Digital Signal Processing for Music

Part 4: Signal Similarity - Correlation

Andrew Beck

Correlation Function

- Indicates (linear) dependencies between two signals
- Shifts the signals to find the dependency for each shift in time

Correlation Function

Compute similarity between two **stationary** signals x, y

$$r_{\mathrm{xy}}(\tau) = \mathcal{E}\{x(t)y(t+\tau)\}$$

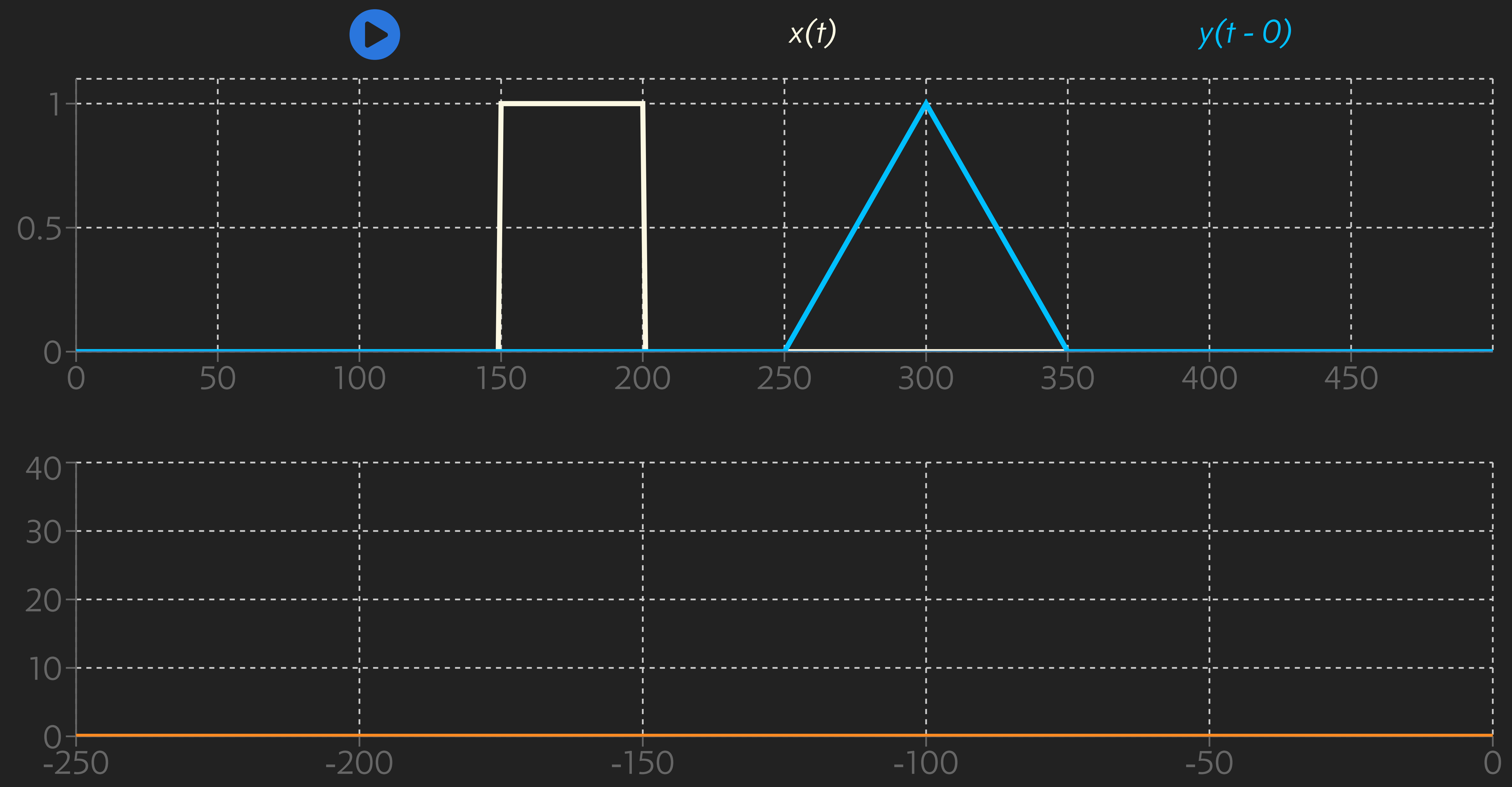
- Continuous:

$$r_{\mathrm{xy}}(\tau) = \int_{-\infty}^{\infty} x(t) \cdot y(t+\tau) dt$$

- Discrete:

$$r_{\mathrm{xy}}(\eta) = \sum_{i=-\infty}^{\infty} x(i) \cdot y(i+\eta)$$

$$r_{\mathrm{xy}}(\tau) = \int_{-\infty}^{\infty} x(t) \cdot y(t+\tau)dt$$



Use Cases

- Find (linear!) similarity between two signals (e.g., clean and noisy)
- Find time shift between two similar signals

