

2. Stationary Processes

2.1 Properties of Stationary Processes

The objective of the previous section was to remove (systematic) deterministic components from an observed series to leave only the stochastic part. In this section, we study the basic properties of *stationary* processes: such processes are inherently stable (in the long run), and form natural models for the stochastic component of observed series.

In the main, we will focus on **weakly** stationary processes, where the mean and covariance structure is stable over time; in general we will not make distributional assumptions.

Initially, we focus on forecasting, and demonstrate the utility of utilizing the covariance structure of the process.

2.1 Properties of Stationary Processes

Consider first a stationary Gaussian process $\{X_t\}$, where for all t and n ,

$$(X_{t+1}, \dots, X_{t+n})$$

is jointly normally distributed, with $\mathbb{E}[X_t] = \mu$, $\text{Var}[X_t] = \sigma^2$, and for $s > t$.

$$\text{Cov}[X_t, X_s] = \gamma_X(s - t) \quad \text{Corr}[X_t, X_s] = \rho_X(s - t) = \frac{\gamma_X(s - t)}{\sigma^2}$$

Then it follows that

$$X_{n+h}|X_n = x_n \sim \mathcal{N}(m_h(x_n), v_h(x_n))$$

where

$$\begin{aligned} m_h(x) &= \mu + \rho_X(h)(x - \mu) \\ v_h(x) &= \sigma^2(1 - \rho_X^2(h)) \end{aligned}$$

2.1 Properties of Stationary Processes

To define “optimal” forecasting, we need to define a criterion via which to assess the quality of predictions.

We use a **minimum mean-square error** (minimum MSE) criterion, and attempt to make the prediction $\hat{X}_{n+h} = \hat{x}_{n+h}$ such that

$$\mathbb{E}[(X_{n+h} - \hat{x}_{n+h})^2 | X_n = x_n]$$

is **minimized**. Here the expectation is taken with respect to the conditional distribution $X_{n+h} | X_n = x_n$.

2.1 Properties of Stationary Processes

This minimization can be carried out analytically to yield

$$\hat{x}_{n+h} = \mathbb{E}[X_{n+h}|X_n = x_n] = m_h(x) = \mu + \rho_X(h)(x_n - \mu)$$

that is, the best prediction (in minimum prediction MSE terms) is the conditional expectation. The prediction MSE corresponding to this prediction is

$$v_h(x) = \sigma^2(1 - \rho_X^2(h))$$

In this case, we could deduce the appropriate form for the optimal predictor from properties of the Gaussian process.

2.1 Properties of Stationary Processes

Best Linear Prediction: For a non-Gaussian process, we seek the again seek the minimum MSE predictor. We restrict attention to **linear** predictors, that is, predictors of the form

$$\hat{X}_{n+h} = l_h(X_n) = aX_n + b$$

for some a and b . Under this restriction, and stationarity, it again transpires that the optimal predictor takes the form

$$\hat{x}_{n+h} = \mathbb{E}[X_{n+h}|X_n = x_n] = m_h(x) = \mu + \rho_X(h)(x_n - \mu)$$

with the same prediction MSE as before.

2.1 Properties of Stationary Processes

Non-negative definite functions: A scalar function g is *non-negative definite* if, for all n , and vectors $a = (a_1, \dots, a_n)$

$$\sum_{i=1}^n \sum_{j=1}^n a_i g(i-j) a_j \geq 0.$$

With a strict inequality, the function is termed *positive definite*.

If G is the $n \times n$ matrix formed as $G = [g(i-j)]_{ij}$, then the above definition becomes that

$$a^T G a \geq 0.$$

and G is a non-negative/positive definite matrix.

2.1 Properties of Stationary Processes

Theorem

*A function g is the autocovariance function (acvf) for a stationary stochastic process if and only if g is **even**, that is*

$$g(h) = g(-h) \quad \text{for all } h \in \mathbb{Z}$$

*and g is **non-negative definite**.*

Note: by elementary properties of covariance, for all h ,

$$\begin{aligned}\gamma(0) &\geq 0 \\ |\gamma(h)| &\leq \gamma(0) \\ \gamma(-h) &= \gamma(h)\end{aligned}$$

