

Cross Correlation Explained With Real Signals

December 1, 2021

Table of Contents

Introduction

Correlation Function

Strong Positive Correlation

Strong Negative Correlation

Weak Positive Correlation

Weak Negative Correlation

Conclusion

Introduction

Cross correlation mathematically measures the similarity of signals. Consider an example where you have a set of data samples represented by $x[n]$ and $y[n]$. Cross correlation is used to measure on a sample by sample basis how similar $x[n]$ is to $y[n]$. Simple examples with plots will demonstrate different combinations of positive, negative, strong and weak correlations.

You might enjoy these other posts:

- Fourier Transform Explanation as a Cross-Correlation
- Cross Correlation: Explaining Time Lags
- Half Band Filter Design: Exceptional Filtering Efficiency!

Correlation Function

Correlation for DSP engineers, referred to as cross-correlation, is slightly different than the equation used by statisticians and mathematicians but they share the same underlying principles. The cross-correlation of sequences $x[n]$ and $y[n]$ is given by [gardner1988, p.212]

$$R_{xy}[\tau] = \sum_n x[n]y^*[n - \tau]. \quad (1)$$

The term τ is referred to as the “time-lag” and controls the relative time delay between the two sequences. The cross-correlation (1) at $\tau = 0$ calculates the similarity when there is no relative time delay,

$$R_{xy}[0] = \sum_n x[n]y^*[n]. \quad (2)$$

A special case of the cross-correlation is when $x[n] = y[n]$ is referred to as autocorrelation,

$$R_x[\tau] = \sum_n x[n]x^*[n - \tau]. \quad (3)$$

A large correlation value means the sequences $x[n]$ and $y[n]$ are similar while a large negative correlation means the sequences are similar but have opposite polarity. Small correlation values means the sequences have weak similarity while a correlation of 0 means the sequences have no similarity.

Strong Positive Correlation

Consider a sequence

$$x_0[n] = \begin{cases} 1, & n = 0 \\ -1, & n = 1 \\ 1, & n = 2 \\ -1, & n = 3 \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

The sequence $x_0[n]$ autocorrelates at $\tau = 0$ (2) according to

$$\begin{aligned} R_{x_0}[0] &= \sum_n x_0[n]x_0^*[n] \\ &= (1 \cdot 1) + (-1 \cdot -1) + (1 \cdot 1) + (-1 \cdot -1) \\ &= 4. \end{aligned} \quad (5)$$

Figure 1 shows that the two sequences are identical at $\tau = 0$ (no relative time delay) and therefore they should have the maximum correlation value, which in this case is 4. The larger the correlation, the larger the similarity.

