Model Types

Typical structure of a time series is

$$X_t = m_t + Y_t + S_t$$

where m_t is a trend, S_t is a seasonal part, and Y_t is a stationary part.

- **White Noise:** Sequence of iid random variables $\{Z_t\}$ with mean 0 and variance σ_Z^2.
- **Random Walk:** $X_t = X_{t-1} + Z_t$, where $\{Z_t\}$ is white noise.
- **MA(q):** Let $\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^2$,
$$X_t = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} = \theta(B) Z_t$$
- **AR(p):** Let $\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p$,
$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t \implies Z_t = \phi(B) X_t$$
- **ARMA(p,q):** $\phi(B)X_t = \theta(B)Z_t$
$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

Autocovariance Functions

The autocovariance function of a time series $\{X_t\}$ is

$$\gamma_X(r,s) = \mathrm{Cov}(X_r, X_s) = E[(X_r - \mu_X(r))(X_s - \mu_X(s))]$$

ACF For Different Models

- **White Noise:** $\gamma_Z(t+h,t) = I(h=0)$
- **Random Walk:** $\gamma_S(t, t+h) = t\sigma_Z^2$
- **MA(1):** $\gamma_X(t+h,t) = \begin{cases} \sigma_Z^2(1+\theta^2) & h = 0 \\ \theta\sigma_Z^2 & h = \pm 1 \\ 0 & |h| > 1 \end{cases}$
- **AR(1):** $\gamma_X(h) = \sigma_Z^2 \dfrac{\phi^h}{1-\phi^2}$
- **ARMA(1,1):** $\gamma_X(0) = \sigma_Z^2\left(1 + \dfrac{(\phi+\theta)^2}{1-\phi^2}\right)$,
$$\gamma(1) = \sigma_Z^2\left[(\phi+\theta) + \frac{(\phi+\theta)^2\phi}{1-\phi^2}\right] \; \gamma(h) = \phi^{h-1}\gamma_X(1)$$
- **AR(2):** $\gamma_X(0) = \dfrac{\sigma_Z^2(1-\phi_2)}{(1+\phi^2)[(1-\phi_2)^2 - \phi_1^2]}$,
$\gamma(1) = \dfrac{\gamma(0)\phi_1}{1-\phi_2}$, $\gamma(h) = \phi_1\gamma(h-1) + \phi_2\gamma(h-2)$
- **MA(q):** $\gamma_X(h) = \sigma_Z^2 \sum_{j=0}^{q-h} \theta_j\theta_{j+h}$

ARMA Models

Solutions to ARMA Models

i) A stationary solution for ARMA(p,q) exists whenever the autoregressive polynomial $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0$ for all $|z| = 1$. In other words, for ARMA to be stationary, if z solves $\phi(z) = 0$, then $|z| \neq 1$.

ii) A stationary and causal solution for ARMA(p,q) exists whenever the autoregressive polynomial $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0$ for all $|z| \leq 1$. In other words, for ARMA to be stationary and causal, if z solves $\phi(z) = 0$, then $|z| > 1$.

Linear Processes

Definition Let $\{Z_t\}$ be white noise. Let $\psi_j,\ j \geq 0$, be a sequence of constants such that $\sum_{j=0}^\infty \psi_j^2 < \infty$. Then,
$$X_t = \sum_{j=0}^\infty \psi_j Z_{t-j}$$
is called a linear process.

Linear Representation of AR(1): AR(1) is obtained by setting $p = 1$ and $q = 0$, so $Z_t = \phi(B)X_t$ where $\phi(z) = 1 - \phi z$. Then, we define
$$\chi(z) = \frac{1}{\phi(z)}$$
This function has the power series expansion when $|\phi| < 1$,
$$\chi(z) = \frac{1}{\phi(z)} = \sum_{j=0}^\infty \phi^j z^j$$
Multiplying this with our equation for Z_t,
$$\chi(B)\phi(B)X_t = \chi(B)Z_t \iff X_t = \chi(B)Z_t$$
Thus, our linear representation of AR(1) is
$$X_t = \chi(B)Z_t = \sum_{j=0}^\infty \phi^j B^j Z_t = \sum_{j=0}^\infty \phi^j Z_{t-j}$$

Linear Representation of ARMA(1,1): Here we have $\phi(B)X_t = \theta(B)Z_t$ where $\phi(z) = 1 - \phi z$ and $\theta(z) = 1 + \theta z$. Then define $\chi(z)$ the same way and multiply on both sides to get $X_t = \chi(B)\theta(B)Z_t$.
$$X_t = \sum_{j=0}^\infty \phi^j B^j (1+\theta B) Z_t = \sum_{j=0}^\infty \phi^j Z_{t-j} + \theta \sum_{j=0}^\infty \phi^j Z_{t-(j+1)}$$

We want to write this in the form:
$$\sum_{j=0}^\infty \psi_j Z_{t-j} = \sum_{j=0}^\infty \phi^j Z_{t-j} + \theta \sum_{j=0}^\infty \phi^j Z_{t-j-1}$$
$$\psi_0 Z_t + \sum_{j=1}^\infty \psi_j Z_{t-j} = \phi^0 Z_t + \sum_{j=1}^\infty (\phi^j + \theta\phi^{j-1}) Z_{j-1}$$
Hence $\psi_0 = 1,\ \psi_j = \phi^{j-1}(\theta+\phi),\ j \geq 1$.