Characterizing the ACVF I

Proof. Suppose first that g is the acvf of a stationary stochastic process, $\{X_t\}$ say. By definition of covariance, g is an even function, as

$$g(-h) = \mathbb{E}[X_t X_{t-h}] = \mathbb{E}[X_{t+h} X_t] = \mathbb{E}[X_t X_{t+h}] = g(h).$$

Let $a \in \mathbb{R}^n$ be an $n \times 1$ vector, and let $X_{n:1} = (x_n, \dots, x_1)^\top$. Then $a^\top X_{n:1}$ is a scalar random variable, and

$$\text{Var}[a^\top X_{n:1}] = a^\top \text{Var}[X_{n:1}] a = \sum_{i=1}^n \sum_{j=1}^n a_i g(i-j) a_j \geq 0$$

so g is non-negative definite.

Characterizing the ACVF II

Now suppose that g is non-negative definite and even. For each $n \geq 1$, define the $n \times n$ matrix Σ by

$$[\Sigma]_{ij} = g(i - j)$$

and consider the multivariate Gaussian distribution with mean zero and variance-covariance matrix Σ . This distribution is the finite dimensional distribution arising from a stationary Gaussian process with the specified covariance structure. $\square$

Note that an autocorrelation function has the same properties, except that $\rho(0) = 1$.

2.1 Properties of Stationary Processes

Example (Simple harmonic process)

Suppose A and B are zero mean, uncorrelated random variables with variance 1. Let $\{X_t\}$ be defined by

$$X_t = A \cos(\omega t) + B \sin(\omega t)$$

Then $\mathbb{E}[X_t] = 0$ for all t , and

$$\begin{aligned} \text{Cov}[X_t, X_{t+h}] &= \mathbb{E}[X_t X_{t+h}] \\ &= \mathbb{E}[(A \cos(\omega t) + B \sin(\omega t)) \\ &\quad (A \cos(\omega(t+h)) + B \sin(\omega(t+h)))] \\ &= \mathbb{E}[A^2] \cos(\omega t) \cos(\omega(t+h)) + \\ &\quad \mathbb{E}[B^2] \sin(\omega t) \sin(\omega(t+h)) \\ &= \cos(\omega t) \cos(\omega(t+h)) + \sin(\omega t) \sin(\omega(t+h)) \\ &= \cos(\omega h) = \cos(\omega(-h)) \end{aligned}$$

2.1 Properties of Stationary Processes

Construction of stationary processes: Let $\{Z_t\}$ be an IID process, and let $\{X_t\}$ be defined by

$$X_t = f(Z_t, Z_{t-1}, \dots, Z_{t-q})$$

for some $q \geq 0$. $\{X_t\}$ is a *filtered* version of $\{Z_t\}$, and $\{X_t\}$ is strictly stationary if $\{Z_t\}$ is strictly stationary.

Note that X_t and X_s are independent if $|t - s| > q$; $\{X_t\}$ is termed *q-dependent*, and $\gamma_X(h) = 0$ if $|h| > q$.

2.1 Properties of Stationary Processes

Moving average processes: Let $\{Z_t\} \sim \text{WN}(0, \sigma^2)$ be a white noise process, and let $\{X_t\}$ be defined by

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q}$$

for some $q \geq 0$, where $\theta_1, \theta_2, \dots, \theta_q$ are real-valued constants.

$\{X_t\}$ is zero mean, stationary and q -dependent; it is termed a *moving average process of order q* (MA(q)).

2.2 Linear Processes

$\{X_t\}$ is a linear process if for all t

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}$$

where

- $\{Z_t\} \sim \text{WN}(0, \sigma^2)$
- $\{\psi_j\}$ is a real sequence with

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$$

Using the backshift operator notation

$$X_t = \psi(B)Z_t$$

where

$$\psi(z) = \sum_{j=-\infty}^{\infty} \psi_j z^j.$$

2.2 Linear Processes

The process is termed *non-anticipating* if $\psi_j = 0$ for all $j < 0$, so that

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}.$$

Note that for the linear process to be sensible we need the sum

$$\sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}$$

to converge in some appropriate fashion; in this context we require *mean-square convergence*, so that in fact

$$X_t \stackrel{\text{m.s.}}{=} \sum_{j=0}^{\infty} \psi_j Z_{t-j}.$$

2.2 Linear Processes

We require for the definition

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\left(X_t - \sum_{j=-n}^n \psi_j Z_{t-j} \right)^2 \right] = 0$$