Example: **Radar**

- Correlate sent signal with received signal
- Pick maximum location and convert to distance of object

Correlation Coefficient

$$r_{\mathrm{xy}}(\tau) = \frac{\mathcal{E}\{(X - \mu_X)(Y - \mu_Y)\}}{\sigma_X \sigma_Y}$$

Special case: **Pearson Correlation Coefficient** $r_{\mathrm{xy}}(0)$ after normalization.

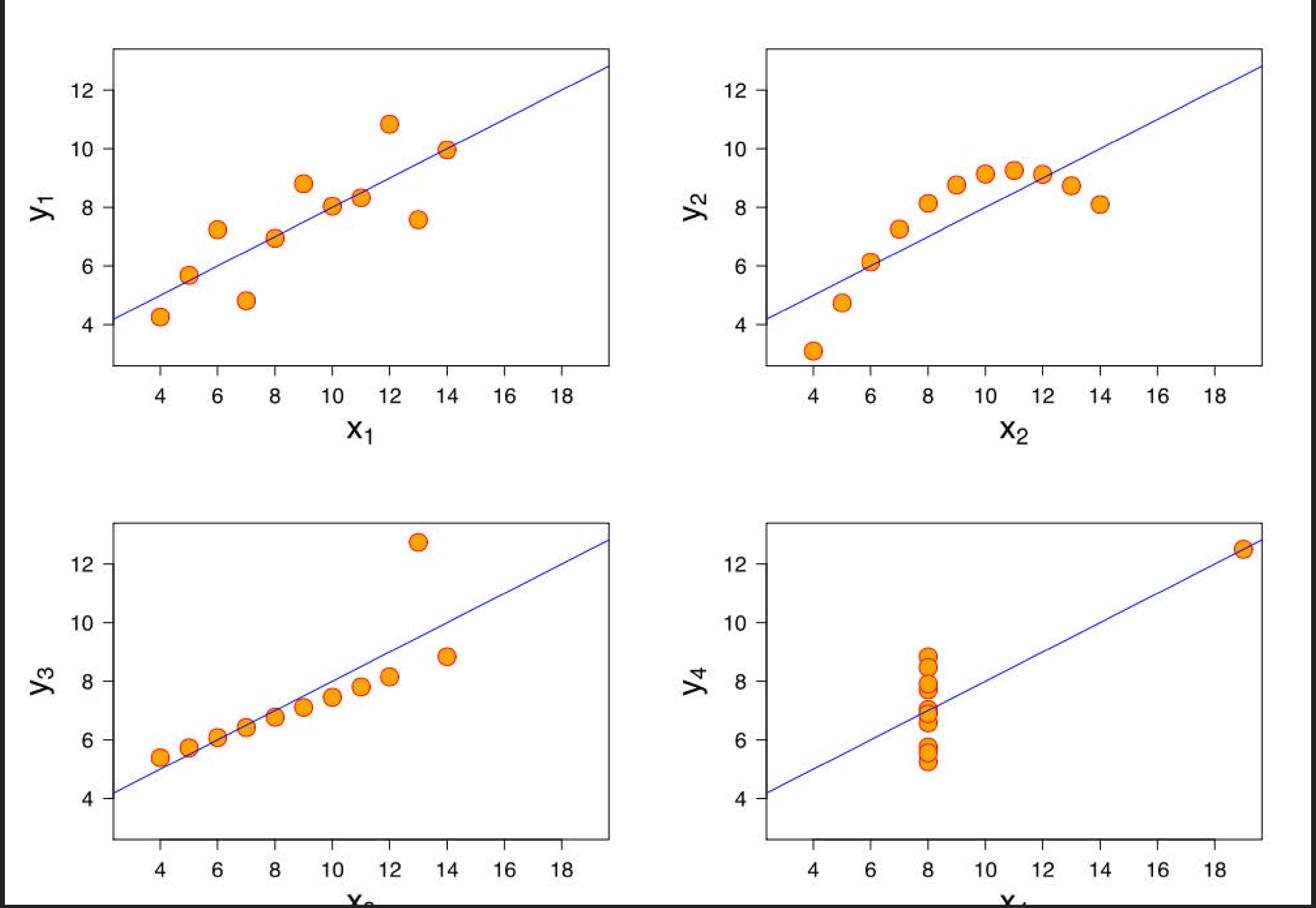
