measure wind speed, and whether a storm is approaching or receding.

How can we measure frequency and phase differences between two sinusoids without having to measure it on an oscilloscope by eye (which is a pain)? Once again, (normalized) correlation comes to our rescue.

B. Computing Phase Difference

Phase Problem: Given the sinusoidal signals $x(t) = A\cos(\omega t + \theta_x)$ and $y(t) = B\cos(\omega t + \theta_y)$, $-\infty < t < \infty$. We wish to compute the *phase difference* $|\theta_x - \theta_y|$ from data $x(t)$ and $y(t)$, without an oscilloscope.

Phase Solution: Compute the **normalized correlation=correlation coefficient**

$$C_N(x, y) = \cos(\theta_x - \theta_y) \to |\theta_x - \theta_y|. \tag{7}$$

C. Derivation of Phase Solution

Using the trig identity $\cos^2 x = \frac{1}{2}(1 + \cos(2x))$ and M(sinusoid)=0, we have

$$E(x) = \int_0^T A^2\cos^2(\omega t + \theta_x) = \int_0^T \frac{A^2}{2}(1 + \cos(2\omega t + 2\theta_x)) = T\frac{A^2}{2} = T\left(\frac{A}{\sqrt{2}}\right)^2 = T(\text{rms}(x))^2 \tag{8}$$

Using the trig identity $\cos(x)\cos(y) = \frac{1}{2}\cos(x+y) + \frac{1}{2}\cos(x-y)$ and M(sinusoid)=0, we have

$$C(x, y) = \int_0^T A\cos(\omega t + \theta_x)B\cos(\omega t + \theta_y)dt$$

$$= \frac{AB}{2}\int_0^T \cos(2\omega t + \theta_x + \theta_y)dt + \frac{AB}{2}\int_0^T \cos(\theta_x - \theta_y)dt = T\frac{AB}{2}\cos(\theta_x - \theta_y).$$

$E(y) = TB^2/2$ using a derivation analogous to that for $E(x) = TA^2/2$. Plugging in,

$$C_N(x,y) = \frac{C(x,y)}{\sqrt{E(x)E(y)}} = \frac{(TAB/2)\cos(\theta_x - \theta_y)}{\sqrt{(TA^2/2)(TB^2/2)}} = \cos(\theta_x - \theta_y). \qquad QED. \tag{9}$$

- $C_N(x,y) = 1$ if $\theta_x = \theta_y$ (in phase); and $C_N(x,y) = -1$ if $\theta_x = \theta_y \pm \pi$ (180^o out-of-phase).
- $C_N(x,y) = 0$ if $|\theta_x - \theta_y| = \frac{\pi}{2}$ (90^o out-of-phase); sin and cos functions are *orthogonal* over one period;
- QED="Quod Erat Demonstradum"=Latin phrase meaning "Thank God that's over."

D. Computing Frequency Difference

Frequency Problem: Given sinusoidal signals $x(t) = A\cos(2\pi f_x t)$ and $y(t) = B\cos(2\pi f_y t), -\infty < t < \infty$. We wish to compute the *frequency difference* $|f_x - f_y|$ from data $x(t)$ and $y(t)$, without oscilloscope.

Frequency Solution: Compute the **normalized correlation=correlation coefficient**

$$C_N(x,y) = \text{sinc}((f_x - f_y)T) \rightarrow |f_x - f_y| \quad \text{if} \quad |f_x - f_y| << |f_x + f_y| \tag{10}$$

Note that $\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$ has its peak at $x = 0$ and smaller peaks at half-integer values of x.

E. Derivation of Frequency Solution

We have shown that $E(x) = T\frac{A^2}{2}$ and $E(y) = T\frac{B^2}{2}$, but $C(x,y)$ is different. Using

$$\frac{1}{T}\int_{-T/2}^{T/2} \cos(\omega t)dt = \frac{\sin(\omega t)}{\omega T}\Big|_{-T/2}^{T/2} = \frac{\sin(\omega T/2)}{\omega T/2} = \text{sinc}(fT) \quad \text{since} \quad \omega = 2\pi f, \tag{11}$$

we now have

$$C(x,y) = AB\int_{-T/2}^{T/2} \cos(2\pi f_x t)\cos(2\pi f_y t)dt$$

$$= \frac{AB}{2}\int_{-T/2}^{T/2} \cos(2\pi(f_x + f_y)t)dt + \frac{AB}{2}\int_{-T/2}^{T/2} \cos(2\pi(f_x - f_y)t)dt = T\frac{AB}{2}[\text{sinc}((f_x + f_y)T) + \text{sinc}((f_x - f_y)T)] \tag{12}$$

We now assume (quite reasonably) that

$$(f_x + f_y) >> (f_x - f_y) \rightarrow \text{sinc}((f_x + f_y)T) << \text{sinc}((f_x - f_y)T) \tag{13}$$

in which case plugging in yields

$$C_N(x,y) = \frac{C(x,y)}{\sqrt{E(x)E(y)}} = \frac{(TAB/2)\text{sinc}(f_x - f_y)T}{\sqrt{(TA^2/2)(TB^2/2)}} = \text{sinc}(f_x - f_y)T \tag{14}$$

so that we can estimate small frequency differences $|f_x - f_y|$ from $C_N(x,y)$.