Note that

$$\left(\sum_{j=-n}^n \psi_j Z_{t-j} \right)^2 = \sum_{j=-n}^n \psi_j^2 Z_{t-j}^2 + \sum_{j=-n}^n \sum_{k \neq j}^n \psi_j \psi_k Z_{t-j} Z_{t-k}$$

and taking expectations, as $\{Z_t\}$ is a white noise process so that $\mathbb{E}[Z_{t-j} Z_{t-k}] = 0$ if $j \neq k$

$$\mathbb{E} \left[\left(\sum_{j=-n}^n \psi_j Z_{t-j} \right)^2 \right] = \sigma^2 \sum_{j=-n}^n \psi_j^2.$$

2.2 Linear Processes

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty \implies \sum_{j=-\infty}^{\infty} \psi_j^2 < \infty$$

so therefore the partial sums (with limits $\pm n$) are also finite for all n . Note also that by the Cauchy-Schwarz inequality

$$\mathbb{E}[|Z_t|] \leq \sqrt{\mathbb{E}[Z_t^2]} = \sigma, \text{ therefore}$$

$$\mathbb{E}[|X_t|] \leq \sum_{j=-\infty}^{\infty} |\psi_j| \mathbb{E}[|Z_{t-j}|] \leq \sigma \sum_{j=-\infty}^{\infty} |\psi_j| < \infty$$

so that

$$\sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}$$

is convergent (almost surely).

2.2 Linear Processes

The operator $\psi(B)$ is termed a *linear filter*; it acts on the stationary (white noise) process $\{Z_t\}$ to produce the stationary (autocorrelated) process $\{X_t\}$.

The operator can also be applied to any stationary process. Suppose that $\{Y_t\}$ is a weakly stationary, zero mean process with ACVF γ_Y , and $\psi(B)$ is the above linear filter with absolutely summable coefficients. Then the process

$$X_t = \psi(B)Y_t = \sum_{j=-\infty}^{\infty} \psi_j Y_{t-j}$$

is also weakly stationary with mean zero. The ACVF of $\{X_t\}$ is

$$\gamma_X(h) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \gamma_Y(h + j - k).$$

2.2 Linear Processes

If $\{Y_t\}$ is a white noise process, so that $\{X_t\}$ is a linear process, then

$$\gamma_X(h) = \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+h} \sigma^2.$$

To see this, note that

$$\mathbb{E}[X_t] \leq \mathbb{E}[|X_t|] \leq \sum_{j=-\infty}^{\infty} |\psi_j| \mathbb{E}[|Y_t|] < \infty$$

so that $\mathbb{E}[X_t] = 0$ for all t . Then

$$\mathbb{E}[X_t X_{t+h}] = \mathbb{E} \left[\left(\sum_{j=-\infty}^{\infty} \psi_j Y_{t-j} \right) \left(\sum_{k=-\infty}^{\infty} \psi_k Y_{t+h-k} \right) \right]$$

2.2 Linear Processes

Multiplying out yields

$$\begin{aligned}\mathbb{E}[X_t X_{t+h}] &= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \mathbb{E}[Y_{t-j} Y_{t+h-k}] \\ &= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \gamma_Y(j+h-k).\end{aligned}$$

If $\{Y_t\}$ is white noise, then $\gamma_Y(h) = \sigma^2$ if $h = 0$ and zero otherwise, in which case

$$\gamma_X(h) = \mathbb{E}[X_t X_{t+h}] = \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+h} \sigma^2.$$

2.2 Linear Processes

If $\alpha(B)$ and $\beta(B)$ are two linear filters, each of which has absolutely summable coefficients, and

$$\psi(B) = \alpha(B)\beta(B) = \beta(B)\alpha(B)$$

then $\psi(B)$ is also a linear filter, with

$$\psi_j = \sum_{k=-\infty}^{\infty} \alpha_k \beta_{j-k} = \sum_{k=-\infty}^{\infty} \beta_k \alpha_{j-k}$$

and if $\{Y_t\}$ is a stationary process, then

$$X_t = \psi(B)Y_t = \alpha(B)\beta(B)Y_t$$

defines a stationary process $\{X_t\}$. Note that $W_t = \beta(B)Y_t$ is also stationary, and $X_t = \alpha(B)W_t$.

