Models for time-series

- Auto-regressive models AR(p):

- Moving average models MA

- ARIMA models (Combines the above two)

- SARIMA: Generalization of ARIMA to handle seasonality.

Auto-regressive models

- AR(p) model: The value of x at time t is a linear function of the value of x at time t−1, t-2,…t-p.

$$AR(p): x_t = \delta + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots \phi_p x_{t-p} + w_t$$

- AR(1) $\quad x_t = \delta + \phi_1 x_{t-1} + w_t$

parameter

errors

$$x_t = x_{t-1} + w_t$$
$$\phi_1 = 1$$

- $w_t \overset{iid}{\sim} N(0, \sigma_w^2)$, meaning that the errors are independently distributed with a normal distribution that has mean 0 and constant variance.
- Properties of the errors w_t are independent of x_t.
- The series x_1, x_2, … is (weakly) stationary. A requirement for a stationary AR(1) is that $|\phi_1| < 1$. We'll see why below.

Finding the best values of parameters: $\delta, \phi_1, \dots, \phi_p$

- Standard least square fit. $D: \{x_1, x_2 - - - - - x_n\}$
- Each t defines a training instance

Our goal is to find $\varphi_1, \varphi_2 - - \varphi_p, \eta, \sigma_w^2$ s.t.
we have a good fit on each x_t

$$x_t = \_ + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + - - \varphi_p x_{t-p} + w_t$$

$$w_t \sim N(0, \sigma_w^2)$$

Example 37.1

- The number of disk access for 50 database queries were measured to be: 73, 67, 83, 53, 78, 88, 57, 1, 29, 14, 80, 77, 19, 14, 41, 55, 74, 98, 84, 88, 78, 15, 66, 99, 80, 75, 124, 103, 57, 49, 70, 112, 107, 123, 79, 92, 89, 116, 71, 68, 59, 84, 39, 33, 71, 83, 77, 37, 27, 30.
- For this data: $\sum_{t=2}^{50} x_t = 3313 \quad \sum_{t=2}^{50} x_{t-1} = 3356$

$$\sum_{t=2}^{50} x_t x_{t-1} = 248147 \quad \sum_{t=2}^{50} x_{t-1}^2 = 272102 \quad n = 49$$

$$\eta \quad \cancel{a_0} = \frac{\sum x_t \sum x_{t-1}^2 - \sum x_{t-1} \sum x_t x_{t-1}}{n \sum x_{t-1}^2 - \left(\sum x_{t-1}\right)^2}$$

$$= \frac{3313 \times 272102 - 3356 \times 248147}{49 \times 272102 - 3356^2} = 33.181$$

Example 37.1 (Cont)

$$
\varphi_1 \; \cancel{a_1} = \frac{n\sum x_t x_{t-1} - \sum x_t \sum x_{t-1}}{n \sum x_{t-1}^2 - \left(\sum x_{t-1}\right)^2}
$$

$$
= \frac{49 \times 248147 - 3313 \times 3356}{49 \times 272102 - 3356^2} = 0.503
$$

- The AR(1) model for the series is:

$$
x_t = 33.181 + 0.503 x_{t-1} + \cancel{e_t} \quad w_t
$$

- The predicted value of x_2 given x_1 is:

$$
\hat{x}_2 = a_0 + a_1 x_1 = 33.181 + 0.503 \times 73 = 69.880
$$

- The actual observed value of is 67. Therefore, the prediction error is:

$$
w_2 \; \cancel{e_2} = x_2 - \hat{x}_2 = 67 - 69.880 = -2.880 \quad w_t
$$

- Sum of squared errors SSE = 32995.57

©2017 Raj Jain

Properties of a time-series following AR(1) model

- The (theoretical) **mean** of x_t is

$$E(x_t) = \mu = \frac{\cancel{\delta}\ \eta}{1 - \phi_1}$$

- The **variance** of x_t is

$$\text{Var}(x_t) = \frac{\sigma_w^2}{1 - \phi_1^2}$$

- The **correlation** between observations h time periods apart is

$$\rho_h = \phi_1^h$$

Proofs.

