Forecasting with ARMA

Linear Predictor

Given data $\{Y_1, \ldots, Y_n\}$ we want to predict Y_{n+1}.

Let $\hat{Y}_n(1)$ be 1-step predictor given $\{Y_1, \ldots, Y_n\}$.

① We only consider linear predictor

$$\hat{Y}_n(1) = a_0 + a_1 Y_n + \cdots + a_n Y_1.$$

for some $(a_0, \ldots, a_n)$.

② We want predictor that minimize <u>mean squared error</u> of prediction

$$E\left\{\left(\underset{\substack{\uparrow \\ \text{prediction}}}{\hat{Y}_n(1)} - \underset{\substack{\uparrow \\ \text{true}}}{Y_{n+1}}\right)^2\right\} = MSE$$

$\hat{Y}_n(1)$ must minimize MSE.

Consider $\hat{Y}_n(h) = a_0 + a_1 Y_n + \cdots + a_n Y_1$. Then,

$$E\left\{\left(\hat{Y}_n(h) - Y_{n+h}\right)^2\right\} = E\left\{\left(a_0 + a_1 Y_n + \cdots + a_n Y_1 - Y_{n+h}\right)^2\right\}$$

take $\frac{d}{da_0}$ and set to 0 to find min.

$\frac{d}{da_0}$ goes inside $E\{\ \}$, and we get

$$E\left\{2\left(a_0 + a_1 Y_n + \cdots + a_n Y_1 - Y_{n+h}\right)\right\} = 0. \quad \text{—— (1)}.$$

If $E(Y_t) = 0$, then $a_0 = 0$

Take $\frac{d}{da_1}$ and get

$$E\left\{2\left(a_0 + a_1 Y_n + \cdots + a_n Y_1 - Y_{n+h}\right)\cdot Y_n\right\} = 0$$

rewrite as

$$\cancel{a_0} + a_1\gamma(0) + \cdots + a_n\gamma(n-1) - \gamma(h) = 0$$

Take $\frac{d}{da_2}$ and get

$$E\left\{ 2\left(a_0 + a_1 Y_n + \cdots + a_n Y_1 - Y_{n+h} \right) \cdot Y_{n-1} \right\} = 0.$$

i.e.

$$\cancel{a_0} + a_1 \gamma(1) + \cdots a_n \gamma(n-2) - \gamma(h+1) = 0$$

If we keep going, we get ...

$$\frac{d}{da_1}: \quad \cancel{\alpha_0} + a_1\gamma(0) + \cdots + a_n\gamma(n-1) = \gamma(h)$$

$$\frac{d}{da_2}: \quad \cancel{\alpha_0} + a_1\gamma(1) + a_2\gamma(0) + \cdots + a_n\gamma(n-2) = \gamma(h+1)$$

$$\vdots$$

$$\frac{d}{da_n}: \quad \cancel{\alpha_0} + a_1\gamma(n-1) + \cdots + a_n\gamma(0) = \gamma(h+n-1)$$

$$\begin{bmatrix} \gamma(0) & \cdots & \gamma(n-1) \\ \vdots & \ddots & \vdots \\ \gamma(n-1) & \cdots & \gamma(0) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \gamma(h) \\ \vdots \\ \gamma(h+n-1) \end{bmatrix}$$

Therefore, we can calculate $a_1, \ldots a_n$ by eq'n

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \gamma(0) & \cdots & \gamma(n-1) \\ \vdots & \ddots & \vdots \\ \gamma(n-1) & \cdots & \gamma(0) \end{bmatrix}^{-1} \begin{bmatrix} \gamma(h) \\ \vdots \\ \gamma(h+n-1) \end{bmatrix}.$$