2.2 Linear Processes

Example (AR(1) process)

Let $\{Z_t\} \sim \text{WN}(0, \sigma^2)$, $|\phi| < 1$, and

$$X_t = \phi X_{t-1} + Z_t$$

with $\mathbb{E}[Z_s X_t] = 0$ for $s > t$. That is, $\{X_t\}$ is the solution of the equation

$$X_t - \phi X_{t-1} = Z_t$$

or

$$(1 - \phi(B))X_t = Z_t.$$

2.2 Linear Processes

Example (AR(1) process)

Consider the linear process

$$X_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j}$$

- note that

$$\sum_{j=0}^{\infty} |\phi|^j = \frac{1}{1 - |\phi|} < \infty$$

hence the conditions of absolute summability hold.

2.2 Linear Processes

Example (AR(1) process)

This is a solution to the equation above as

$$\begin{aligned}X_t &= \phi X_{t-1} + Z_t \\&= \phi(\phi X_{t-2} + Z_{t-1}) + Z_t \\&\vdots \\&= \lim_{n \rightarrow \infty} \left\{ \phi^n X_{t-n} + \sum_{j=0}^{n-1} \phi^j Z_{t-j} \right\} \\&= \sum_{j=0}^{\infty} \phi^j Z_{t-j}\end{aligned}$$

2.2 Linear Processes

Example (AR(1) process)

Also, for $h \geq 0$,

$$\gamma_X(h) = \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+h} \sigma^2 = \sum_{j=0}^{\infty} \phi^j \phi^{j+h} \sigma^2 = \sigma^2 \sum_{j=0}^{\infty} \phi^{2j+h}$$

that is

$$\gamma_X(h) = \frac{\sigma^2 \phi^h}{1 - \phi^2}$$

2.2 Linear Processes

Example (AR(1) process)

Now suppose that $\{Y_t\}$ is another solution to the AR(1) equation. Then, for any $k \geq 0$

$$\begin{aligned}\mathbb{E} \left[\left(Y_t - \sum_{j=0}^k \phi^j Z_{t-j} \right)^2 \right] &= \mathbb{E} \left[\left(\sum_{j=k+1}^{\infty} \phi^j Z_{t-j} \right)^2 \right] \\ &= \mathbb{E} \left[\left(\sum_{l=0}^{\infty} \phi^{l+k+1} Z_{t-k-1-l} \right)^2 \right] \\ &= \mathbb{E} \left[\left(\sum_{l=0}^{\infty} \phi^l Z_{t-k-1-l} \right)^2 \right] \\ &= \phi^{2(k+1)} \mathbb{E} \left[(Y_{t-k-1})^2 \right]\end{aligned}$$

2.2 Linear Processes

Example (AR(1) process)

The right hand side of this equation converges to zero as $k \rightarrow \infty$, as $|\phi| < 1$, and $\mathbb{E}[Y_{t-k-1}^2] < \infty$.

Therefore Y_t and X_t are equal in mean square

$$Y_t \stackrel{\text{m.s.}}{=} X_t$$

and therefore $\{X_t\}$ is essentially the unique stationary solution to the AR(1) equation

$$X_t - \phi X_{t-1} = Z_t$$

with $|\phi| < 1$ and $\{Z_t\} \sim \text{WN}(0, \sigma^2)$. Therefore for any stationary AR process there is an essentially unique MA(∞) representation, with $\psi_j = \phi^j$ for $j \geq 0$, and $\psi_j = 0$ for $j < 0$.

2.2 Linear Processes

Example (Non-stationary AR(1) process)

If $|\phi| = 1$, say

$$X_t = X_{t-1} + Z_t \quad \text{or} \quad X_t = -X_{t-1} + Z_t$$

then X_t is not stationary. Clearly

$$\mathbb{V}ar[X_t] = \mathbb{V}ar[\pm X_{t-1}] + \mathbb{V}ar[Z_t] > \mathbb{V}ar[X_{t-1}]$$

for all t .