- Easy proofs here

$x_t = \eta + \varphi_1 x_{t-1} + \omega_t$

$E[x_t] = E(\eta) + E(\varphi_1 x_{t-1}] + E(\omega_t)$

$= \eta + \varphi E(x_{t-1}) + 0$

Assume series is stationary $E[x_t] = E(x_{t-1})$

$$E(x_t) = \frac{\eta}{1-\varphi_1}$$

$V(x_t) = E$

$$ACF(1) \quad \frac{E[(x_t-\mu_x)(x_{t-1}-\mu_x)]}{Var(x_t)}$$

$$= E[(\varphi_1 x_{t-1}+\eta+w_t-\mu_x)(x_{t-1}-\mu_x)]$$

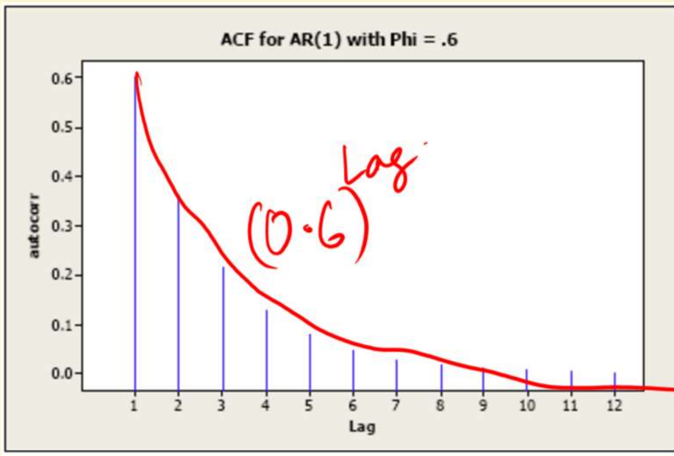
$$= \varphi E[x_{t-1}^2] + \eta E[x_{t-1}) + E[x_{t-1} w_t]$$

Shape of ACF of a series following AR(1) model

Following is the ACF of an AR(1) with $\phi_1 = 0.6$, for the first 12 lags.

Note!

The tapering pattern:

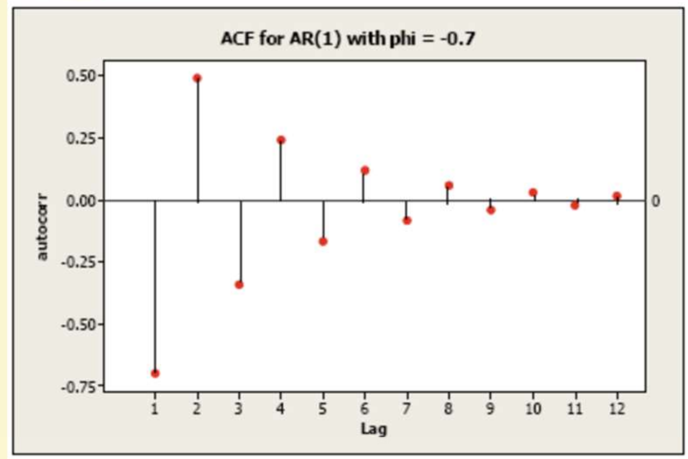


The ACF of an AR(1) with $\phi_1 = -0.7$ follows.

ϕ_1^h $h = \text{lag}$

Note!

The alternating and tapering pattern.



Choosing the p for which AR(p) provides a good fit

Partial Auto Correlation Function (PACF), also called conditional ACF.

- Auto-correlation after adjusting for the intervening values
- Conditional correlation: correlation between x_t and x_{t-h} under known values of x-s in-between them.
- ACF might show that x_t and x_{t-h} are correlated but that might be because they are both correlated with the values in-between. PACF corrects for that. $h = 1$

x_t is correlated with x_{t-1}
x_{t-1} is " " x_{t-2}
$\Rightarrow x_t$ is correlated with x_{t-2}

- The 1st order partial autocorrelation will be defined to equal the 1st order autocorrelation.
- The 2nd order (lag) partial autocorrelation is

$$\frac{\text{Covariance}(x_t, x_{t-2}|x_{t-1})}{\sqrt{\text{Variance}(x_t|x_{t-1})\text{Variance}(x_{t-2}|x_{t-1})}}$$

Computing Partial Auto Correlation Function (PACF)

- To compute PACF between x_t and x_{t-h} fit an AR(h) model. The coefficient ϕ_h is the measure of PACF

$$x_t = \eta + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots \phi_h x_{t-h} + w_t$$

$$\phi_h = PACF(x_t, x_{t-h})$$

But

$$\phi_2 \neq PACF(x_t, x_{t-2}) \text{ for } h > 2$$

ARMA models

- Extending auto-regressive models with smoother noise.

  In AR model for each t, we associate an independent noise w_t

Rice production in Maharashtra

AR(1)

MA terms.

$$x_t = \eta + \phi_1 x_{t-1} + w_t + \theta_1 w_{t-1}$$

$$x_t = \eta + \phi_1 x_{t-1} + \cdots \phi_p x_{t-p} + \boxed{\theta_1 w_{t-1} + \theta_2 w_{t-2} + \cdots \theta_p w_{t-p}} + w_t$$

Need smoother handling of noise.

  A moving average (MA) model provides that.

- ARMA models: AR models + MA models

Moving average models (MA models)

- A value x_t in a time-series sometimes cannot be explained just in terms of its past values.
- External (unknown) variables might be influencing the values
  - Example: Total wheat export of India in 2023 can be determined by wheat export in 2022, but also other external factors like weather patterns, war, exchange rates, etc.
- External variables are also time-varying → errors at each position cannot be independent.
- Moving average models capture dependency on such external unknowns.