$h=1$
$\Rightarrow$ Y-W eq'n

Then,

$\hat{Y}_n(h) = a_1 X_n + \cdots + a_n X_1$ is the linear predictor

with minimum MSE

$$E\left\{\left(\hat{Y}_n(h) - Y_{n+h}\right)^2\right\}.$$

when $E(Y_n) \neq 0$

Suppose $E(Y_n) = \mu$. then $Z_n = Y_n - \mu$ has

$E(Z_n) = 0$. $\hat{Z}_n(h) = a_1 Z_n + \cdots + a_n Z_1$

can be calculated by minimizing MSE

$$E\left(Z_n - \hat{Z}_n(h)\right).$$

Then $\boxed{\hat{Y}_n(h) = \mu + \hat{Z}_n(h)}$

$$E\left\{\left(Y_n - \hat{Y}_n(h)\right)^2\right\}$$

$$= E\left\{\left(Z_n + \mu - \left(\hat{Z}_n(h) + \mu\right)\right)^2\right\}$$

$$= E\left\{\left(Z_n - \hat{Z}_n(h)\right)^2\right\}.$$

minimized.

One - step prediction of AR(1)

$h = 1$, $\qquad \gamma(h) = \frac{\sigma^2}{1-\phi_1^2} \phi_1^h$

$$\begin{bmatrix} \gamma(0) & \cdots & \gamma(n-1) \\ \vdots & \ddots & \vdots \\ \gamma(n-1) & \cdots & \gamma(0) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \gamma(1) \\ \vdots \\ \gamma(n) \end{bmatrix}$$

Y-W eq'n.

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \phi_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \hat{Y}_n(1) = a_1 Y_n = \phi_1 Y_n$$

$$\begin{aligned} MSE &= E(\phi_1 Y_n - Y_{n+1})^2 \\ &= E(\phi_1^2 Y_n^2 - 2\phi_1 Y_n Y_{n+1} + Y_{n+1}^2) \\ &= \phi_1^2 \gamma(0) - 2\phi_1 \gamma(1) + \gamma(0) = \end{aligned}$$

$$= \phi_1^2\left(\frac{\sigma^2}{1-\phi_1^2}\right) - 2\phi_1\left(\frac{\sigma^2}{1-\phi_1^2}\right)\phi_1 + \frac{\sigma^2}{1-\phi_1^2}$$

$$= \left(\frac{\sigma^2}{1-\phi_1^2}\right)\left[\phi_1^2 - 2\phi_1^2 + 1\right]$$

$$= \frac{\sigma^2}{1-\phi_1^2}\left[1-\phi_1^2\right]$$

$$= \sigma^2$$ ← MSE of 1-step Forecasting.

$\hat{Y}_n(1) - Y_n = e_t$ ← 1-step error.

One-step prediction of AR(p)

$$\begin{bmatrix} \gamma(0) & \cdots & \gamma(n-1) \\ \vdots & \ddots & \vdots \\ \gamma(n-1) & \cdots & \gamma(0) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \gamma(1) \\ \vdots \\ \gamma(n) \end{bmatrix}$$

Y-W.

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \phi_1 \\ \vdots \\ \phi_p \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\hat{Y}_n(1) = \boxed{\phi_1 Y_n + \phi_2 Y_{n-1} + \cdots + \phi_p Y_{n-p+1}}$$

Innovations Algorithm.

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \gamma(0) & \cdots & \gamma(n-1) \\ \vdots & \ddots & \vdots \\ \gamma(n-1) & \cdots & \gamma(0) \end{bmatrix}^{-1} \begin{bmatrix} \gamma(h) \\ \vdots \\ \gamma(h+n \end{bmatrix}$$

what do you do when $n = 1000$?

Get inverse of 1000×1000 matrix ?

$\to$ recursive algorithm.

innovations : one-step prediction errors.

$$X_n - \hat{X}_n$$

$n=1$ $\quad X_1 - \hat{X}_1(1) \qquad \hat{X}_1(1) = 0.$

$n=2$ $\quad X_2 - \hat{X}_2(1) \qquad \hat{X}_2(1) = a_{11} X_1$

$n=3$ $\quad X_3 - \hat{X}_3(1) \qquad \hat{X}_3(1) = a_{21} X_2 + a_{22} X_1$

$n=4$ $\quad X_4 - \hat{X}_4(1) \qquad \hat{X}_4(1) = a_{31} X_3 + a_{32} X_2 + a_{33} X_1$

