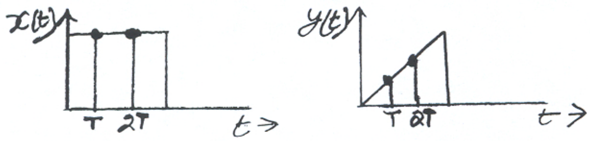
Comparison of Two Signals Using Cross-Correlation

[Note: Please read Sections of 2.5.1, 2.6 of Lathi's text for more details about this topic.]

Introduction

Cross-correlation is a very useful tool to quantify or measure the degree of similarity between two signals. Consider, for instance, the following two continuous-time signals, x(t) and y(t), as well as their discretized forms, x(nT) and y(nT), where T denotes the sampling interval.



A commonly used measure of similarity between x(t) and y(t) is called their cross-correlation, which is given by:

$$\text{Corr}[x(t),y(t)] = \int_{-\infty}^{\infty} x(t)y(t)dt \qquad (1a)$$

(which is sometimes also denoted by <x(t), y(t)>)

In discretized form, we can write equation (1) as:

$$\text{Corr}[x(n),y(n)] = \sum_{n=-\infty}^{\infty} x(n)y(n) \qquad (1b)$$

where x(n) and y(n) are commonly used alternative notations for x(nT) and y(nT), respectively.

In the example in Fig. 1, T is chosen such that we get just two samples of x(t) and y(t). Thus, we obtain

$$\text{Corr}[x(n),y(n)] = \sum_{n=1}^{2} x(n)y(n) \qquad (2)$$

Interpretation in terms of vectors

Suppose

$$X = \begin{bmatrix} x(1) \\ x(2) \end{bmatrix} \text{ and } Y = \begin{bmatrix} y(1) \\ y(2) \end{bmatrix}$$

denote two vectors in a rectangular coordinate space (with axes representing the values at T and 2T, respectively). Then we can think of equation (2) as the dot product of the vectors, X and Y, i.e.,

$$\text{Corr}[x(n),y(n)] = X \cdot Y \qquad (3)$$

Also, we know that

$$X \cdot Y = |X||Y|\cos(\theta), \tag{4a}$$

where θ denotes the angle between the vectors, X and Y and

$$|X| = \sqrt{[x(1)^2 + x(2)^2]} \text{ and } |Y| = \sqrt{[y(1)^2 + y(2)^2]} \tag{4b}$$

The quantity,

$$\cos(\theta) = (X \cdot Y)/(\,|X||Y|) \tag{5}$$

is taken to be a measure of similarity between the vectors, X and Y, because

$\theta = 0$ => $\cos(\theta) = 1$, i.e, X and Y are similar in the sense, $Y = \alpha X$, where $\alpha > 0$ is a constant;

$\theta = 180^\circ$ => $\cos(\theta) = -1$, i.e, X and Y are similar in the negative sense, $Y = -\alpha X$, $\alpha > 0$;

$\theta = 90^\circ$ => $\cos(\theta) = 0$, i.e, X and Y are orthogonal to each other.

Notice that the quantity in the numerator of equation (5), i.e., $X \cdot Y$, which is also denoted as the inner product, $X^T Y$, of two column vectors X and Y, equals the actual (raw) cross-correlation between X and Y. It is denoted by:

$$r_{xy} = X^T Y \tag{6a}$$

Also, the quantity in the denominator of equation (5), i.e., $|X||Y|$, just serves as a normalization factor. Thus, equation (5) can also be rewritten as:

$$\rho_{xy} = (X^T Y)/(\,|X||Y|) \tag{6}$$

which is known as the normalized cross-correlation, or correlation coefficient of X and Y. In view of what was said above,

$\rho_{xy} = 1$ => maximum positive correlation between X and Y (i.e., they are similar),

$\rho_{xy} = -1$ => maximum negative correlation between X and Y (similar in negative sense),

$\rho_{xy} = 0$ => X and Y are uncorrelated (or, dissimilar).

Generalization of the above idea

Instead of just two samples, we can collect N samples of x(t) and y(t) and let

$$X = \begin{bmatrix} x(1) \\ x(2) \\ \cdot \\ \cdot \\ \cdot \\ x(N) \end{bmatrix} \quad Y = \begin{bmatrix} y(1) \\ y(2) \\ \cdot \\ \cdot \\ \cdot \\ y(N) \end{bmatrix}$$

Then we can define

$$\rho_{xy} = (X^TY)/(\,|X||Y|) = \frac{\sum_{n=1}^{N} x(n)y(n)}{\sqrt{[\sum_{n=1}^{N} x^2(n) \sum_{n=1}^{N} y^2(n)]}} \qquad (7)$$

Notice that the above equation can also be written as:

$$\rho_{xy} = (X^TY)/(\,|X||Y|) = \frac{\sum_{n=1}^{N} x(n)y(n)}{\sqrt{(ExEy)}} \qquad (8)$$

where Ex and Ey denote the energies of x(n) and y(n), respectively.

Equation (8) is used to compute the normalized cross-correlation of two discretized signals, x(n) and y(n). Notice that if the sampling interval T is small, and $N \rightarrow \infty$, we get the cross-correlation of the corresponding analog signals, x(t) and y(t), i.e.,

$$\rho_{xy} = (X^TY)/(\,|X||Y|) = \frac{\int_{-\infty}^{\infty} x(t)y(t)}{\sqrt{[\int_{-\infty}^{\infty} x^2(t) \int_{-\infty}^{\infty} y^2(t)]}} \qquad (9)$$

Applications of cross-correlation

Cross-correlation is used to solve a variety of signal processing and communication problems, such as:

1. Target range estimation in a Radar
2. Detection of binary data in a Correlation Receiver
3. Detection of abnormalities in data
4. Identification and detection of known types of signals (or signatures) in presence of noise, etc.

Two simple examples of applications 1 and 2 are explored in Lab #7.