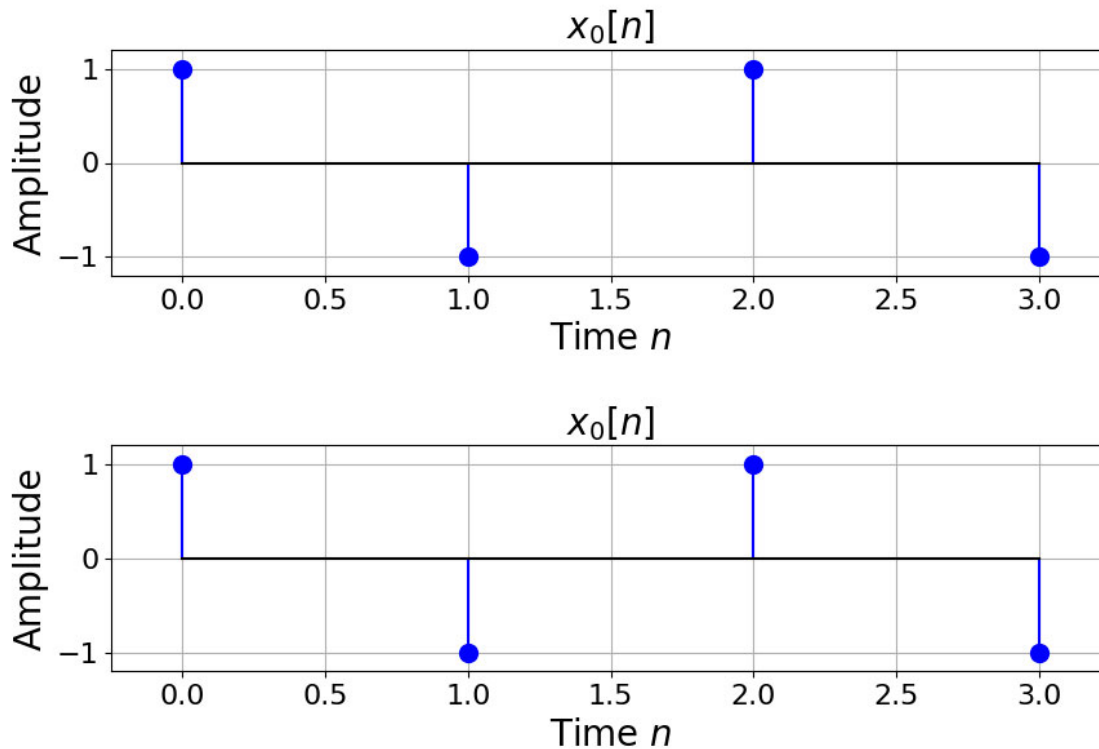


Figure 1: The two sequences for the autocorrelation of $x_0[n]$ and $x_0[n]$.

Strong Negative Correlation

Consider a sequence which is the negative of $x_0[n]$,

$$x_1[n] = -x_0[n] \quad (6)$$

such that

$$x_1[n] = \begin{cases} -1, & n = 0 \\ 1, & n = 1 \\ -1, & n = 2 \\ 1, & n = 3 \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

The cross-correlation between $x_0[n]$ and $x_1[n]$ at $\tau = 0$ from (2) is

$$\begin{aligned}
 R_{x_0 x_1}[0] &= \sum_n x_0[n] x_1^*[n] \\
 &= (1 \cdot -1) + (-1 \cdot 1) + (1 \cdot -1) + (-1 \cdot 1) \\
 &= -4.
 \end{aligned} \tag{8}$$

A negative correlation value means that the two sequences are similar at $\tau = 0$ (no relative time delay) but have opposite polarity. Figure 2 shows that the two sequences are the same with opposite polarity which is why the cross-correlation in (8) is the maximum negative value, -4.

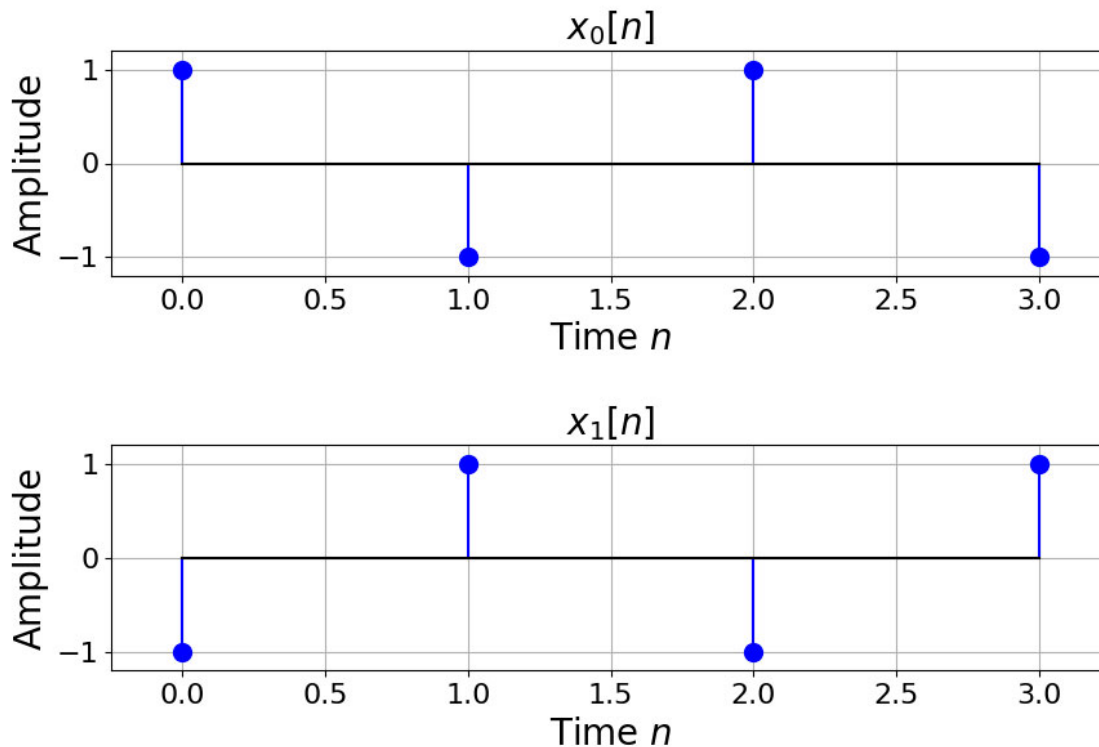


Figure 2: The two sequences for the cross correlation of $x_0[n]$ and $x_1[n]$.

