Temporal Time Series Analysis

Joaquín Rapela

Gatsby Computational Neuroscience Unit
University College London

January 12, 2025

Contents

Contents

Background

- Last Spring 2023 I helped in the discussion sessions of this course.
- Suggested to Klara Olofsdotter (SWC PhD program coordinator) and Sonja Hofer (SWC PhD program faculty coordinator) to ask SWC PhD students to take this course. They liked the idea.
- I volunteered to lead discussions and do grading with Gatsby Unit PhD students and postdoctoral scholars.

A few motivations to run this course

1. Gain more teaching experience.
2. Provide SWC PhD students with essential neural data-analysis tools.
3. Contribute to better interactions between the SWC and the Gatsby Unit.

Course structure

Week 01	Jan 11	The t-test and randomisation tests	Joaquin Rapela	tutorial
Week 02	Jan 18	Power spectra	Joaquin Rapela Yousef Mohammadi Joe Zimminski	tutorial
Week 03	Jan 25	Spectrograms and coherence	Joaquin Rapela Yousef Mohammadi Joe Zimminski	tutorial
Week 04	Feb 01	Circular statistics	Joaquin Rapela	tutorial
Week 05	Feb 08	Singular value decomposition	Will Dorrell	tutorial
Week 06	Feb 15 Feb 16	Linear regression	Lior Fox	lecture tutorial
Week 07	Feb 22 Feb 23	Linear dynamical systems	Aniruddh Galgali Joaquin Rapela	lecture tutorial
Week 08	Feb 29	no class (CoSyNe)		
Week 09	Mar 07 Mar 08	Artificial neural networks	Erin Grant	lecture tutorial
Week 10	Mar 14 Mar 15	Experimental control with Bonsai	Goncalo Lopes Joaquin Rapela	lecture tutorial
Week 11	Mar 21 Mar 22	Reinforcement learning		lecture tutorial
Week 12 Week 15	Mar 28 Apr 25	Project development		
Week 16	May 02	Project presentations		

Teaching assistants: Kira Dusterwald, Sihao (Daniel) Liu

Graded worksheets

Every Thursday we will assign you a worksheet that is due on the second Monday after the assignment.

Contents

Contents

What is time series analysis?

- Time series analysis characterizes **data that is correlated in time**.
- These correlations serverly **restrict the applicability of conventional technniques** assuming data samples that are independent and identically distributed.
- These correlations allow to **forecast** future values of a time series based on present and past values.

Relevance of time series analysis

economics: daily stock market quotations, monthly unemployment figures.

social scientists: birthrates, school enrollment.

epidemology: number of influenza cases observed over some time period.

medicine: blood pressure measurements traced over time.

Temporal vs spectral time series analysis

termporal time series analysis focuses on the analysis of lagged relationship (e.g., how does what happened today affect what will happen tomorrow?).

spectral time series analysis centers on the analysis of rhythms (e.g., can we observe rhythmic activity in local field potentials recorded from human brains?)

An example time series

In the examples below we will use spontaneous EEG (i.e., no task) recorded from a human subject (Liu et al., 2024).

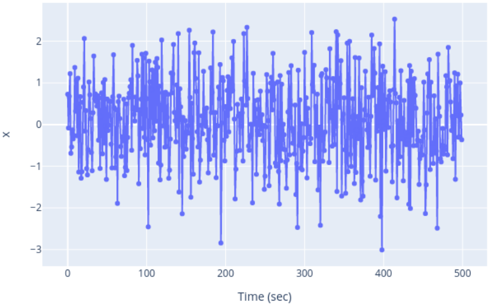
Contents

Generation of time series: white noise

The first step to generate time series is to generate **white noise**, $\{w_t\}$ (i.e., independent Gaussian random variables with zero mean and fixed variance, example).

$$E\{w_t\} = 0$$

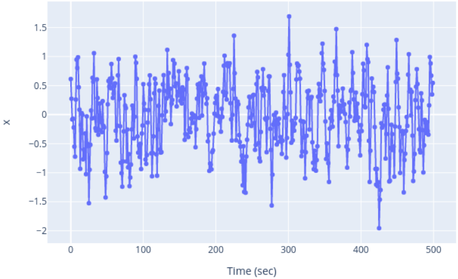
$$\mathrm{Cov}\{w_t, w_s\} = \begin{cases} \sigma_w^2 & s = t \\ 0 & s \neq t \end{cases}$$



Generation of time series: moving average model

In a white noise stochastic process x, for any pairs of time points, t_1 and t_2, the random variables x_{t_1} and x_{t_2} are uncorrelated. The **moving average model** adds serial correlation to white noise (example).

