EECS 206 Lectures on Correlation, Fall 2001

D.L. Neuhoff

Remaining topics of course

- Signal correlation

  A time division concept that is the basis of

  I. many detection/classification/decision systems

  II. Fourier theory

- More discussion of discrete-time Fourier theory: Fourier series, Discrete Fourier transform, z transforms. Much of this discussion is based on the concept of signal correlation.
- Time frequency methods
- Introduction to continuous-time systems and Fourier theory
- Introduction to topics and courses that follow this course

  second course in signals and systems 306; signal processing 451,452; communications 353,455; control 460; probability and random processes (statistical signal processing) 401, ...

There are three kinds of correlation

A. Event correlation

B. Value correlation

C. Signal correlation

Though we will focus only on signal correlation, we mention the other types of correlation to decrease the likelihood of confusing signal correlation with the other types.

A. <u>Event correlation</u>:

In every day parlance --

Two events E_1 and E_2 are "correlated" if the frequency with which E_1 occurs in situations where E_2 is known to occur differs measurably from the frequency with which E_1 occurs when E_2 is known not to occur (or vice versa). Two events are said to "uncorrelated" if they are not correlated.

Example of correlated events:

E_1 = event that a patient smokes, E_2 = even that a patient has lung cancer

Example of uncorrelated events:

E_1 = event that a tossed coin shows heads, E_2 = event that another tossed coin shows heads

B. <u>Value correlation</u>:

In every day parlance --

Two measured values X and Y are "correlated" if knowledge of one value changes the likely values of the other.

Example if correlated values: X = height of parent, Y = height of first child

More concretely if $(X_1,Y_1), (X_2,Y_2), \ldots, (X_N,Y_N)$ is a sequence of pairs of X and Y values, the "correlation between X and Y" is

$$c(X,Y) = \frac{1}{N} \sum_{i=1}^{N} X_i \, Y_i$$

Statisticians say that X and Y are "uncorrelated" if $c(X,Y) = m(X)\, m(Y)$, where $m(X) = \frac{1}{N}\sum_{i=1}^{N} X_i$ = mean of X, $m(Y) = \frac{1}{N}\sum_{i=1}^{N} Y_i$ = mean of Y.

Or equivalently, if c(X-m(X),y-m(Y)) = 0 .

X and Y are said to be "correlated" if they are not uncorrelated.

Event and value correlation are things we think about in every day life. They are studied scientifically in courses on probability and statistics, e.g. EECS 401, Math 425. We will focus on "signal correlation", as defined below.

C. <u>Signal correlation</u>:

In signal processing, correlation is the following measure of the similarity of two signals:

Continuous-time: The "correlation between signals x(t) and y(t)" is

$$c(x(t),y(t)) = \int_{t_1}^{t_2} x(t)\ y(t)\ dt$$

Discrete-time: The "correlation between signals x[n] and y[n]" is

$$c(x[n],y[n]) = \sum_{n=n_1}^{n_2} x[n]\ y[n]$$

The limits of the integral/sums are whatever you want them to be. That is, you correlate over the time interval of interest for your particular problem.

Some authors multiply the integral and sum by $1/(t_2-t_1)$ and $1/(n_2-n_1)$, respectively. Some authors use other names for correlation, such as "inner product" or "dot product".

Large positive correlation means the signals are very similar.
Correlation zero or near zero means the signals are very different.
Large negative correlation means that each signal is very similar to the negative of the other.

If x and y have infinite duration, the correlation may be infinite or undefined. In such cases, we use a limiting definition such as

$$c(x(t),y(t)) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} x(t)\ y(t)\ dt$$

$$c(x[n],y[n]) = \lim_{N\to\infty} \frac{1}{2N}\sum_{n=-N}^{N} x[n]\ y[n]$$

Notice the similarity of the definition of signal correlation to that of value correlation.

From now on "correlation" means "signal correlation", unless stated otherwise.
As a shortcut, we will often write x for x(t) or x[n] and write correlation as c(x,y).

