Introduction to Signal Processing

Davide Bacciu

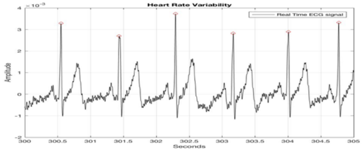
Dipartimento di Informatica
Università di Pisa
bacciu@di.unipi.it

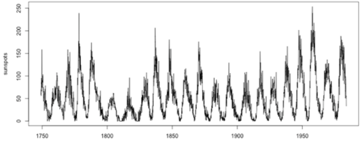
Intelligent Systems for Pattern Recognition

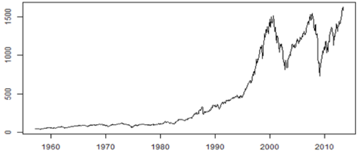
Signals = Time series

A sequence of measurements in time

- Medicine
- Financial
- Meteorology
- Geology
- Biometrics
- Robotics
- IoT
- Biometrics
- ...







Formalization

A time series $\mathbf{x}$ is a sequence of measurements in time t

$$\mathbf{x} = x_0, x_1, \ldots, x_t, \ldots, x_N$$

where x_t (or $x(t)$) is the measurement at time t.

- Observations can be observable at irregular time intervals
- Time series analysis assumes weakly stationary (or second-order stationary) data
  - $\mathbb{E}[x_t] = \mu$ for all t
  - $Cov(x_{t+\tau}, x_t) = \gamma_\tau$ for all t (γ does only depend on lag τ)

Goals

- Description - Summary statistics, graphs
- Analysis - Identify and describe dependencies in data
- Prediction - Forecast the next values given information up to time t
- Control - Adjust the parameters of the generative process to make the time series fit a target

The goal of this lecture is providing knowledge on some basic techniques that can be useful as

- Baseline
- Preprocessing
- Building blocks

Key Methods

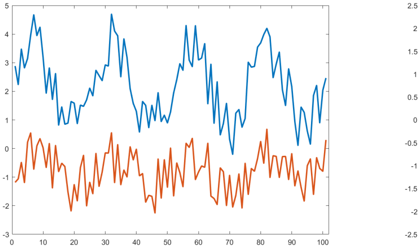
- Time domain analysis - Assesses how a signal changes over time
  - Correlation and Convolution
  - Autoregressive models
- Spectral domain analysis - Assesses the distribution of the signal over a range of frequencies
  - Fourier Analysis
  - Wavelets

Mean and Autocovariance

Some interesting estimators for time series statistics are

Sample mean

$$\hat{\mu} = \frac{1}{N}\sum_{t=1}^{N} x_t$$



(Sample) Autocovariance for lag $-N \leq \tau \leq N$

$$\hat{\gamma}_{\mathbf{x}}(\tau) = \frac{1}{N}\sum_{t=1}^{N-|\tau|}(x_{t+|\tau|} - \hat{\mu})(x_t - \hat{\mu})$$

Autocorrelation

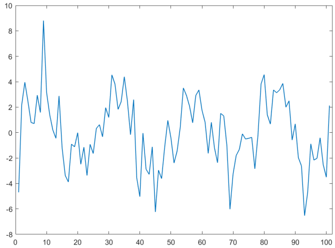
Autocovariance serves to compute autocorrelation, i.e. the correlation of a signal with itself

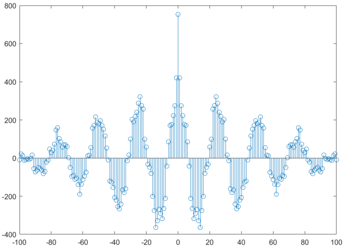
$$\hat{\rho}_{\mathbf{x}}(\tau) = \frac{\hat{\gamma}_{\mathbf{x}}(\tau)}{\hat{\gamma}_{\mathbf{x}}(0)}$$

Autocorrelation analysis can reveal repeating patterns such as the presence of a periodic signal hidden by noise

Autocorrelation Plot

A revealing view on timeseries statistics





What do you see in this time series?