Weak Positive Correlation

Consider a sequence $x_2[n]$ which has 1 data point in difference from $x_0[n]$ such that

$$x_2[n] = \begin{cases} -1, & n = 0 \\ -1, & n = 1 \\ 1, & n = 2 \\ -1, & n = 3 \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

The cross-correlation between $x_0[n]$ and $x_2[n]$ at $\tau = 0$ from (2) is

$$\begin{aligned}
 R_{x_0 x_2}[0] &= \sum_n x_0[n] x_2^*[n] \\
 &= (1 \cdot -1) + (-1 \cdot -1) + (1 \cdot 1) + (-1 \cdot -1) \\
 &= 2.
 \end{aligned} \tag{10}$$

A weak correlation value of 2 in (10) as compared to 4 in (5) means that the two sequences share some similarity at $\tau = 0$ but are not the exact same. Figure 3 shows that the two sequences are similar but with a single difference at $n=0$ which is why the cross-correlation in (10) is only 2.

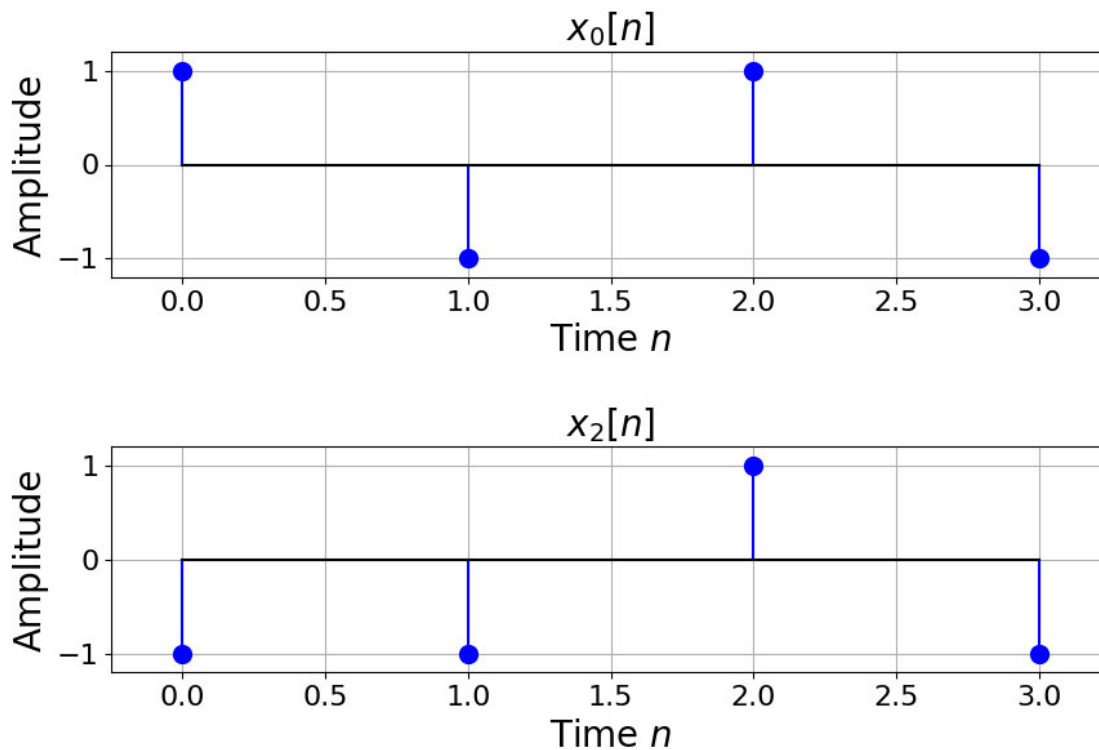


Figure 3: The two sequences for the cross correlation of $x_0[n]$ and $x_2[n]$.

