2.1: Stationary Processes: Basic Properties

Taylor

University of Virginia

Motivation

Why are autocovariances and autocorrelations so important for stationary time series data? What do we use them for?

Before answering that question in this chapter, this section discusses properties of the autocovariance function, stationarity, and the options for forecasting/predicting.

Properties of Autocovariance

All autocovariances are "tied together" in special ways.

1. $\gamma(0) \geq 0$
2. for any h, $|\gamma(h)| \leq \gamma(0)$
3. for any h, $\gamma(h) = \gamma(-h)$
4. for any n and any $a_1, \ldots, a_n$, $\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \gamma(|i-j|) \geq 0$

Theorem

A real valued function κ defined on the integers is an autocovariance function of a stationary time series if and only if it is even (3) and non-negative definite (4).

Show (1) and (2) are extra.

Example 1

Prove that $\kappa(h) = \cos(\omega h)$ is a valid autocovariance function, where ω is any real number.

Example 1

Prove that $\kappa(h) = \cos(\omega h)$ is a valid autocovariance function, where ω is any real number.

This is the autocovariance function for $X_t = A\cos(\omega t) + B\sin(\omega t)$ provided A and B are uncorrelated each with mean 0 and variance 1. It's also discussed in your homework.

Example 2

Prove that

$$\kappa(h) = \begin{cases} 1, h = 0 \\ \rho, h = \pm 1 \\ 0, \text{else} \end{cases}$$

is an autocovariance function for $|\rho| \leq \frac{1}{2}$.

Example 2

Prove that

$$\kappa(h) = \begin{cases} 1, h = 0 \\ \rho, h = \pm 1 \\ 0, \text{else} \end{cases}$$

is an autocovariance function for $|\rho| \leq \frac{1}{2}$.

It is the autocovariance function of $X_t = Z_t + \theta Z_{t-1}$ if $Z_t \sim$ IID Noise with variance σ^2. Notice we can only solve for θ and σ^2 when $|\rho| \leq \frac{1}{2}$.

Example 3

Prove that

$$\kappa(h) = \begin{cases} 1, h = 0 \\ \rho, h = \pm 1 \\ 0, \text{else} \end{cases}$$

can't be an autocovariance function for $|\rho| > \frac{1}{2}$.

Example 3

Prove that

$$\kappa(h) = \begin{cases} 1, h = 0 \\ \rho, h = \pm 1 \\ 0, \text{else} \end{cases}$$

can't be an autocovariance function for $|\rho| > \frac{1}{2}$.

Take $\mathbf{a} = (1, -1, 1, \ldots)'$. Then

$$\begin{aligned} \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \kappa(|i-j|) &= \sum_{i=1}^{n} (\pm 1)^2 \kappa(0) + \sum_{i=1}^{n-1} (1)(-1)\rho + \sum_{i=2}^{n} (1)(-1)\rho \\ &= n - \rho(n-1) - \rho(n-1) \\ &= n(1-2\rho) + 2\rho \end{aligned}$$

Which could be negative if $\rho > 1/2$. For the case when $\rho < 1/2$, take $\mathbf{a} = (1, 1, 1, \ldots)'$

Stationarity

definition

For $h, n \in \mathbb{Z}^+$, a **strictly stationary** time series is one where

$$(X_1, \ldots, X_n)' \stackrel{d}{=} (X_{1+h}, \ldots, X_{n+h})'$$

Necessary Properties

For each stationary time series:

- each X_t is identically distributed
- $(X_t, X_{t+h})' \stackrel{d}{=} (X_1, X_{1+h})'$ for all t, h
- If $E[X_t^2] < \infty$ for all t, then the series is weakly stationary, too

Other Stuff

- Weak stationarity does not imply strict stationarity
- IID random samples are strictly stationary (non-time series data)

Stationarity: Example

Let $Z_1, Z_2, \ldots \sim$ IID Noise. Prove that

$$X_t = g(Z_t, Z_{t-1}, \ldots, Z_{t-q})$$

is (strictly) stationary, for any function g (not necessarily linear)...

Stationarity: Example

Let $Z_1, Z_2, \ldots \sim$ IID Noise. Prove that

$$X_t = g(Z_t, Z_{t-1}, \ldots, Z_{t-q})$$

is (strictly) stationary, for any function g (not necessarily linear)...