Autocorrelogram reveals a sine wave

Cross-Correlation (Discrete)

A measure of similarity of $\mathbf{x}^1$ and $\mathbf{x}^2$ as a function of a time lag τ

$$\phi_{\mathbf{x}^1\mathbf{x}^2}(\tau) = \sum_{t=\max\{0,\tau\}}^{\min\{(T^1-1+\tau),(T^2-1)\}} x^1(t-\tau)\cdot x^2(t)$$

- $\tau \in [-(T^1-1),\dots,0,\dots,(T^1-1)]$
- The maximum $\phi_{\mathbf{x}^1\mathbf{x}^2}(\tau)$ w.r.t. τ identifies the displacement of $\mathbf{x}^1$ vs $\mathbf{x}^2$

Cross-Correlation (Discrete)

Normalized cross-correlation returns an amplitude independent value

$$\overline{\phi}_{\mathbf{x}^1\mathbf{x}^2}(\tau) = \frac{\phi_{\mathbf{x}^1\mathbf{x}^2}(\tau)}{\sqrt{\sum_{t=0}^{T^1-1}(x^1(t))^2 \sum_{t=0}^{T^2-1}(x^2(t))^2}} \in [-1, +1]$$

- $\overline{\phi}_{\mathbf{x}^1\mathbf{x}^2}(\tau) = +1 \Rightarrow$ The two time-series have the exact same shape if aligned at time τ
- $\overline{\phi}_{\mathbf{x}^1\mathbf{x}^2}(\tau) = -1 \Rightarrow$ The two time-series have the exact same shape but opposite sign if aligned at time τ
- $\overline{\phi}_{\mathbf{x}^1\mathbf{x}^2}(\tau) = 0 \Rightarrow$ Completely uncorrelated signals

Cross-Correlation - Something already seen...

What is this?

$$(f * g)[n] = \sum_{t=-M}^{M} f(n-t)g(t)$$

- Discrete convolution on finite support $[-M, +M]$
- Similar to cross-correlation but one of the signals is reversed (i.e. $-t$ in place of t)
- Convolution can be seen as a smoothing operator (commutative!)

A View of Time Domain Operators

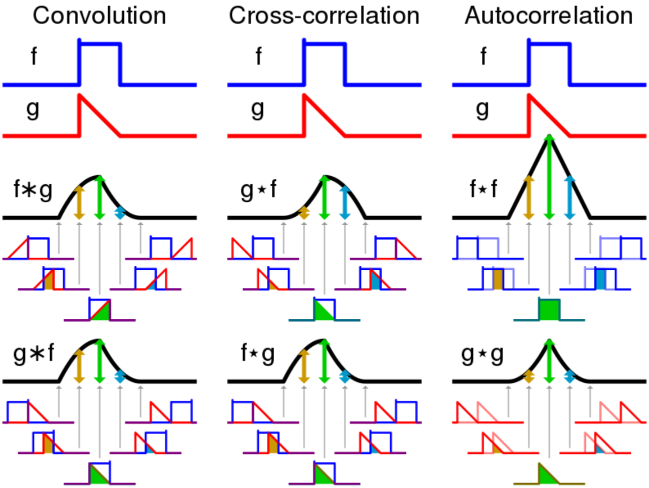


Image Credit: Wikipedia

Autoregressive Process

A timeseries Autoregressive process (AR) of order K is the linear system

$$x_t = \sum_{k=1}^{K} \alpha_k x_{t-k} + \epsilon_t$$

- Autoregressive $\Rightarrow$ x_t regresses on itself
- α_k $\Rightarrow$ linear coefficients s.t. $|\alpha| < 1$
- ϵ_t $\Rightarrow$ sequence of i.i.d. values with mean 0 and fixed variance

ARMA

Autoregressive with Moving Average process (ARMA)

$$x_t = \sum_{k=1}^{K} \alpha_k x_{t-k} + \sum_{q=1}^{Q} \beta_q \epsilon_{t-q} + \epsilon_t$$

- $\epsilon_t \Rightarrow$ Random white noise (again)
- The current timeseries value is the result of a regression on its past values plus a term that depends on a combination of stochastically uncorrelated information
- Denotes new information or shocks at time t

Estimating Autoregressive Models

- Need to estimate
  - The values of the linear coefficients α_t (and β_t)
  - The order of the autoregressor K (and Q)
- Estimation of the α is performed with the Levinson-Durbin recursion, e.g. in Matlab `a = levinson(x,K)`
- The order is often estimated with a Bayesian model selection criterion, e.g. BIC, AIC, etc.