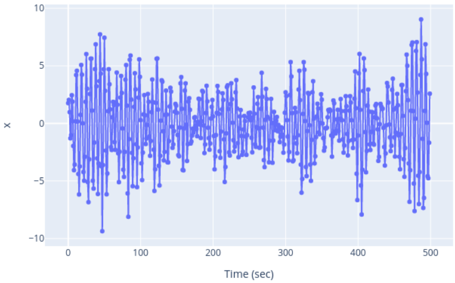
$$\nu_t = \frac{1}{3}(w_{t-1} + w_t + w_{t+1}) \tag{1}$$



Generation of time series: autoregressive model

Many neural time series, like local field potential recordings, exhibit oscillations of the type of sine waves. The **autoregressive model** generates oscillations (example). Lecture on linear dynamical systems.

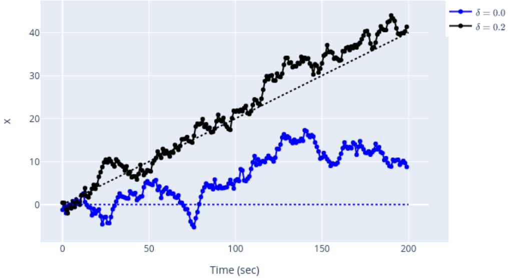
$$x_t = x_{t-1} - 0.9x_{t-2} + w_t$$



Generation of time series: random walk with drift

The **random walk with noise** model is used to characterize trends in time series (example).

$$x_t = \delta + x_{t-1} + w_t \tag{2}$$



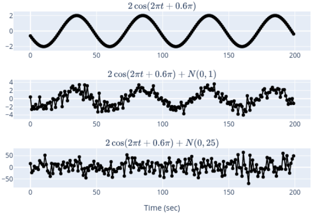
Generation of time series: signal plus noise

Many realistic models of time series assume an underlying signal with a periodic variation contaminated by adding a random noise (example).

$$x_t = 2\cos(2\pi\frac{t}{50} + 2\pi\frac{15}{50}) + w_t$$

$$A\cos(2\pi\omega t + \phi)$$

where $A = 2, \omega = 1/50, \phi = 2\pi 15/50$. Lecture on spectral analysis of time series.



Contents

Mean function

Definition 1 (Mean function)

The mean function, μ_t, is defined as $\mu_t = E\{x_t\}$.

Example (Mean function of a moving average model)

Calculate the mean function of the moving average model in Eq. 5.

$$E\{\nu_t\} = \frac{1}{3}(E\{w_{t-1}\} + E\{w_t\} + E\{w_{t+1}\}) = \frac{1}{3}(0 + 0 + 0) = 0$$

Mean function

Example (Mean function of the autoregressive model of order 1)

Calculate the mean function of the autoregressive model of order 1, AR(1), in Eq. 3.

$$x_t = \phi x_{t-1} + w_t \tag{3}$$

An AR(1) model (Eq. 3) can be represented as a moving average of ininite order MA(∞). See MA(∞) representation of AR(1) random process in Appendix. Then

$$x_t = \sum_{i=0}^{\infty} \phi^i w_{t-i}$$

$$E\{x_t\} = \sum_{i=0}^{\infty} \phi^i E\{w_{t-i}\} = \sum_{i=0}^{\infty} \phi^i 0 = 0$$

Mean function

Example (Mean function of the random walk with drift model)

Calculate the mean function of the random noise with drift model, in Eq. 2.

The random noise with drift model in Eq. 2 can be represented as

$$x_t = t\delta + \sum_{i=0}^{\infty} w_{t-i}$$

$$E\{x_t\} = \delta t + \sum_{i=0}^{\infty} E\{w_{t-i}\} = \delta t + \sum_{i=0}^{\infty} 0 = \delta t$$

See the figure of samples of the random noise with drift random process.

Autocovariance function

Definition 2 (Autocovariance function)

The autocovariance function is defined as
$\gamma(s,t) = cov(x_s, x_t) = E\{(x_s - \mu_s)(x_t - \mu_t)\}$.

Note

For $s = t$ the autocovariance reduces to the variance, because
$\gamma(t,t) = E\{(x_t - \mu_t)^2\} = var(x_t)$.

Definition 3 (Autocorrelation function)

The autocorrelation function is defined as $\rho(s,t) = \frac{\gamma(t,s)}{\sqrt{\gamma(t,t)\gamma(s,s)}}$.