2.2 Linear Processes

Non-causal AR(1): In the AR(1) formulation, suppose $|\phi| > 1$, that is

$$X_t = \phi X_{t-1} + Z_t$$

so that

$$\frac{1}{\phi} X_t = X_{t-1} + \frac{1}{\phi} Z_t$$

or

$$X_{t-1} = \phi_* X_t + Z_t^*$$

where $\{Z_t\} \sim \text{WN}(0, \sigma^2/\phi^2)$, and $\phi_* = 1/\phi$, with $|\phi_*| < 1$.

2.2 Linear Processes

Therefore, by considering a reverse time formulation, we **can** construct a stationary solution to the AR(1) equation. It takes the linear process form

$$X_t = - \sum_{j=0}^{\infty} \phi_*^j Z_{t+j}^*$$

that is, X_t is defined in terms of future values of $\{Z_t\}$.

The ACVF can be computed by the usual methods. Such a process is termed **non-causal**.

2.2 Linear Processes

Note: we have that

$$(1 - \phi B)X_t = Z_t \quad \therefore \quad X_t = (1 - \phi B)^{-1}Z_t$$

and by a geometric series expansion,

$$X_t = \left\{ \sum_{j=0}^{\infty} \phi^j B^j \right\} Z_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j}$$

as before.

2.3 Autoregressive Processes

To generalize the previous formulation is straightforward. Suppose that

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t$$

is used to define (implicitly) $\{X_t\}$, where as usual $\{Z_t\} \sim \text{WN}(0, \sigma^2)$. This definition does not ensure (or even suggest) that the solution can be found, or is unique or stationary.

This process is termed the autoregressive process of order p (AR(p)). Here $\mathbb{E}[Z_t] = 0$, so $\mathbb{E}[X_t] = 0$.

2.3 Autoregressive Processes

Example (AR(2) process)

If $p = 2$ we have

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} = Z_t \quad \text{or} \quad \Phi(B)X_t = Z_t.$$

where $\Phi(B) = (1 - \phi_1 B - \phi_2 B^2)$, and the following linear process representation obtained by noting that

$$X_t = (1 - \phi_1 B - \phi_2 B^2)^{-1} Z_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}$$

where ψ_j is the coefficient of B^j in the series expansion of

$$(1 - \phi_1 B - \phi_2 B^2)^{-1}$$

2.3 Autoregressive Processes

Example (AR(2) process)

Consider the factorization

$$\Phi(z) = (1 - \phi_1 z - \phi_2 z^2) = (1 - \xi_1 z)(1 - \xi_2 z).$$

where z is an arbitrary complex number, and (ξ_1, ξ_2) are a complex conjugate pair. In general,

$$(1 - \xi z)^{-1}$$

has radius of convergence $|\xi| < 1$ when $|z| = 1$.

2.3 Autoregressive Processes

Example (AR(2) process)

(ξ_1, ξ_2) are the *reciprocal roots* of the equation

$$\Phi(z) = 0$$

and

$$\{(1 - \xi_1 z)(1 - \xi_2 z)\}^{-1}$$

is convergent for $|z| = 1$ if and only if

$$|\xi_1| = |\xi_2| < 1.$$

2.3 Autoregressive Processes

Example (AR(2) process)

We have that

$$\Phi(B)X_t = Z_t \quad \therefore \quad X_t = \{\Phi(B)\}^{-1}Z_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}$$

where ψ_j is the coefficient of z^j in the series expansion of

$$\{\Phi(z)\}^{-1} = \{(1 - \xi_1 z)(1 - \xi_2 z)\}^{-1} = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \xi_1^j \xi_2^k z^{j+k}$$

that is

$$\psi_j = \sum_{k=0}^j \xi_1^k \xi_2^{j-k}$$

2.3 Autoregressive Processes

Example (AR(2) process)

Note that

$$\begin{aligned} |\psi_j| &\leq \sum_{k=0}^j |\xi_1|^k |\xi_2|^{j-k} \\ &= (j+1)M^j \end{aligned}$$

where $M = \max\{|\xi_1|, |\xi_2|\} < 1$. Therefore

$$\sum_{j=0}^{\infty} |\psi_j| \leq \sum_{j=0}^{\infty} (j+1)M^j = \frac{1}{(1-M)^2} < \infty.$$

Hence $\{X_t\}$ is a linear process with absolutely summable coefficients.