Partial Autocorrelation Functions

Finding the partial autocorrelation function between X_1 and X_2 when removing the effect of X_3:

i) First regress X_1 on X_3, compute the regression coefficient β_{13},
$$X_1 = \beta_{13}X_3 + Z$$
where Z has mean zero and is independent of X_3. Then multiply by X_3 and take expected value
$$E(X_1X_3) = \beta_{13}E(X_3^2) + E(ZX_3)$$
$$\gamma_X(2) = \beta_{13}\gamma_X(0) + E(ZX_3) = \beta_{13}\gamma_X(0)$$
This gives us $\beta_{13} = \dfrac{\gamma_X(2)}{\gamma_X(0)} = \rho_X(2)$.

ii) Then similarly, we regress X_2 on X_3,
$$X_2 = \beta_{23}X_3 + V$$
and we get $\beta_{23} = \dfrac{\gamma_X(1)}{\gamma_X(0)} = \rho_X(1)$.

iii) Finally, we define the partial autocorrelation function (PACF) between X_1 and X_2 when removing the effect of X_3 as
$$\begin{aligned}\rho_{12.3} &= \mathrm{Cor}(X_1 - \beta_{13}X_3, X_2 - \beta_{23}X_3) \\ &= \frac{\mathrm{Cov}(X_1 - \beta_{13}X_3, X_2 - \beta_{23}X_3)}{\sqrt{\mathrm{Var}(X_1 - \beta_{13}X_3)}\sqrt{\mathrm{Var}(X_2 - \beta_{23}X_3)}}\end{aligned}$$
We can rewrite the covariance using linearity of covariance as $\gamma_X(1)(1-\rho_X(2))$, and similarly for the variances
$$\begin{aligned}\mathrm{Var}(X_1 - \beta_{13}X_3) &= \mathrm{Cov}(X_1 - \beta_{13}X_3, X_1 - \beta_{13}Y_3) \\ &= \mathrm{Cov}(X_1, X_1) + \mathrm{Cov}(\beta_{13}X_3, \beta_{13}X_3) \\ &\quad - \mathrm{Cov}(X_1, \beta_{13}X_3) - \mathrm{Cov}(\beta_{13}X_3, X_1) \\ &= \gamma_X(0)(1 - \rho_X^2(2)) \\ \mathrm{Var}(X_2 - \beta_{23}X_3) &= \gamma_X(0)(1 - \rho_X^2(1))\end{aligned}$$
Thus, the partial autocorrelation function is
$$\begin{aligned}\rho_{12.3} &= \frac{\gamma_X(1)(1-\rho_X(2))}{\gamma_X(0)\sqrt{(1-\rho_X^2(2))(1-\rho_X^2(1))}} \\ &= \frac{\mathrm{Cor}(X_1,X_2) - \mathrm{Cor}(X_2,X_3)\,\mathrm{Cor}(X_1,X_3)}{\sqrt{(1-\mathrm{Cor}^2(X_1,X_3))(1-\mathrm{Cor}^2(X_2,X_3))}}\end{aligned}$$

Interpreting PACF/ACF Plots

	AR(p)	MA(q)
ACF	Tails off (Geometric Decay)	Significant at lag q Cuts off after lag q
PACF	Significant at each lag p Cuts off after lag p	Tails off (Geometric Decay)

Yule Walker Equations

Denote $P_n X_{n+k}$ as the predicted value of X_{n+k} given that we have n observations.

$$P_n X_{n+k} = a_0 + a_1 X_n + a_2 X_{n-1} + \cdots a_n X_1$$

To find coefficients $a_0, \ldots, a_n$, solve the system

$$\Gamma_n \mathbf{a}_n = \gamma(n;k),\ \ \gamma(n;k) = (\gamma_X(k), \ldots, \gamma_X(k+n-1))$$

$$\Gamma_n = [\gamma_X(i-j)]_{i,j=1}^{n},\ \ \mathbf{a}_n = (a_1, \ldots, a_n)^T$$

The mean squared prediction error is

$$\begin{aligned}\mathrm{MSPE}_n(k) &= E\left[\left(X_{n+k} - \sum_{i=1}^{n} a_i X_{n+1-i}\right)^2\right] \\ &= \gamma(0) - \mathbf{a}_n \gamma(n;k)\end{aligned}$$