Autocovariance function

Example (Autocovariance function of moving average)

Calculate the autocovariance function of the moving average model in Eq. 5.

$$\gamma_\nu(s,t) = cov(\nu_s,\nu_t) = cov(\frac{1}{3}(w_{s-1}+w_s+w_{s+1}),\frac{1}{3}(w_{t-1}+w_t+w_{t+1}))$$

If s=t:

$$\begin{aligned}
\gamma_\nu(t,t) &= cov(\nu_t,\nu_t) = cov(\frac{1}{3}(w_{t-1}+w_t+w_{t+1}),\frac{1}{3}(w_{t-1}+w_t+w_{t+1})) \\
&= \frac{1}{9}\left(cov(w_{t-1},w_{t-1})+cov(w_t,w_t)+cov(w_{t+1},w_{t+1})\right) \\
&= \frac{1}{9}\left(\sigma_w^2+\sigma_w^2+\sigma_2^2\right) = \frac{3}{9}\sigma_w^2
\end{aligned}$$

Autocovariance function

Example (Autocovariance function of moving average)

If s=t+1:

$$
\begin{aligned}
\gamma_\nu(t+1,t) &= cov(\nu_{t+1},\nu_t)\\
&= cov(\frac{1}{3}(w_t+w_{t+1}+w_{t+2}),\frac{1}{3}(w_{t-1}+w_t+w_{t+1}))\\
&= \frac{1}{9}\left(cov(w_t,w_t)+cov(w_{t+1},w_{t+1})\right)\\
&= \frac{1}{9}\left(\sigma_w^2+\sigma_w^2\right)=\frac{2}{9}\sigma_w^2
\end{aligned}
$$

Autocovariance function

Example (Autocovariance function of moving average)

If s=t+2:

$$\begin{aligned}\gamma_\nu(t+2,t) &= cov(\nu_{t+2},\nu_t)\\ &= cov(\frac{1}{3}(w_{t+1}++w_{t+2}+w_{t+3}),\frac{1}{3}(w_{t-1}+w_t+w_{t+1}))\\ &= \frac{1}{9}\left(cov(w_{t+1},w_{t+1})\right)\\ &= \frac{1}{9}\sigma_w^2\end{aligned}$$

Autocovariance function

Example (Autocovariance function of moving average)

$$\gamma_\nu(s,t) = \begin{cases} \frac{3}{9}\sigma_w^2 & \text{if} \quad s = t, \\ \frac{2}{9}\sigma_w^2 & \text{if} \quad |s-t| = 1, \\ \frac{1}{9}\sigma_w^2 & \text{if} \quad |s-t| = 2, \\ 0 & \text{if} \quad |s-t| > 2. \end{cases}$$

Autocovariance function

Example (Autocovariance function of AR(1))

Calculate the autocovariance function of the autoregressive model of order 1 in Eq. 4.

An AR(1) model (Eq. 3) can be represented as a moving average of ininite order MA(∞). See MA(∞) representation of AR(1) random process in Appendix.

$$x_t = \sum_{i=0}^{\infty} \phi^i w_{t-i}$$

Autocovariance function

Example (Autocovariance function of AR(1))

$$\begin{aligned}
\gamma(t-h,t) &= E\{(x_{t-h}-\mu_{t-h})(x_t-\mu_t)\} = E\{x_{t-h}x_t\} = E\left\{\left(\sum_{i=0}^{\infty}\phi^i w_{t-h-i}\right)\left(\sum_{j=0}^{\infty}\phi^j w_{t-j}\right)\right\}\\
&= E\left\{\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\phi^i\phi^j w_{t-h-i}w_{t-j}\right\} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\phi^i\phi^j E\{w_{t-h-i}w_{t-j}\}\\
&= \sum_{i=0}^{\infty}\phi^i\phi^{i+h}E\{w_{t-h-i}^2\} = \phi^h\sigma_w^2\sum_{i=0}^{\infty}\phi^{2i} = \phi^h\sigma_w^2\frac{1}{1-\phi^2}, \quad \text{if } |\phi|<1
\end{aligned}$$

Contents

Strictly stationary time series

Definition 4 (Strict stationarity)

A **strictly stationary time series** is one for which the probabilistic behavior of every collection of values

$$\{x_{t_1}, \ldots, x_{t_n}\}$$

is identical to that of any shifted set

$$\{x_{t_1+h}, \ldots, x_{t_n+h}\}$$

That is

$$P(x_{t_1} < c_1, \ldots, x_{t_k} < c_k) = P(x_{t_1+h} < c_1, \ldots, x_{t_k+h} < c_k)$$

for all $k = 1, 2, \ldots$, all time points $t_1, t_2, \ldots, t_k$, all numbers $c_1, c_2, \ldots, c_k$, and all time shifts $h = 0, \pm 1, \pm 2, \ldots$.