2.3 Autoregressive Processes

The extension to the $AR(p)$ case is now straightforward: we write

$$\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p = \prod_{j=1}^p (1 - \xi_j z)$$

and by an identical argument to above, to ensure stationarity, we require that the reciprocal roots

$$\xi_1, \xi_2, \dots, \xi_p$$

all satisfy $|\xi_j| < 1$. Equivalently, if we let $\eta_j = 1/\xi_j$ denote the roots of $\Phi(z) = 0$, then we have

$$\Phi(z) = \prod_{j=1}^p (\eta_j - z)$$

and require $|\eta_j| > 1$ for each j .

2.3 Autoregressive Processes

If $|\xi_j| > 1$ for at least one j , then no causal stationary solution to the $\text{AR}(p)$ equation exists, but a similar approach to the $\text{AR}(1)$ case allows us to define a (unique) **non-causal** stationary solution.

Hence the only case that leads to **non-stationary** solutions is when $|\xi_j| = 1$ for at least one j . In the simplest case $\xi_j = 1$, which corresponds to the factor $(1 - B)$ in the AR polynomial; however any complex root with modulus 1 yields a non-stationary solution, that is

$$\xi_j = e^{\imath\omega}$$

for $-\pi < \omega \leq \pi$, where $\imath = \sqrt{-1}$.

2.3 Autoregressive Processes

Example

Suppose

$$X_t - \frac{13}{4}X_{t-1} + \frac{3}{4}X_{t-2} = Z_t.$$

Then

$$\Phi(z) = 1 - \frac{13}{4}z + \frac{3}{4}z^2 = (1 - 3z)(1 - z/4)$$

Therefore the reciprocal roots are 3 and 1/4. Therefore no causal stationary solution exists.

2.3 Autoregressive Processes

Example

Suppose

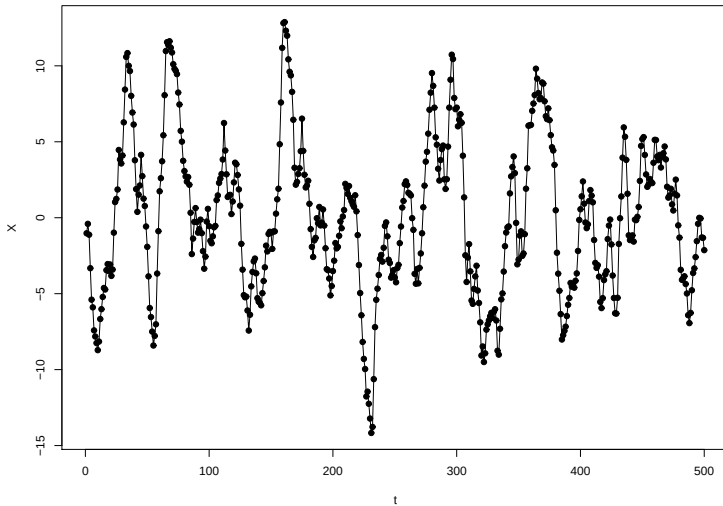
$$X_t - \frac{31}{20}X_{t-1} + \frac{3}{5}X_{t-2} = Z_t.$$

Then

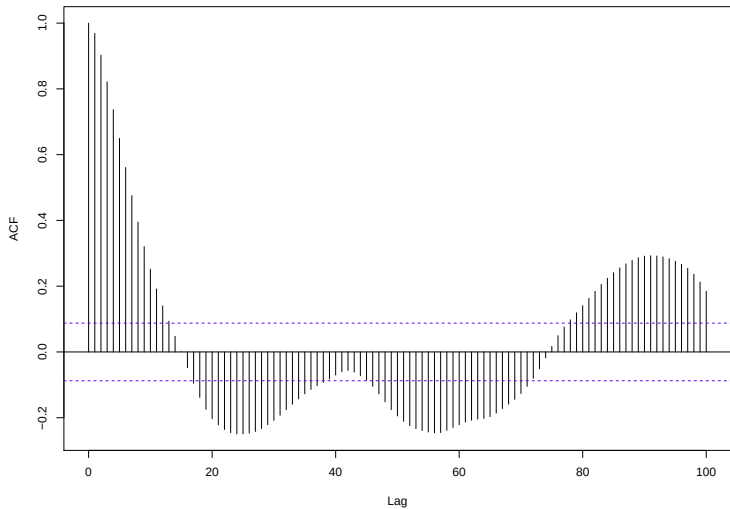
$$\Phi(z) = 1 - \frac{31}{20}z + \frac{3}{5}z^2 = (1 - 4z/5)(1 - 3z/4)$$

Therefore the reciprocal roots are $4/5$ and $3/4$. Therefore a causal stationary solution does exist.