The set of autoregressive parameters $\alpha_1^i, \dots, \alpha_K^i$ fitted to a specific timeseries $\mathbf{x}^i$ is used to confront it with other timeseries

Comparing Timeseries by AR

- Timeseries clustering

$$d(\mathbf{x}^1, \mathbf{x}^2) = \|\alpha^1 - \alpha^2\|_M^2$$

- Novelty/anomaly detection

$$\text{Test} \;\; Err(x_t, \hat{x}_t) < \xi$$

  where $\hat{x}_t$ is the AR predicted value
- Encode time series as a set of α^i vectors and feed them to a flat ML model

Spectral Analysis

Analysing time series in the frequency domain

Key Idea

Decompose a time series into a linear combination of sinusoids (and cosines) with random and uncorrelated coefficients

- Time domain - Regression on past values of the time series
- Frequency domain - Regression on sinusoids

Use the framework of Fourier Analysis

Fourier Transform

- Discrete Fourier transform (DFT)
- Transforms a time series from the time domain to the frequency domain
- Can be easily inverted (back to the time domain)
- Useful to handle periodicity in the time series
  - Seasonal trends
  - Cyclic processes

Representing Functions

We (should) know that, given an orthonormal system $\{\mathbf{e}_1; \mathbf{e}_2, \dots\}$ for E, we can represent any function $f \in E$ by a linear combination of the basis

$$\sum_{k=1}^{\infty} < f, \mathbf{e}_k > \mathbf{e}_k.$$

Given the orthonormal system

$$\left\{ \frac{1}{\sqrt{2}}, \sin(x), \cos(x), \sin(2x), \cos(2x), \dots \right\}$$

the linear combination above becomes the Fourier Series

$$a_0/2 + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \sin(kx)]$$

with a_k, b_k being coefficients resulting from integrating $f(x)$ with the sin and cos functions

Representing Functions in Complex Space

Using $\cos(kx) - i\sin(kx) = e^{-ikx}$ with $i = \sqrt{-1}$ we can rewrite the Fourier series as

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

on the orthonormal system

$$\left\{1, e^{ix}, e^{-ix}, e^{i2x}, e^{-i2x}, \dots\right\}$$

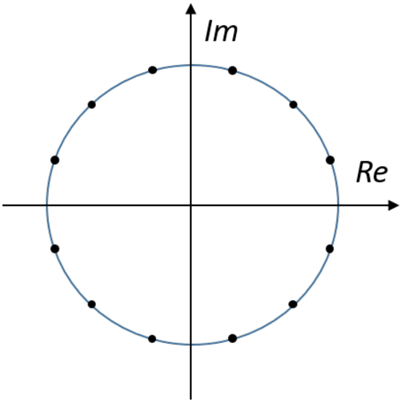
and c_k integrates $f(x)$ with e^{-ikx}.

Representing Discrete Time series

1. Consider a discrete time series $\mathbf{x} = x_0, x_1, \ldots, x_{N-1}$ of length N and $x_n \in \mathbb{R}$
2. Using the exponential formulation the orthonormal system is made of $\{\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_{N-1}\}$ vectors $\mathbf{e}_k \in \mathbb{C}^N$
3. The n-th component of the k-th vector is

$$[\mathbf{e}_k]_n = e^{-\frac{2\pi ink}{N}}$$

Graphically



A basis $\mathbf{e}_k$ at frequency k has N elements sampled from the roots of the unitary circle in imaginary-real space

Discreet Fourier Transform

Given a time series $\mathbf{x} = x_0, x_1, \ldots, x_{N-1}$ its Discrete Fourier Transform (DFT) is the sequence (in frequency domain)

$$X_k = \sum_{n=1}^{N-1} x_n e^{\frac{-2\pi ink}{N}}$$

The DFT has an inverse transform

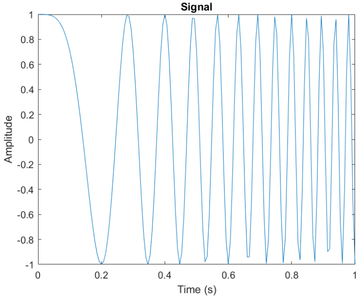
$$x_n = \frac{1}{N} \sum_{k=1}^{N-1} X_k e^{\frac{2\pi ink}{N}}$$

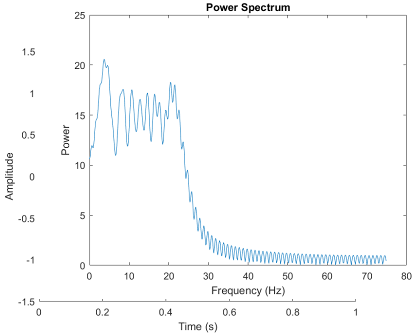
to go back to the time domain.

