743- Regression and Time Series

Mamikon S. Ginovyan

The Partial Autocorrelation Function (PACF)

The Partial Autocorrelation Function

Motivation.

- Let $\{X_t, t = 0, \pm1, \ldots\}$ be a second-order stationary process with mean $\mu = 0$, ACVF $\gamma(h)$ and ACF $\rho(h)$.

- In addition to the autocorrelation between X_t and X_{t+k}:

$$\text{ACF} = \rho(k) = \text{Corr}\,(X_t, X_{t+k}),$$

it is of interest to consider the **conditional correlation** between X_t and X_{t+k} after removing their mutual linear dependency on the intervening variables $X_{t+1}, X_{t+2}, \ldots, X_{t+k-1}$.

The conditional correlation:

$$P_k := Corr\left(X_t, X_{t+k} \middle| X_{t+1}, \ldots, X_{t+k-1}\right)$$

is called the **Partial Autocorrelation Function (PACF)** of X_t.

The Partial Autocorrelation Function

Definition based on Prediction. Let

$$\hat{X}_{t+k} = P\left(X_{t+k} \mid X_{t+k-1}, \ldots, X_{t+1}\right) = \text{ forward predictor}$$

$$\hat{X}_{t} = P\left(X_{t} \mid X_{t+k-1}, \ldots, X_{t+1}\right) = \text{ backward predictor}$$

$$(X_{t+k} - \hat{X}_{t+k}) \text{ and } (X_t - \hat{X}_t) = \text{ prediction errors.}$$

The **Partial Autocorrelation** (P_k) between $\boldsymbol{X_t}$ and $\boldsymbol{X_{t+k}}$ is defined to be the **ordinary autocorrelation** between $(X_t - \hat{X}_t)$ and $(X_{t+k} - \hat{X}_{t+k})$, that is,

$$P_k = \frac{Cov[(X_t - \hat{X}_t), (X_{t+k} - \hat{X}_{t+k})]}{\sqrt{Var(X_t - \hat{X}_t)}\sqrt{Var(X_{t+k} - \hat{X}_{t+k})}}.$$

The Partial Autocorrelation Function

$$
\phi_{kk} = P_k = \frac{\begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{pmatrix}}{\begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{pmatrix}} \qquad (4).
$$

Remark –Notation: As a function of $\boldsymbol{k}$, $\boldsymbol{\phi}_{kk}$ is called the **partial autocorrelation function (PACF)**, and is denoted by

$$\boldsymbol{\alpha(k) = \phi_{kk} = P_k} .$$

Examples

1. White Noise: $\boldsymbol{X_t} = \boldsymbol{Z_t} = \boldsymbol{\varepsilon_t} \sim \boldsymbol{WN}(0, \sigma^2)$.

The autocorrelation and partial autocorrelation functions:

$$ACF = r_k = \begin{cases} 1, & k = 0, \\ 0, & k \neq 0, \end{cases}$$

$$PACF = f_{kk} = \begin{cases} 1, & k = 0, \\ 0, & k \neq 0. \end{cases}$$

Examples

2. AR(1) Model: $X_t - \phi_1 X_{t-1} = Z_t$.

The autocorrelation and partial autocorrelation functions:

$$r_0 = 1;$$

$$r_k = j_1 r_{k-1},$$

$$r_k = f_1^{|k|} \quad \text{for} \quad k = \pm 1, \pm 2, \ldots$$

$$f_{kk} = \begin{cases} r_1 = f_1, & \text{for} \quad k = 1, \\ 0, & \text{for} \quad k \geq 2, \end{cases}$$

Examples

3. AR(2) Model: $X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} = Z_t$.

The autocorrelation and partial autocorrelation functions:

$$r_1 = \frac{f_1}{1-f_2}; \; r_2 = f_2 + \frac{f_1^{\,2}}{1-f_2};$$

$$r_k = f_1 r_{k-1} + f_2 r_{k-2}, \text{ for } k = \pm 3, \pm 4, \ldots$$

$$f_{11} = r_1 = \frac{f_1}{1-f_2};$$

$$f_{22} = f_2;$$

$$f_{kk} = 0 \quad \text{for all } \; k \geq 3.$$

Examples

4. AR(p) Model: $X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = Z_t$.

Because $r_k = f_1 r_{k-1} + f_2 r_{k-2} + \mathbf{L} + f_p r_{k-p}$ for $k > 0$,
we can easily see that when $k > p$ the last column of the matrix in the numerator of ϕ_{kk} in (4) can be written as a linear combination of previous columns of the same matrix.