Possible reasons for normalization

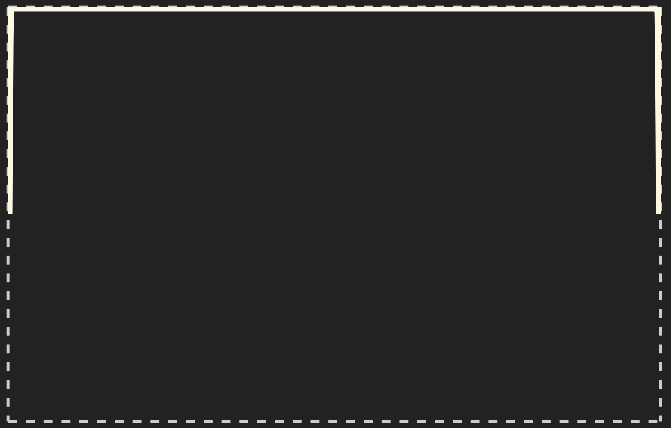
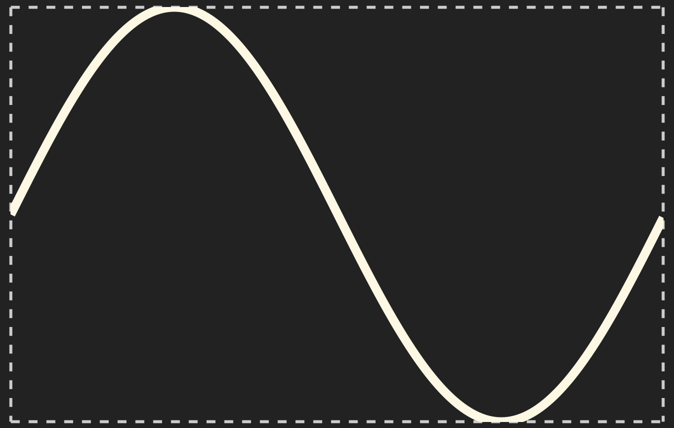
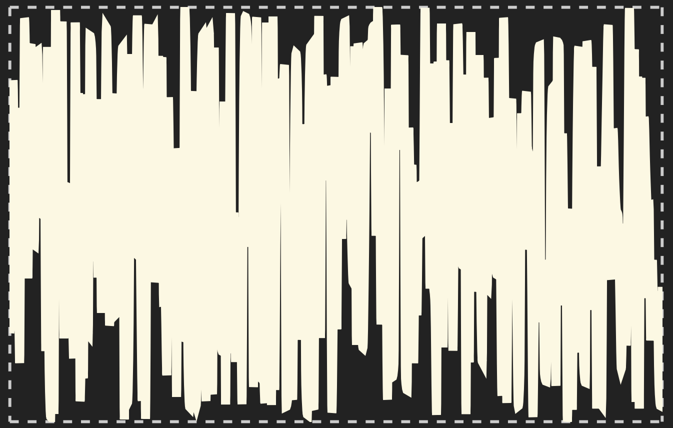
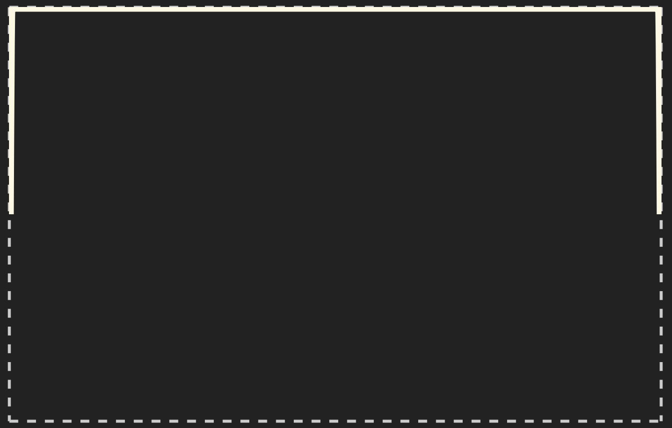
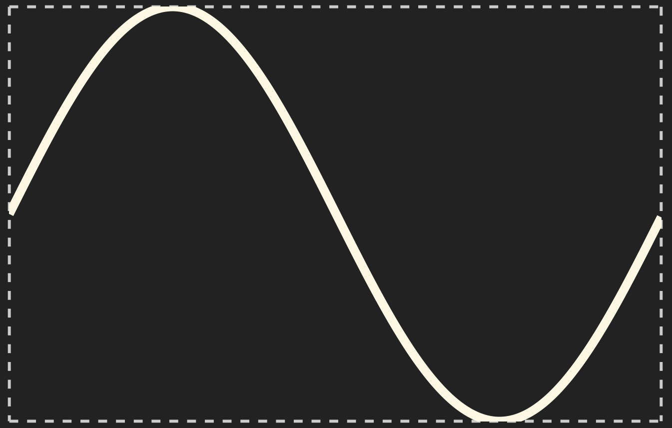
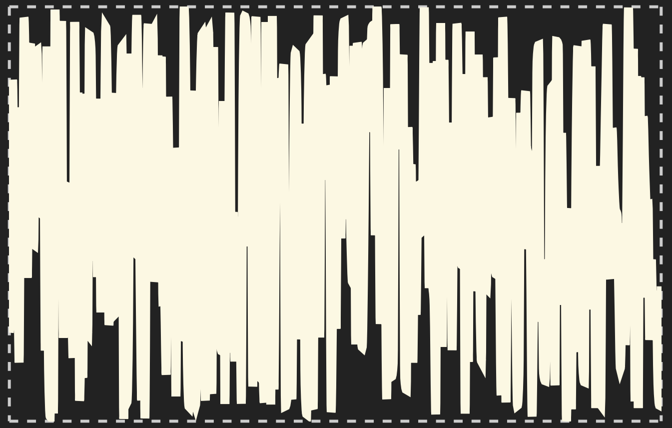
- Ensuring that function will always be between -1 and 1
- Shifting and scaling one signal will not change the coefficient