2.3 Autoregressive Processes



2.3 Autoregressive Processes



2.4 ARMA Processes

Recall the two main classes of stationary linear processes: if $\{Z_t\} \sim \text{WN}(0, \sigma^2)$

- **Moving Average:**

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

- stationary
- q -dependent; $\gamma_X(h) = 0$ for $h > q$

- **Autoregressive:**

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t$$

- stationary provided the roots of the AR polynomial

$$\Phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p = \prod_{j=1}^p (1 - \xi_j z)$$

- satisfy $|\xi_j| \neq 1$ for each j ; causal if $|\xi_j| < 1$ for each j .
- yields an $\text{MA}(\infty)$ representation.

2.4 ARMA Processes

$\{X_t\}$ is an $ARMA(p, q)$ process if it is stationary and satisfies

$$\Phi(B)X_t = \Theta(B)Z_t$$

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

$$\begin{aligned}\Phi(z) &= 1 - \phi_1 z - \cdots - \phi_p z^p = \prod_{j=1}^p (1 - \xi_j z) \\ \Theta(z) &= 1 + \theta_1 z + \cdots + \theta_q z^q = \prod_{j=1}^q (1 - \omega_j z)\end{aligned}$$

2.4 ARMA Processes

Example (ARMA(1, 1))

Suppose $\{X_t\}$ satisfies

$$(1 - \phi B)X_t = (1 + \theta B)Z_t$$

This is the ARMA(1, 1) process. Suppose that $|\phi| < 1$, and let

$$\alpha(B) = \{\Phi(B)\}^{-1} = \sum_{j=0}^{\infty} \phi^j B^j$$

Then

$$X_t = \alpha(B)\Theta(B)Z_t = \Psi(B)Z_t$$

where

$$\Psi(z) = \alpha(z)\Theta(z) = (1 + \phi z + \phi^2 z^2 + \cdots)(1 + \theta z)$$

2.4 ARMA Processes

Example (ARMA(1, 1))

Therefore

$$\begin{aligned}\psi_0 &= 1 \\ \psi_j &= (\theta + \phi)\phi^{j-1}, j \geq 1.\end{aligned}$$

so

$$X_t = Z_t + (\theta + \phi) \sum_{j=1}^{\infty} \phi^{j-1} Z_{t-j}$$

Note: if $|\phi| > 1$, can construct a stationary non-causal version.

2.4 ARMA Processes

Invertibility: If

$$\Phi(B)X_t = \Theta(B)Z_t$$

then we may also write

$$\{\Theta(B)\}^{-1}\Phi(B)X_t = Z_t$$

provided the expansion

$$\beta(B) = \{\Theta(B)\}^{-1} = \sum_{j=0}^{\infty} \beta_j B^j$$

converges. From above, we have

$$\Theta(z) = \prod_{j=1}^q (1 - \omega_j z)$$

so therefore expansion exists provided $|\omega_j| < 1$ converges. Such a process is termed **invertible**.

2.4 ARMA Processes

Example (ARMA(1, 1))

In this case

$$\begin{aligned}\Theta(B) &= 1 + \theta B \\ \beta(B) &= 1 - \theta B + \theta^2 B^2 - \dots\end{aligned}$$

and we require $|\theta| < 1$. If

$$\pi(B) = \beta(B)\Phi(B)$$

where $\pi_0 = 1$, and

$$\pi_j = -(\theta + \phi)(-\theta)^{j-1}, \quad j \geq 1.$$

We thus have an $\text{AR}(\infty)$ representation

$$\pi(B)X_t = Z_t.$$

2.4 ARMA Processes

Estimation: Estimation for the ARMA processes studied above, and indeed all stationary processes, can be achieved in a non-parametric fashion using moment-based estimation. For example, we may estimate expectation $\mu = \mathbb{E}[X_t]$ using estimator

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_t = \bar{X}$$

Recall that we require an estimator to have good statistical properties. In the non IID case, properties of this estimator are more difficult to study; it is clear that, under stationarity,

$$\mathbb{E}[\bar{X}] = \mu$$

so we have an unbiased estimator, but the variance of this estimator must be computed with care.