Amplitude

A measure of relevance/strenght of a target frequency k

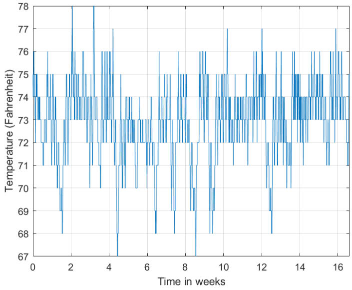
$$A_k = Re^2(X_k) + Im^2(X_k)$$

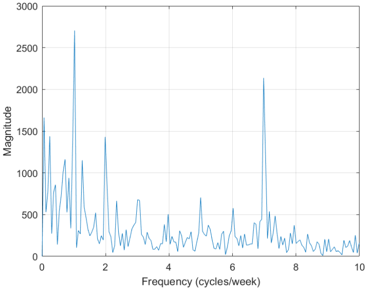




DFT in Action

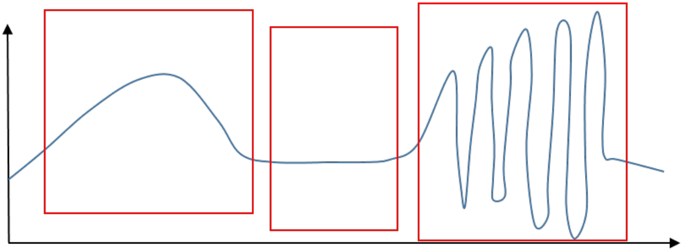
- Use the DFT elements $X_1, \ldots, X_K$ as representation of the signal to train predictor/classifier
- Representation in spectral domain can reveal patterns that are not clear in time domain



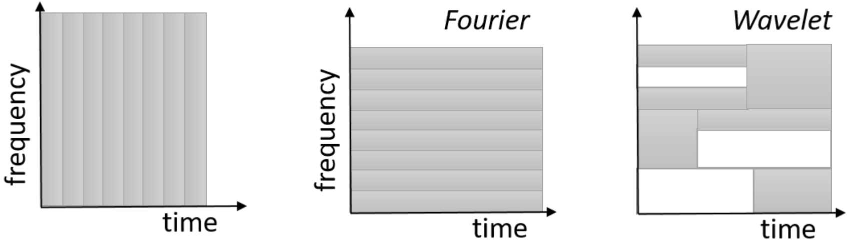


Limitations of DFT

Sometimes we might need localized frequencies rather than global frequency analysis



Graphical Intuition



Split signal in frequency bands only if they exist in specific time-intervals

Wavelets

Split the signal using an orthonormal basis generated by translation and dilation of a mother wavelet

$$f(\mathbf{x}) = \sum_{t \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \Psi_{t,k}(\mathbf{x})$$

Terms k and t regulate scaling and shifting of the wavelet

$$\Psi_{t,k}(\mathbf{x}) = 2^{k/2}\Psi(2^k\mathbf{x} - t)$$

with respect to the mother $\Psi(\cdot)$. E.g.

$k < 1$ Compresses the signal

$k > 1$ Dilates the signal

Many different possible choices for the mother wavelet function: Haar, Daubechies, Gabor,...

Take Home Messages

- Old-school pattern recognition on timeseries is about learning coefficients that describe properties of the time series
    - Autoregressive coefficients (time domain)
    - Fourier coefficient (frequency domain)
- Often linear methods
    - Autocorrelation reveals similitude of a signal with delayed versions of itself
    - Cross-correlation provides hints on time series similarity and how to align them
- Fourier analysis allows to identify recurring patterns and to identify key frequencies in the signal

Next Lecture

- Introduction to image processing
- Representing images and visual content
- Identify informative parts of an image
- Spectral analysis in 2D