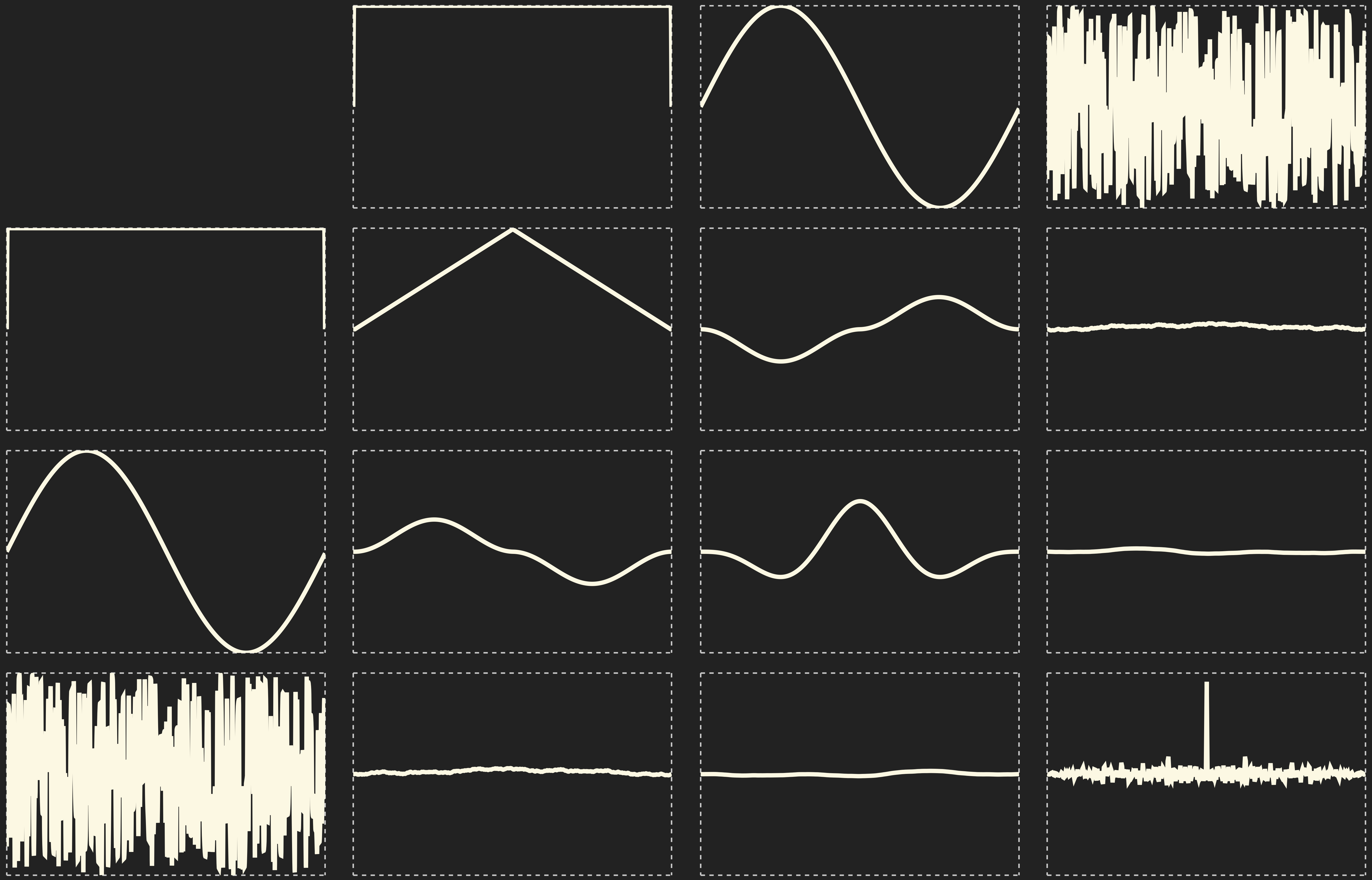
Problems with correlation as summary statistic