Weakly stationary time series

Definition 5 (Weak or wide-sense stationarity)

A **weakly** or **wide-sense stationary time series** is a finite-variance process such that:

1. the mean function, μ_t, is contant and does not depend on time t, and
2. the autocovariance function, $\gamma(s, t)$, depends on s and t only through their difference $|s - t|$.

Contents

Sample mean, autocovariance and autocorrelation

Definition 6 (Sample mean)

Let $x_1, \ldots, x_n$ be observations from a time series. The **sample mean** of $x_1, \ldots, x_n$ is

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Definition 7 (Sample autocovariance)

The **sample autocovariance function** is

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{i=1}^{n-|h|} (x_{i+|h|} - \bar{x})(x_i - \bar{x}), \quad -n < h < n$$

Sample mean, autocovariance and autocorrelation

Definition 6 (Sample autocorrelation)

The **sample autocorrelation function** is

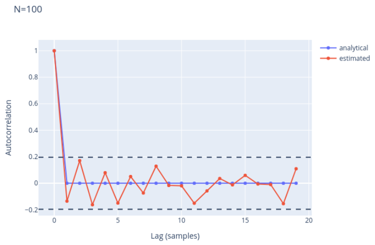
$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$$

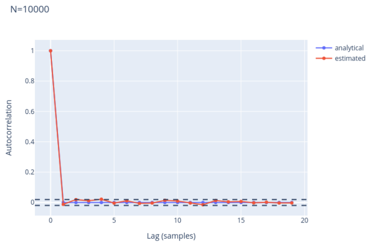
Theorem 7 (Distribution of sample autocorrelation for white noise)

For white noise, and a sample of size n, the sample autocorrelations, $\hat{\gamma}(h), h > 0$, are approximately independent and identically distributed $N(0, 1/\sqrt{n})$, for large n (Brockwell and Davis, 1991). Hence 95% of the sample autocorrelations should fall between the bound $\pm 1.96/\sqrt{n}$

Analytical and estimated autocorrelation function for white noise

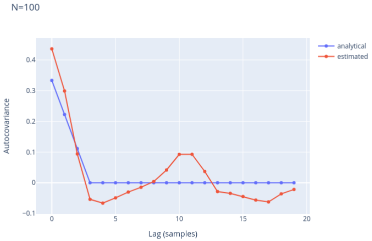
Simulate a white noise time series with N=100 and N=100,000 samples. For each N, plot the analytical and estimated autocorrelation function. Include the 95% confidence interval of the autocorrelation function.
Solution.

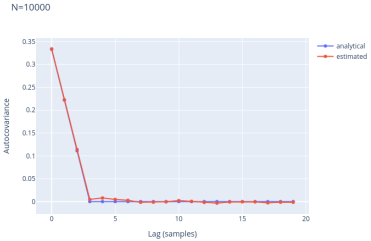




Analytical and estimated autocovariance function for MA

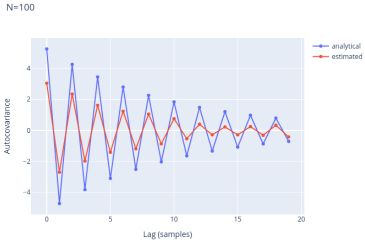
Simulate the previous moving average time series with N=100 and N=100,000 samples. For each N, plot the analytical and estimated autocovariance function.

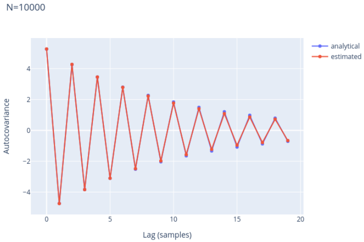




Analytical and estimated autocovariance function for AR(1)

Simulate an AR(1) time series with N=100 and N=100,000 samples, $\phi = 0.9$ and $\sigma_w = 1.0$. For each N, plot the analytical and estimated autocovariance function.