Usually, one subtracts the mean, i.e. the DC value, from x and y before taking the correlation. That is, we consider c(x-m(x),y-m(y)), where

the "mean value of x" is defined by

Continuous-time:

$$M(x) = \frac{1}{t_2-t_1}\int_{t_1}^{t_2} x(t)\ dt$$

Discrete-time:

$$M(x) = \frac{1}{n_2-n_1}\sum_{n_1}^{n_2} x[n]$$

From now on we assume that, unless otherwise stated, the signals we consider have the mean value subtracted, i.e. we assume $M(x) = 0$.

Two signals x and y are called "uncorrelated" or "orthogonal" if $c(x,y) = 0$,. "correlated" if $c(x,y) \neq 0$.

Uncorrelated signals are considered to be unrelated or independent.

"Energy" is the correlation of x with itself, except in the infinite duration case in which case "power" is the correlation of x with itself.

Continuous-time:

$$E(x) = \int_{t_1}^{t_2} x^2(t)\, dt$$

Discrete-time:

$$E(x) = \sum_{n_1}^{n_2} x^2[n]$$

Again the limits of the integral/sum are the time intervals of interest.

It is easy to show that uncorrelated signals have

$$E(x+y) = E(x) + E(y)$$

How large is "large correlation"? This is answered by the following theorem.

Schwarz Inequality Theorem:

$$|c(x,y)| \leq \sqrt{E(x)}\,\sqrt{E(y)}$$

with equality if and only if x and y differ only by a constant factor, i.e. if and only if $y = \alpha x$ for some constant α.

We won't give a proof of this. It's a basic theorem of linear algebra.

It follows from the C-S Inequality Theorem that x and y are highly correlated if

$$|c(x,y)| \cong \sqrt{E(x)}\,\sqrt{E(y)}\,,$$

Or equivalently if

$$\frac{c(x,y)}{\sqrt{E(x)}\,\sqrt{E(y)}} \cong \pm 1 \qquad \text{(normalized correlation)}$$

Moreover large correlation implies that one signal is, approximately, a scalar multiple of the other. Large negative correlation means that x is very similar to the negative of y.

Difference energy -- another similarity measure

Continuous-time: The "difference energy between signals $x(t)$ and $y(t)$" is

$$E(x-y) = \int_{t_1}^{t_2} (x(t)-y(t))^2\, dt$$

Discrete-time: The "difference energy between signals $x[n]$ and $y[n]$" is

$$E(x-y) = \sum_{n=n_1}^{n_2} (x[n]-y[n])^2$$

Which is the better measure of signal similarity, $c(x,y)$ or $E(x-y)$?

Observe their close relationship:

$$E(x\text{-}y) = \int_{t_1}^{t_2} (x(t)-y(t))^2\, dt = \int_{t_1}^{t_2} \left(x^2(t) - 2\,x(t)\,y(t) + y(t)^2\right) dt$$

$$= \int_{t_1}^{t_2} x^2(t) - 2\int_{t_1}^{t_2} x(t)\,y(t)\,dt + \int_{t_1}^{t_2} y(t)^2\, dt$$

$$= E(x) - 2\,c(x,y) + E(y)$$

So, for example, a large positive $c(x,y)$ implies a small $E(x\text{-}y)$. The same relation holds for discrete-time signals

Correlation $c(x,y)$ is preferred over difference energy $E(x\text{-}y)$ in situations where one signal, say x, is large and the other, y, is small. In this case $E(x\text{-}y) \cong E(x)$. And since difference energy depends very weakly on y, it is very sensitive to noise and computational roundoff errors. In contrast, $c(x,y)$ is greatly affected by y. For example, when y is much smaller than x, doubling y causes $c(x,y)$ to double but has little effect on $E(x\text{-}y)$. Thus correlation is less sensitive to noise and roundoff errors

Signals that are highly correlated, they have similar spectra, but the converse is false, as evidences by various examples. Thus high correlation means much more than just similar spectra.