2.4 ARMA Processes

$$\begin{aligned}\mathbb{V}ar[\bar{X}] &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \mathbb{C}ov[X_i, X_j] = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \gamma_X(|i-j|) \\ &= \frac{1}{n^2} \sum_{h=-n}^n (n - |h|) \gamma_X(h) \\ &= \frac{1}{n} \sum_{h=-n}^n \left(1 - \frac{|h|}{n}\right) \gamma_X(h)\end{aligned}$$

As $n \rightarrow \infty$, this variance should stay finite, otherwise the performance of the estimator will be poor. A necessary condition is that $\gamma_X(h) \rightarrow 0$ as $h \rightarrow \infty$, but this is not sufficient, for example if $\gamma_X(h) \sim 1/h$ as $h \rightarrow \infty$, the sum diverges.

2.4 ARMA Processes

If $\text{Var}[\bar{X}] \longrightarrow 0$, then $\bar{X} \longrightarrow \mu$ in mean square. If

$$\sum_{h=-\infty}^{\infty} |\gamma_X(h)| < \infty$$

then the asymptotic variance is the probability limit of $n\text{Var}[\bar{X}]$,

$$n\text{Var}[\bar{X}] \xrightarrow{p} \sum_{h=-\infty}^{\infty} \gamma_X(h)$$

Note that it is not necessarily the case that

$$\sum_{h=-\infty}^{\infty} |\gamma_X(h)| < \infty.$$

2.4 ARMA Processes

For a Gaussian series $\{X_t\}$

$$\sqrt{n}(\bar{X} - \mu) \sim \mathcal{N}\left(0, \sum_{h=-n}^n \left(1 - \frac{|h|}{n}\right) \gamma_X(h)\right)$$

whereas for non-Gaussian stationary series this result holds as an asymptotic approximation, provided the autocovariance series is absolutely summable.

In practice, $\gamma_X(h)$ must be estimated from the data, and in the asymptotic case, the infinite sum must be truncated;

$$\sum_{h=-l}^l \left(1 - \frac{|h|}{n}\right) \hat{\gamma}(h)$$

commonly this is done at $l = \sqrt{n}$.

2.4 ARMA Processes

Example (AR(1) process)

Suppose that $\{Z_t\} \sim \text{WN}(0, \sigma^2)$, and

$$X_t - \mu = \phi(X_{t-1} - \mu) + Z_t$$

with $|\phi| < 1$. Then

$$\gamma_X(h) = \frac{\sigma^2 \phi^{|h|}}{1 - \phi^2}$$

so that

$$\sum_{h=-\infty}^{\infty} |\gamma_X(h)| = \sum_{h=-\infty}^{\infty} \gamma_X(h) = \frac{\sigma^2}{1 - \phi^2} \left[1 + 2 \sum_{h=1}^{\infty} \phi^h \right] = \frac{\sigma^2}{(1 - \phi)^2}$$

2.4 ARMA Processes

Estimation of the autocovariance function $\gamma_X(h)$ can also be achieved by moment-based methods

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_{t+h} - \bar{x})(x_t - \bar{x}).$$

This is a biased estimator of $\gamma_X(h)$, but the bias is low (order $o(1/n)$).

Note that the $k \times k$ symmetrix matrix $\hat{\Gamma}_k$ with

$$[\hat{\Gamma}_k]_{ij} = \hat{\gamma}(|i - j|)$$

is non-negative definite for each $k \geq 1$ (see B & D p. 59).

2.4 ARMA Processes

The autocorrelation function $\rho_X(h)$ can be estimated by

$$R_k = \frac{1}{\hat{\gamma}(0)} \hat{\Gamma}_k$$

The statistical properties of this estimator can be established with care (see B & D pp. 60-63) and can be used to construct tests for each h of the hypothesis that $\rho_X(h)$ is zero based on a z-score. Global (or *portmanteau*) tests, that assess multiple h simultaneously can be constructed based on a Chi-squared distribution approximation. See for example the *Box-Ljung* test (`Box.test` in R).

Note: Γ_k , $\hat{\Gamma}_k$ and other related matrices are symmetric *Toeplitz* matrices (that is, they are constant along diagonals).