Hence, the PACF ϕ_{kk} will vanish after *lag* p:

$$\phi_{kk} = 0 \quad \text{for all} \quad k > p.$$

Recall a similar propery for **MA(q)** models:

$$\rho_k = 0 \quad \text{for all} \quad k > q.$$

This properties are useful in identifying an AR and MA models as generating processes for time series.

Examples

5. MA(1) Model: $X_t = Z_t - \theta_1 Z_{t-1}$.

The autocorrelation and partial autocorrelation functions:

$$r_{\pm 1} = \frac{-q_1}{1+q_1^2}; \qquad r_k = 0 \quad \text{for} \quad k = \pm 2, \pm 3, \ldots.$$

$$f_{11} = r_1 = \frac{-q_1}{1+q_1^2} = \frac{-q_1(1-q_1^2)}{1-q_1^4}$$

$$f_{22} = -\frac{r_1^2}{1-r_1^2} = \frac{-q_1^2}{1+q_1^2+q_1^4} = \frac{-q_1^2(1-q_1^2)}{1-q_1^6}$$

$$f_{33} = \frac{r_1^3}{1-2r_1^2} = \frac{-q_1^3}{1+q_1^2+q_1^4+q_1^6} = \frac{-q_1^3(1-q_1^2)}{1-q_1^8}.$$

$$f_{kk} = \frac{-q_1^k(1-q_1^2)}{1-q_1^{2(k+1)}} = \frac{-(q_1)^k}{1+q_1^2+\cdots+q_1^{2k}} \qquad \text{for} \quad k \geq 1.$$

Examples

<u>6. MA(2) Model:</u> $\boldsymbol{X_t = Z_t - \theta_1 Z_{t-1} - \theta_2 Z_{t-2}}$.

The autocorrelation and partial autocorrelation functions:

$$r_1 = \frac{q_1(q_2 - 1)}{1 + q_1^2 + q_2^2}; \quad r_2 = \frac{-q_2}{1 + q_1^2 + q_2^2};$$

$$r_k = 0 \quad \text{for} \quad k = \pm 3, \pm 4, \ldots.$$

$$f_{11} = r_1$$

$$f_{22} = \frac{r_2 - r_1^2}{1 - r_1^2}$$

$$f_{33} = \frac{r_1^3 - r_1 r_2 (2 - r_2)}{1 - r_2^2 - 2 r_1^2 (1 - r_2)}$$

M

Elements of Spectral Analysis

Spectral representation of covariance function

We consider a discrete-time stationary time-series $\boldsymbol{X_t}, t \in Z = \{0, \pm 1, \ldots\}$ with covariance function $g(t), t \in Z$.

First recall the characterization theorem of covariance functions.

Theorem 1.
A complex-valued function $g(t), t \in Z$, is the covariance function of a (generally, complex-valued) stationary time series $\boldsymbol{X_t}$ if and only if γ(t) is non-negative definite, that is, for all $n \in N$ and $z_1, \mathbf{L}, z_n \in C$

$$\sum_{kj=1}^{n} r(k-j) z_k \overline{z_j} \geq 0 \qquad (1)$$

Spectral representation of covariance function

The **Herglotz's theorem**, which we are going to state, characterizes the class of non-negative definite functions as the functions which can be represented in the form of **Fourier – Stieltjes integral (transform)** of some distribution function.

Theorem2. **(Herglotz).** A complex-valued function $g(t)$, $t \in Z$ is non-negative definite if and only if there exists a distribution function $F(l)$, $l \in [-p, p]$ (that is, ***F***(λ) is right-continuous, non-decreasing, bounded function), such that for every $t \in Z$

$$g(t) = \int_{-p}^{p} e^{ilt} dF(l) \, .$$

Spectral representation of covariance function

Combining Theorems 1 and 2, we obtain the following fundamental result of the time-series analysis.

Theorem 3 (Spectral Representation of the ACVF).
A complex-valued function $g(t), t \in Z$, is the covariance function of a stationary time-series $X(t), t \in Z$, **if and only if** $\boldsymbol{\gamma}$(t) admits the representation

$$g(t) = \int_{-p}^{p} e^{ilt} dF(l), \qquad (2)$$

where $\boldsymbol{F}(\lambda)$ is a right-continuous, non-decreasing, bounded function on $[-\pi, \pi]$ and $\boldsymbol{F}(-\pi)=0$.

Spectral representation of covariance function

Remark 1.

- The function $F(\lambda)$ is called **spectral distribution** function of both ACVF $\gamma(t)$ and process $X(t)$.

- The representation (2) is called **spectral representation** of ACVF $\gamma(t)$.