Contents

Forecasting

Forecasting is the problem of predicting the value of $X_{n+h}, h > 0$, of a stationary time series, in term of the previous m values$\{X_n, \ldots, X_{n-(m-1)}\}$. The mean of such predictor is

$$\text{mean}(\text{pred}(X_{m+h}|X_n, \ldots, X_{n-(m-1)})) = \mu + \mathbf{a_m}^\intercal \begin{bmatrix} x_n - \mu \\ \ldots \\ x_{n-(m-1)} - \mu \end{bmatrix}$$

and its variance is

$$\text{var}(\text{pred}(X_{n+h}|X_n, \ldots, X_{n-(m-1)})) = \gamma(0) - \mathbf{a_m}^\intercal \gamma_m(h)$$

Forecasting

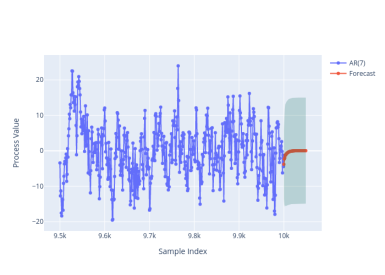
with

$$
\begin{aligned}
\Gamma_m \mathbf{a_m} &= \gamma_m(h) \\
\Gamma_m = [\gamma(i-j)]_{i,j=1}^m &= \begin{bmatrix}
\gamma(0) & \gamma(1) & \gamma(2) & \gamma(3) & \dots & \gamma(m-1) \\
\gamma(1) & \gamma(0) & \gamma(1) & \gamma(2) & \dots & \gamma(m-2) \\
\gamma(2) & \gamma(1) & \gamma(0) & \gamma(1) & \dots & \gamma(m-3) \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\gamma(m-1) & \gamma(m-2) & \gamma(m-3) & \gamma(m-4) & \dots & \gamma(0)
\end{bmatrix} \\
\mathbf{a_m} &= [a_1, \dots, a_m]^\intercal \\
\gamma_m(h) &= [\gamma(h), \gamma(h+1), \dots, \gamma(h+m-1)]^\intercal
\end{aligned}
$$

AR(p) forecasting example

Example (Forecasting with an AR(p) model)

Simulate N=10,000 samples from an AR(7) stochastic process with $\phi = [5.0/6, -1.0/6, 0.5/6, -0.25/6, 0.5/6, -0.1/6, 0.05/6]$ and $\sigma_w = 5.0$. Use the last 500 samples to forecast 50 samples (i.e., $n = 10,000, m = 500, h = 1, \ldots, 50$).



Summary

Contents

MA(∞) representation of AR(1) random process

Claim 1

Let $|\phi| < 1$, *then*

$$x_t = \phi x_{t-1} + w_t \quad \textit{if and only if} \tag{4}$$

$$x_t = \sum_{i=0}^{\infty} \phi^i w_{t-i} \tag{5}$$

MA(∞) representation of AR(1) random process

Proof.

We first show that x_t, as defined in Eq. 5, satisfies Eq. 4.

$$\phi x_{t-1} = \phi \sum_{i=0}^{\infty} \phi^i w_{t-1-i} = \phi \sum_{j=1}^{\infty} \phi^{j-1} w_{t-j} = \sum_{j=1}^{\infty} \phi^j w_{t-j}$$

$$\phi x_{t-1} + w_t = \sum_{j=0}^{\infty} \phi^j w_{t-j} = x_j$$

MA(∞) representation of AR(1) random process

Proof.

We now show that Eq. 5 is the unique solution to Eq. 3. Suppose y_t is stationary and satisfies Eq. 3, then

$$\begin{aligned}
y_t &= \phi y_{t-1} + w_t \\
&= \phi(\phi y_{t-2} + w_{t-1}) + w_t = \phi^2 y_{t-2} + \phi w_{t-1} + w_t \\
&= \phi^{t+1} y_{t-(t+1)} + \phi^t w_t + \ldots + \phi w_{t-1} + w_t \\
&= \phi^{k+1} y_{t-(k+1)} + \sum_{i=0}^{k} \phi^k w_{t-i}
\end{aligned}$$

$$E\left\{\left(y_t - \sum_{i=0}^{k} \phi^i w_{t-i}\right)^2\right\} = \phi^{2k+2} E\{y^2_{t-(k+1)}\} = \phi^{2k+2}\sigma^2$$

$$E\left\{\left(y_t - \sum_{i=0}^{\infty} \phi^i w_{t-i}\right)^2\right\} = \lim_{k\to\infty} \phi^{2k+2}\sigma^2 = 0$$

MA(∞) representation of AR(1) random process

Proof.

Thus y_t equals $\sum_{i=0}^{\infty} \phi^i w_{t-i}$ in the mean-squared sense. $\square$

Brockwell, P. J. and Davis, R. A. (1991). *Time series: Theory and methods*. Springer-Verlag, 2nd edition.

Liu, Q., Jia, S., Tu, N., Zhao, T., Lyu, Q., Liu, Y., Song, X., Wang, S., Zhang, W., Xiong, F., et al. (2024). Open access eeg dataset of repeated measurements from a single subject for microstate analysis. *Scientific Data*, 11(1):379.