Correlation between complex-valued signals:

Continuous-time: The "correlation between signals $x(t)$ and $y(t)$" is

$$c(x,y) = \int_{t_1}^{t_2} x(t)\; y^*(t)\; dt$$

Discrete-time: The "correlation between signals $x[n]$ and $y[n]$" is

$$c(x,y) = \sum_{n=n_1}^{n_2} x[n]\; y^*[n]$$

Why the complex conjugate? So relations like $E(x) = c(x,x)$ continue to hold.

Note that correlation between complex signals is a complex quantity that is not quite symmetric:

$$c(y,x) = \sum_{n=n_1}^{n_2} y[n]\; x^*[n] = \left(\sum_{n=n_1}^{n_2} x[n]\; y^*[n]\right)^* = c^*(x,y)$$

It's a bit unfortunate that for complex signals, correlation is not quite symmetric.
The Schwarz-Inequality Theorem holds for complex signals as well as real ones.

Examples:

1. x = constant signal and y is any signal with mean value equal to zero.

   $c(x,y) = 0.$

2. A noise waveform x and almost any other signal y:

   $c(x,y) \cong 0$

3. Two sinewaves of the same frequency and different phases.

   discrete-time:

   $$x[n] = A_1 \cos(\hat{\omega}_o n+\phi_1), \quad y[n] = A_2 \cos(\hat{\omega}_o n+\phi_2)$$

   $$\frac{c(x,y)}{\sqrt{E(x)}\sqrt{E(y)}} \cong \cos(\phi_2-\phi_1) \quad \text{if } n_2-n_1+1 >> N_o = \frac{2\pi}{\hat{\omega}_o}$$

   $$= \cos(\phi_2-\phi_1) \quad \text{if } n_2-n_1+1 \text{ is a multiple of } N_o/2.$$

   Notice how correlation depends on the phase but not the frequency.

4. Two sinewaves with differing frequencies.

   discrete-time:

   $$x[n] = A_1 \cos(\hat{\omega}_1 n+\phi_1), \quad y[n] = A_2 \cos(\hat{\omega}_2 n+\phi_2)$$

   $$\frac{c(x,y)}{\sqrt{E(x)}\sqrt{E(y)}} \cong 0 \quad \text{if } n_2-n_1+1 >> N_+, \quad \text{where } N_+ = \frac{2\pi}{\hat{\omega}_1+\hat{\omega}_2}$$

   $$= 0 \quad \text{if } n_2-n_1+1 \text{ is a multiple of } N_+/2 \text{ and } N_-/2,$$

   $$\text{where } N_- = \frac{2\pi}{\hat{\omega}_2-\hat{\omega}_1}$$

   Notice that when the frequencies are different, the correlation is approximately or exactly zero, independent of the phases.

5. Two complex exponentials with the same frequency and different phases.

   discrete-time:

   $$x[n] = A_1 e^{j(\hat{\omega}_o n+\phi_1)}, \quad y[n] = A_2 e^{j(\hat{\omega}_o n+\phi_2)},$$

   $$\frac{c(x,y)}{\sqrt{E(x)}\sqrt{E(y)}} = e^{j(\phi_1-\phi_2)} \quad \text{(which is never zero)}$$

6. Two complex exponentials with differing frequencies.

   discrete-time:

   $$x[n] = A_1 e^{j(\hat{\omega}_1 n+\phi_1)}, \quad y[n] = A_2 e^{j(\hat{\omega}_1 n+\phi_2)},$$

   $$\frac{c(x,y)}{\sqrt{E(x)}\sqrt{E(y)}} \cong 0 \quad \text{if } n_2-n_1+1 >> N_-, \quad \text{where } N_- = \frac{2\pi}{\hat{\omega}_2-\hat{\omega}_1}$$

   $$= 0 \quad \text{if } n_2-n_1+1 \text{ is a multiple of } N_-/2$$