- The set of **jump points** of $F(\lambda)$ is called the **spectrum** of the process $X(t)$.

Spectral representation of covariance function

- If $F(\lambda)$ is **absolute continuous**, that is,

$$F(l) = \int_{-p}^{l} f(n)dn, \quad -p \le l \le p, \qquad (3)$$

then the function

$$f(l) := F'(l) = \frac{dF(l)}{dl}, \quad -p \le l \le p, \qquad (4)$$

is called **spectral density function** of $X(t)$.

- In this case (2) becomes (since $dF(\lambda) = f(\lambda)d\lambda$)

$$g(t) = \int_{-p}^{p} e^{ilt} f(l)dl,$$

that is, $g(t), t \in Z$ are the **Fourier coefficients** of $f(\lambda)$.

Spectral representation of covariance function

Remark 3.

If $\boldsymbol{X}(\boldsymbol{t})$ is a **real-valued** stationary process then the covariance $\boldsymbol{\gamma}(\boldsymbol{t})$ is real-valued and even ($\boldsymbol{\gamma}(\boldsymbol{t}) = \boldsymbol{\gamma}(-\boldsymbol{t})$) function, implying that its spectral density function is even: $\boldsymbol{f}(\lambda) = \boldsymbol{f}(-\lambda)$.

In this case, we have

$$g(t) = \int_{-p}^{p} \cos(l\,t)\,dF(l) = 2\int_{0}^{p} \cos(l\,t)\,f(l)\,dl$$

Spectral representation of covariance function

Remark 4.

The remarkable result in Frequency-domain analysis is that:
Every zero-mean stationary process can be represented as

$$X_t = \int_{(-p\ ,p\ ]} e^{ih\,l}\, dZ(\,l\,), \qquad (5)$$

where $Z(\lambda), -\pi < \lambda \leq \pi\}$ is a complex-valued process with orthogonal (or uncorrelated) increments.

The representation (5) of a stationary process $\{X_t\}$ is called the **spectral representation of the process,** and should be compared with the corresponding spectral representation (2) of the ACVF $\gamma(t\,)$.

Spectral Densities

We naturally rise the following questions:

1. Under what conditions the **spectral density function** $f(\lambda)$ of a time series $X(t)$ exists? Or more precisely: Describe the class of covariance functions for which the spectral density $f(\lambda)$ exists.

2. Under what conditions given function $f(\lambda)$ is a **spectral density function** of a time series $X(t)$?

3. List the properties of **spectral density function** $f(\lambda)$, and write explicitly the spectral densities of common **ARMA** models.

Spectral Densities

The following proposition characterizes spectral densities.

Proposition 1.

A real-valued function $\boldsymbol{f}(\lambda)$ defined on $(-\pi, \pi]$ is the spectral density of a stationary process $\{X_t, t \in Z\}$ **if and only if**

(i) $\boldsymbol{f}(\lambda) \geq 0$ and

(ii) $\int_{-p}^{p} f(l)\,d\,l < \infty.$

Remark 1.
The underlying process $\{X_t, t \in Z\}$ is **real-valued** if and only if $\boldsymbol{f}(\lambda)$ is an **even** function: $\boldsymbol{f}(\lambda) = \boldsymbol{f}(-\lambda)$.

Spectral Densities

Theorem (Existence of spectral density).

An absolutely summable function $g(h): \sum_{h=-\infty}^{\infty} |g(h)| < \infty$ is the ACVF of a real-valued stationary time series $\{X_t, t \in Z\}$ **if and only if** it is **even:** $\boldsymbol{f}(\lambda) = \boldsymbol{f}(-\lambda)$, and

$$f(l) = \frac{1}{2p} \sum_{h=-\infty}^{\infty} e^{-ihl} g(h) \geq 0 \quad \text{for all } l \in (-p, p],$$

in which case $\boldsymbol{f}(\lambda)$ is the spectral density of $\gamma(h)$.

Spectral Densities

Example 1.

Show that the function

$$k(h) = \begin{cases} 1, & \text{if } h = 0 \\ r, & \text{if } h = \pm 1 \\ 0, & \text{otherwise} \end{cases}$$

is the ACVF of a stationary time series if and only if $|r| \leq 1/2$.

