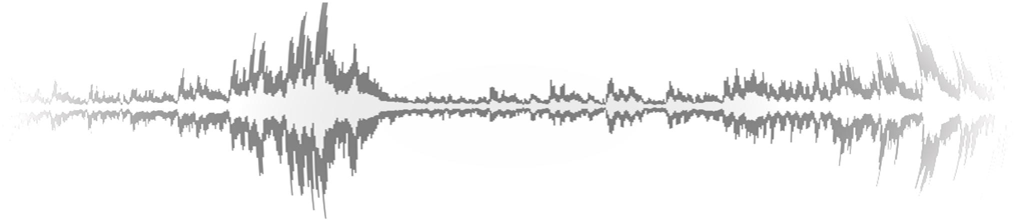
Digital Signal Processing for Music

Part 5: Signal Similarity — Correlation

alexander lerch



Georgia Tech Center for Music Technology
College of Design

signal similarity
introduction

- correlation function
  - indicates (linear) dependencies between two signals
  - shifts the signals to find the dependency for each shift in time

signal similarity
correlation function

Georgia Tech Center for Music Technology College of Design

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

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

- **continuous**:

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

- **discrete**:

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

signal similarity

correlation function

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

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

- **continuous**:

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

- **discrete**:

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

signal similarity
correlation function: animation

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

signal similarity
correlation function: use cases

Georgia Tech Center for Music Technology
College of Design

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

- example: **radar**
  - correlate sent signal with received signal
  - pick maximum location and convert to distance of object

signal similarity

correlation function: use cases

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

- example: **radar**
  - correlate sent signal with received signal
  - pick maximum location and convert to distance of object

signal similarity

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

What are possible reasons for normalization

signal similarity
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

What are possible reasons for normalization

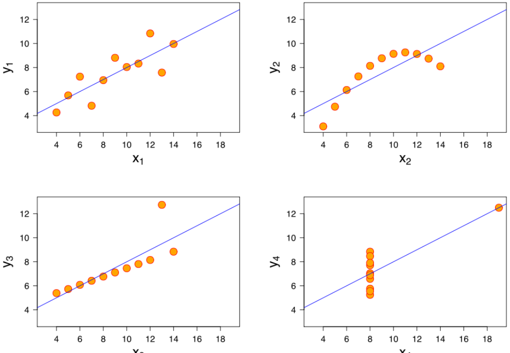
- ensuring that function will always be between -1 and 1
- shifting and scaling one signal will not change the coefficient

signal similarity

correlation function: problems as summary statistic

Anscombe's quartet:

- identical mean: 7.5
- identical variance: 4.2
- identical **Pearson correlation coefficient**: 0.816



correlation function

examples 1/3

Describe the (cross) correlation function between the following signals

correlation function

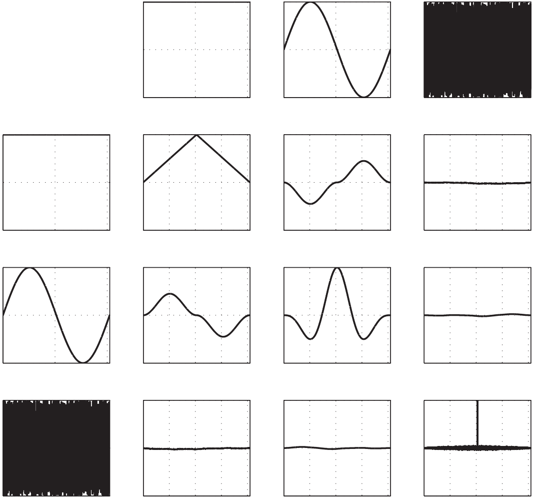
examples 1/3

Describe the (cross) correlation function between the following signals

- rectangular window vs.
- windowed sine vs.
- noise

correlation function

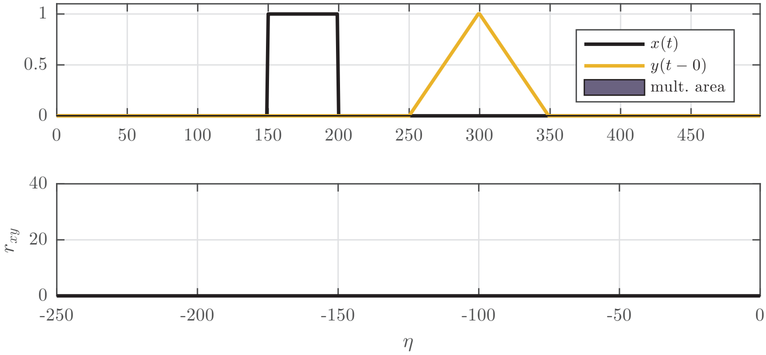
examples 2/3



correlation function

animation

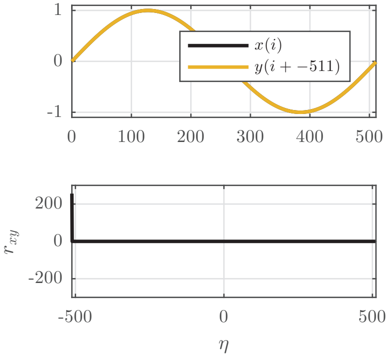
$$r_{\mathrm{xy}}(\eta) = \sum_{i=-\infty}^{\infty} x(i) \cdot y(i+\eta)$$

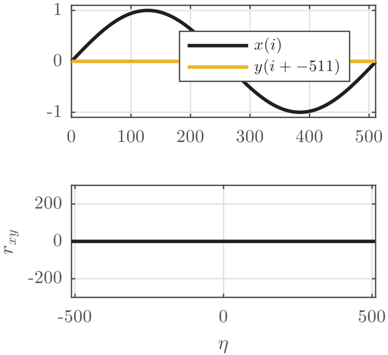


matlab source: [matlab/animateCorrelation.m]

correlation function

blocked correlation: animation





matlab source: matlab/animateBlockedCorrelation.m

correlation function

auto correlation 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)

correlation function

auto correlation function

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

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)

correlation function

auto correlation function

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

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)

correlation function

auto correlation 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)

correlation function
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)

correlation function

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)

correlation function

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)