Answer:

$$\begin{aligned}(X_{q+1}, \ldots, X_{q+1+n})' &= (g(Z_{q+1}, \ldots, Z_1), \ldots, g(Z_{q+1+n}, \ldots, Z_{n+1}))' \\ &= (g(Z_{q+1+h}, \ldots, Z_{1+h}), \ldots, g(Z_{q+1+n+h}, \ldots, Z_{n+1+h})) \\ &= (X_{q+1+h}, \ldots, X_{q+1+n+h})'\end{aligned}$$

Stationarity: Example

If $Z_1, Z_2, \ldots \sim$ IID Noise, then

$$X_t = g(Z_t, Z_{t-1}, \ldots, Z_{t-q})$$

is **q-dependent**. This means that X_t and X_s are independent if and only if $|t-s| > q$.

We say a time series $\{X_t\}$ is **q-correlated** if $\gamma(h) = 0$ for $|h| > q$. This is a weaker condition.

A common class of models is when we take g to be a linear combination/transformation/filter.

The Moving Average Process

MA(q) Process

$\{X_t\}$ is a **moving-average process of order q** if

$$X_t = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

where $\{Z_t\} \sim \text{WN}(0, \sigma^2)$, $\{\theta_i\} \in \mathbb{R}$

Example

The definition was:

$$X_t = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$, $\{\theta_i\} \in \mathbb{R}$

Prove that

1. $\{X_t\}$ is stationary
2. if we upgraded $\{Z_t\}$ to IID Noise, then it's strictly stationary

Important: every stationary, mean 0, q-correlated process is an MA(q) process.

Forecasting

Now let's discuss forecasting. What are our options?

Conditional Distributions of Multivariate Normals

Assume

$$\mathbf{x} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} \sim \text{Normal}\left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma_1} & \boldsymbol{\Sigma_{12}} \\ \boldsymbol{\Sigma_{21}} & \boldsymbol{\Sigma_2} \end{bmatrix} \right)$$

Then

$$\mathbf{X}_1 | \mathbf{X}_2 = \mathbf{x}_2 \sim \text{Normal}\left(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma_{12}}\boldsymbol{\Sigma_2}^{-1}[\mathbf{x}_2 - \boldsymbol{\mu}_2], \boldsymbol{\Sigma_1} - \boldsymbol{\Sigma_{12}}\boldsymbol{\Sigma_2}^{-1}\boldsymbol{\Sigma_{21}} \right)$$

In particular,

bivariate case

$$X_1 | X_2 = x_2 \sim \text{Normal}\left(\mu_1 + \frac{\sigma_1}{\sigma_2}\rho[x_2 - \mu_2], \sigma_1^2(1 - \rho^2) \right)$$

where $\rho = \frac{\sigma_{1,2}}{\sqrt{\sigma_1^2 \sigma_2^2}}$ is the correlation, not the autocorrelation.

Watch out!

Note that the formula is different when $\rho(h) = \gamma(h)/\gamma(0)$ and we assume the time series is stationary

general bivariate case

$$X_1|X_2 = x_2 \sim \text{Normal}\left(\mu_1 + \frac{\sigma_1}{\sigma_2}\rho[x_2 - \mu_2], \sigma_1^2(1 - \rho^2)\right)$$

becomes

stationary time series case

$$X_{n+h}|X_n = x_n \sim \text{Normal}\left(\mu + \rho(h)[x_n - \mu], \sigma^2[1 - \rho(h)^2]\right)$$

Two Different Questions

At time n we know $X_n = x_n$.

Whats the function (possibly nonlinear) $m(x_n)$ that minimizes

$$E[(X_{n+h} - m(X_n))^2],$$

or in other words, minimizes **mean squared error**?

What's the best linear function $\ell(x_n) = ax_n + b$ that minimizes MSE?

$$E[(X_{n+h} - aX_n - b)^2]$$

Two Different Questions

If our time series $X_1, X_2, \ldots$ is jointly Normal, then $\ell() = m()$, and

Linear Gaussian Prediction

$$\ell(X_n) = m(X_n) = E[X_{n+h}|X_n] = \mu + \rho(h)[X_n - \mu] \tag{1}$$

and

$$\begin{aligned} \text{MSE}[m(X_n)] &= E[(X_{n+h} - m(X_n))^2] \\ &= E\left\{\text{Var}[X_{n+1}|X_n = x_n]\right\} \\ &= \sigma^2[1 - \rho(h)^2]. \end{aligned}$$

If our time series is not jointly Normal, $\ell() \neq m()$, but $\ell(X_n) = \mu + \rho(h)[X_n - \mu]$, still, and it has the same MSE.