Spectral Densities

Indeed:

Since $\boldsymbol{k}(h)$ is even and nonzero only at lags **0, ±1,** it follows from Theorem that $\boldsymbol{k}(h)$ is an ACVF if and only if the function

$$f(l) = \frac{1}{2p}\sum_{h=-\infty}^{\infty} e^{-ihl} g(h)$$

$$= \frac{1}{2p}\left[1 + r e^{-il} + r e^{il}\right] \quad (\text{using } \; e^{-il} + e^{il} = 2\cos(l))$$

$$= \frac{1}{2p}\left[1 + 2r\cos l\right] \quad \text{is nonnegative for all } \; l \in [-p, p].$$

But this occurs if and only if $|r| \le 1/2.$

Spectral Densities

Remark. **Not all ACVF's have a spectral density.**

For example, the stationary time series

$$\boldsymbol{X_t} = \boldsymbol{Z_1}\, cos(\omega t) + \boldsymbol{Z_2}\, sin(\omega t)$$

where $\boldsymbol{Z_1}$ and $\boldsymbol{Z_2}$ are uncorrelated random variables with

mean **0** and variance **1**, has ACVF $\boldsymbol{\gamma(h) = cos(\omega h)}$,

which is not expressible as $\int_{-p}^{p} e^{ihl} f(l)\,dl$, with $\boldsymbol{f}$ a function

on $(-\pi, \pi]$.

Spectral Densities

Nevertheless, the function $\boldsymbol{\gamma(h) = cos(\omega h)}$ can be written as the Fourier transform of the discrete distribution function

$$F(l) = \begin{cases} 0 & \text{if} \quad l < -w \\ 0.5 & \text{if} \quad -w \le l < w \\ 1.0 & \text{if} \quad l \ge w \end{cases}$$

that is,

$$g\,(\mathrm{h}) = \cos\left(wh\right) = \int_{-p}^{p} e^{ihl}\, dF(l).$$

Examples.

Example 1. (White noise)

If $\{X_t\} \sim WN(0, \sigma^2)$, then $\gamma(0) = \sigma^2$ and $\gamma(h) = 0$ for all $|h| > 0$.

This process has a **flat spectral density** :

$$f(l) = \frac{s^{2}}{2p}, \quad -p \leq l \leq p.$$

A process with this spectral density is called **white noise**, since each frequency in the spectrum contributes equally to the variance of the process.

Examples.

Example 2. (Spectral density of an AR(1) process).

If $X_t = \phi X_{t-1} + Z_t$, where $\{Z_t\} \sim WN(0, \sigma^2)$, then X_t has a spectral density

$$f(\lambda) = \frac{\sigma^2}{2\pi(1-\phi^2)}\left(1 + \sum_{h=1}^{\infty} \phi^h \left(e^{-ih\lambda} + e^{ih\lambda}\right)\right)$$

$$= \frac{\sigma^2}{2\pi(1-\phi^2)}\left(1 + \frac{\phi e^{i\lambda}}{1-\phi e^{i\lambda}} + \frac{\phi e^{-i\lambda}}{1-\phi e^{-i\lambda}}\right)$$

$$= \frac{\sigma^2}{2\pi}\left(1 - 2\phi\cos\lambda + \phi^2\right)^{-1}.$$

Examples.

Example 3. (Spectral density of an MA(1) process).

If $\boldsymbol{X_t} = \boldsymbol{Z_t} + \boldsymbol{\theta Z_{t\text{-}1}}$, where $\{\boldsymbol{Z_t}\} \sim \boldsymbol{WN}(0, \sigma^2)$, then $\boldsymbol{X_t}$ has a spectral density

$$f(l) = \frac{s^2}{2p}\left(1 + q^2 + q\left(e^{-il} + e^{il}\right)\right)$$

$$= \frac{s^2}{2p}\left(1 + 2q \cos l + q^2\right).$$

Examples.

Example 4. (Spectral density of an ARMA(*p*,*q*) process) .

If $\boldsymbol{X(t)}$ is the solution of equation

$$X_t - f_1 X_{t-1} - \ldots - f_p X_{t-p} = Z_t + q_1 Z_{t-1} + \ldots + q_q Z_{t-q}, \qquad (1)$$

where $\{\boldsymbol{Z_t}\} \sim \boldsymbol{WN}(0, \sigma^2)$ and the polynomials

$$f(z) = 1 - f_1 z - \ldots - f_p z^p \text{ and } q(z) = 1 + q_1 z + \ldots + q_q z^q$$

have no common zeros, and $\boldsymbol{\phi}$**(z)** has no zeros on the unit circle **{z: |z|=1}**, then (1) has a stationary solution $\boldsymbol{X(t)}$ with **spectral density function:**

$$f(l) = \frac{s^2}{2p} \cdot \frac{\left|q(e^{-il})\right|^2}{\left|f(e^{-il})\right|^2}, \quad -p \le l \le p.$$