Properties of a series following MA(1) model $p = 0, q = 1$

$x_t = \eta + \theta_1 \omega_{t-1} + \omega_t$

$E(x_t) =$

- Mean is $E(x_t) = \mu$
- Variance is $Var(x_t) = \sigma_w^2(1 + \theta_1^2)$
- Autocorrelation function (ACF) is: $E(x_t x_{t-1})$

$$\rho_1 = \frac{\theta_1}{1 + \theta_1^2}, \text{ and } \rho_h = 0 \text{ for } h \geq 2$$

Proofs here: https://online.stat.psu.edu/stat510/lesson/2/2.1#paragraph--264

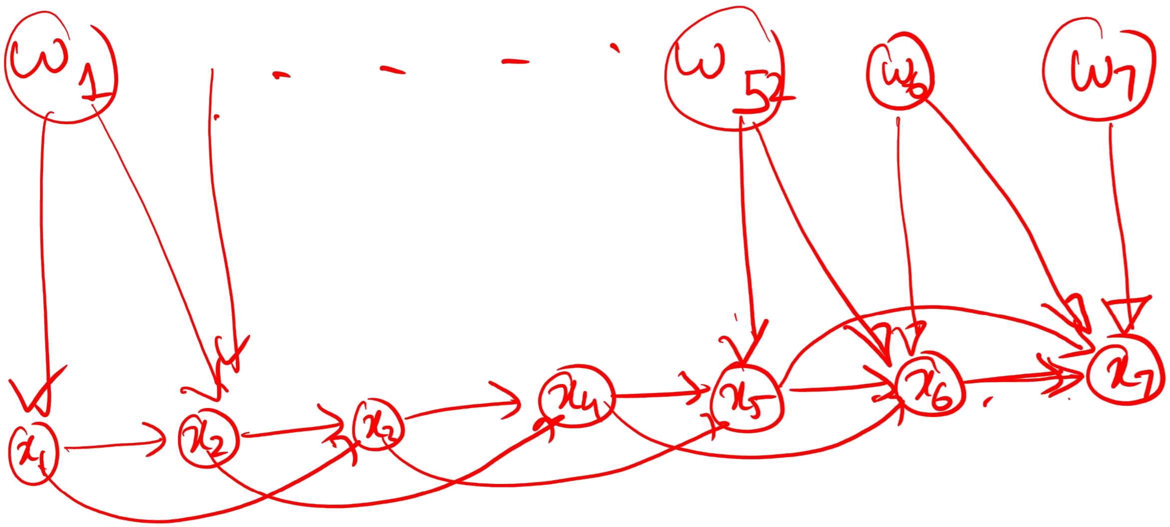
Pictorial representation of dependency.

$$E(x_t x_{t-1}) = E\big((n + \theta_1 w_{t-1} + w_t)\, x_{t-1}\big)$$

$$= n E(x_{t-1}) + \theta_1 E(w_{t-1} \cdot x_{t-1}) + E(w_t x_{t-1})$$

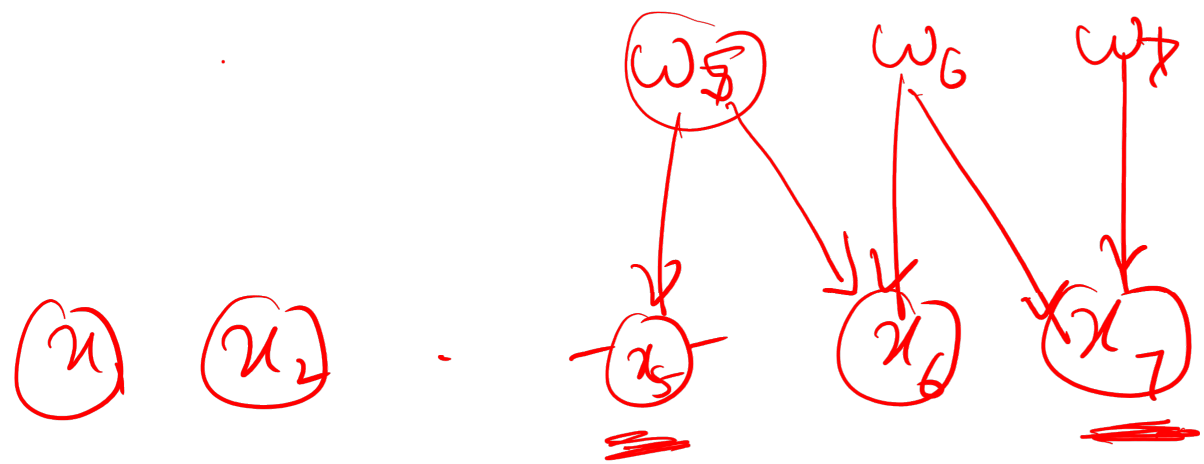
$$=$$

MA( )



AR(2)

$p = 0, \quad q = 1$



ARMA (p,q) model

Each xt depends on p previous x-values, and q-previous error values

$$x_t = \eta + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q}$$

Estimating all the parameters of this model is not as straightforward as least-square regression since the w_t values are not observed (Not covered)