Weak Negative Correlation

Consider a sequence $x_3[n]$ which has 1 data point the same as $x_0[n]$ but the other 3 are opposite polarity such that

$$x_3[n] = \begin{cases} -1, & n = 0 \\ 1, & n = 1 \\ -1, & n = 2 \\ -1, & n = 3 \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

The cross-correlation between $x_0[n]$ and $x_3[n]$ at $\tau = 0$ from (2) is

$$\begin{aligned}
 R_{x_0 x_3}[0] &= \sum_n x_0[n] x_3^*[n] \\
 &= (1 \cdot -1) + (-1 \cdot 1) + (1 \cdot -1) + (-1 \cdot -1) \\
 &= -2.
 \end{aligned} \tag{12}$$

A weak negative correlation value of -2 means that the two sequences share some similarity at $\tau = 0$ with opposite polarity but are not the exact same. Figure 4 shows that the two sequences are similar but with a single sample at $n=3$ in common, while the other three samples at $n=0, 1, 2$ are the opposite polarity which is why the cross-correlation in (12) is only -2.

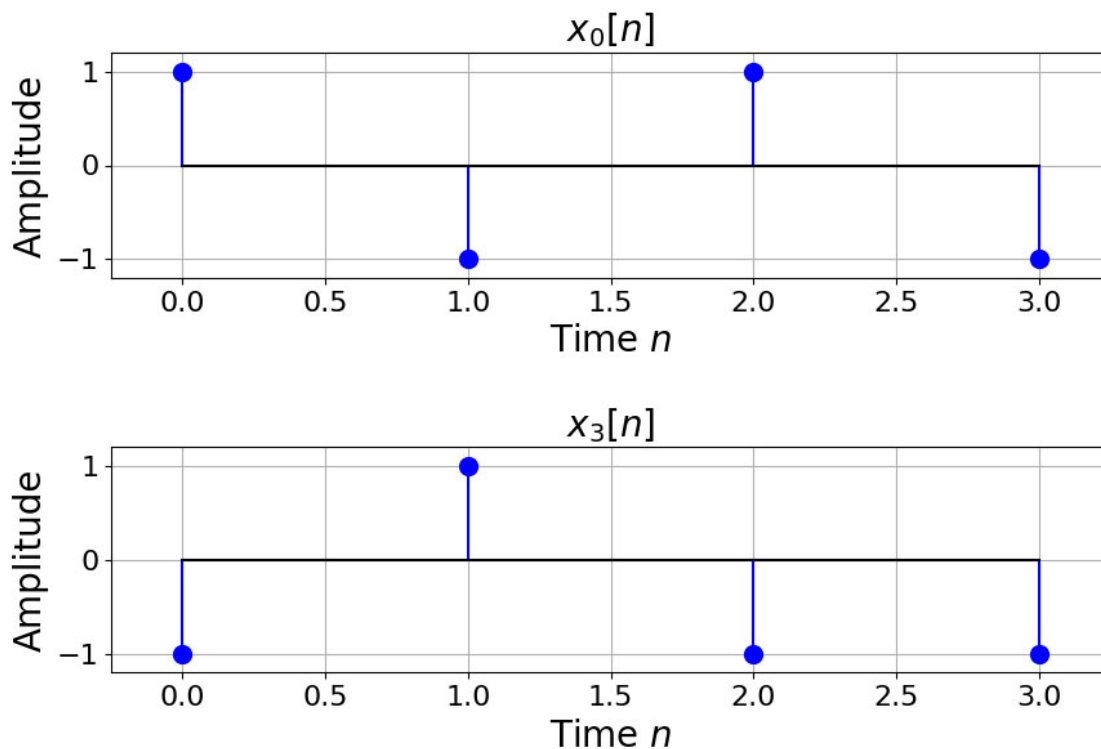


Figure 4: The two sequences for the cross correlation of $x_0[n]$ and $x_3[n]$.

Conclusion

Correlation is a way to mathematically measure similarity of two sequences. A large positive correlation means the two sequences are similar whereas a large negative correlation means the two sequences are similar but have opposite polarity. A small correlation value, positive or negative, means the two sequences share few similarities. A correlation value of zero means the two sequences do not share any similarities.

This post covered a subset of the cross-correlation, only for $\tau = 0$, in order to simplify the examples in this introduction. A future blog post will describe why the cross-correlation is computed over all time lags τ and how cross-correlation is applied in DSP algorithms.

You might enjoy these other posts:

- Fourier Transform Explanation as a Cross-Correlation
- Cross Correlation: Explaining Time Lags
- Half Band Filter Design: Exceptional Filtering Efficiency!

I hope you enjoyed this post explaining correlation, please check out others in the DSP Math series!