Anscombes quartet

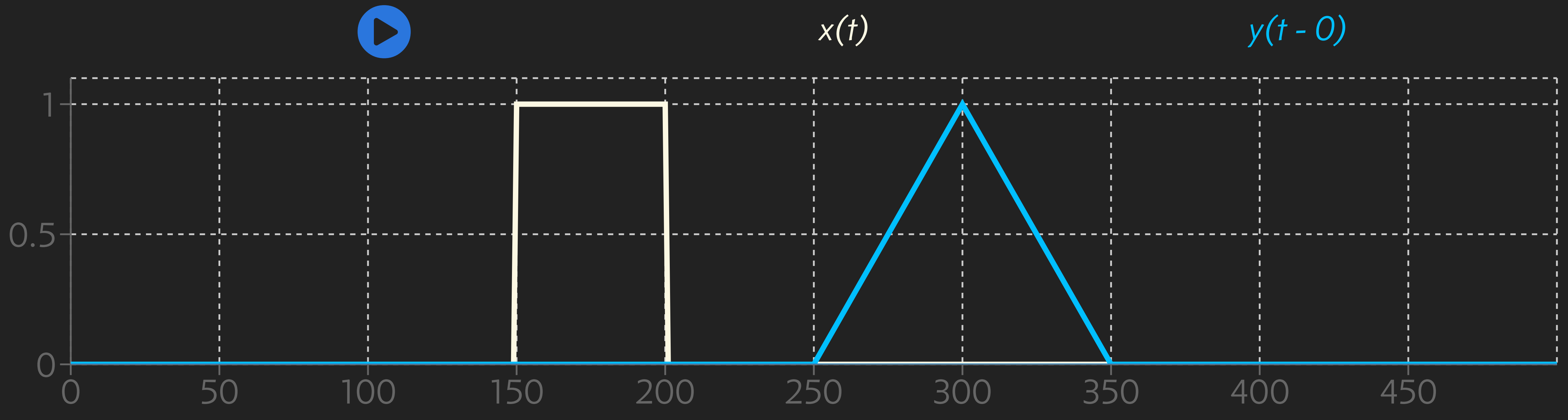
- Identical Mean: 7.5
- Identical Variance: 4.2
- Identical **Pearson correlation coefficient**: 0.816



			
	Rect Window x Rect Window	Rect Window x Windowed Sine	Rect Window x Noise
	Windowed Sine x Rect Window	Windowed Sine x Windowed Sine	Windowed Sine x Noise
	Noise x Rect Window	Noise x Windowed Sine	Noise x Noise



$$r_{\mathrm{xy}}(\tau) = \int_{-\infty}^{\infty} x(t) \cdot y(t+\tau)dt$$



Autocorrelation Function

$$r_{\mathrm{xx}}(\tau) = \mathcal{E}\{x(t)x(t+\tau)\}$$

Autocorrelation function properties

- **Power:** $r_{xx}(0) = \mathcal{E}\{X^2\}$
- **Symmetry:** $r_{xx}(\tau) = r_{xx}(-\tau)$

  (substitute $t = t' + \tau$)
- **Global Max:** $r_{xx}(\tau) \leq r_{xx}(0)$
- **Periodicity:**

  The ACF of a periodic signal is periodic (period length of input signal)

Summary

- Correlation Function is useful tool to

Determine the similarity between two signals (CCF)

Identify a shift/latency between two similar signals (CCF)

Identify periodicity vs. noisiness in a signal (ACF)

- Continues to be standard approach for all applications related to the above tasks
- Note: CCF or ACF do not display time information (lost in integration)