$\vdots$

$$
-\begin{bmatrix}
-1 & & & & & \\
a_{11} & -1 & & & & \\
a_{22} & a_{21} & -1 & & & \\
a_{33} & a_{32} & a_{31} & -1 & & \\
\vdots & & & & -1 & \\
 & & & & & \ddots \\
a_{n-1,n-1} & & & & & -1
\end{bmatrix}
\begin{bmatrix}
x_1 \\ x_2 \\ \vdots \\ x_n
\end{bmatrix}
$$

$$
\underline{U}_n \quad = \quad A_n \quad \underline{x}_n
$$

lower triangular matrix.

$$A_n^{-1} = \begin{bmatrix} 1 & & & & \\ \theta_{11} & 1 & & & \\ \theta_{22} & \theta_{21} & 1 & & \\ & & & \ddots & \\ \theta_{n-1,n-1} & \cdots & \cdots & \theta_{n-1,1} & 1 \end{bmatrix}$$

$$\underline{U}_n = A_n \underline{X}_n = \underline{X}_n - \underline{\hat{X}}_n$$

$$\underline{\hat{X}}_n = \underline{X}_n - \underline{U}_n$$

$$= A_n^{-1} \underline{U}_n - \underline{U}_n = (A_n^{-1} - I)\, \underline{U}_n$$

$$\underline{\hat{X}}_n = \underbrace{\left(A_n^{-1} - \mathbb{I}\right)}_{\Theta} \underline{U}_n$$

$$\begin{bmatrix} \hat{X}_1 \\ \hat{X}_2 \\ \vdots \\ \hat{X}_n \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & & & & \\ \theta_{11} & 0 & & & \\ \theta_{22} & \theta_{21} & 0 & & \\ & & & \ddots & \\ \theta_{n-1,n-1} & & \cdots & \theta_{n-1,1} & 0 \end{bmatrix}}_{\Theta} \underbrace{\begin{bmatrix} X_1 - \hat{X}_1 \\ X_2 - \hat{X}_2 \\ \vdots \\ X_n - \hat{X}_n \end{bmatrix}}_{\underline{U}_n}$$

$$\hat{X}_{n+1} = \begin{cases} 0 & \text{if } n = 0 \\ \sum_{j=1}^{n} \theta_{nj} (X_{n+1-j} - \hat{X}_{n+1-j}) & \text{if } n = 1, 2, \ldots \end{cases}$$

Innovations Algorithm.

→ recursive

→ Can be used for non-stationary series.

→ If X_t is invertible, $\theta_{nj} \rightarrow \theta_j$ $j = 1, \ldots, q$.

MA param.

Durbin - Levinson Algo.

Innovations Algorithm (for stationary series)

$$\nu_0 = \gamma(0)$$

$$\begin{cases} \theta_{n,n-k} = \dfrac{1}{\nu_k}\left[\gamma(n-k) - \displaystyle\sum_{j=0}^{k-1} \theta_{k,k-j}\,\theta_{n,n-j}\,\nu_j\right] \\ k = 0, \dots, n \\ \nu_n = \gamma(0) - \displaystyle\sum_{j=0}^{n-1} \theta_{n,n-j}^2\,\nu_j \end{cases}$$

Compute as ν_0.

θ_{11}, ν_1

θ_{22}, θ_{21}, ν_2

θ_{33}, θ_{32}, θ_{31}, ν_3.

$\vdots$

One-Step predictor for MA(1)

Cryer p.199

$$\hat{Y}_t(1) = -\theta e_t$$

but e_t not observable.

$$e_t = \sum_{j=0}^{\infty} \pi_j Y_{t-j}$$ (invertible representation)

formula for π_j

$$\pi_j = -\sum_{k=1}^{q} \theta_k \pi_{j-k} - \phi_j$$ (ARMA(p,q))

$-\theta$ from Cryer.

For MA(1) $\quad Y_t = e_t - \theta_1 e_{t-1}$.

$$\pi_j = - \cancel{\text{theta}}(-\theta_1)\,\pi_{j-1}.$$

$$\pi_j = \theta_1^j$$

$$e_t = \sum_{j=0}^{\infty} \theta_1^j \, Y_{t-j} \qquad \text{(invertible representation)}$$

So Cryer's formula

$$\hat{Y}_t(1) = -\theta e_t \qquad \text{really means}$$

$$= -\theta_1 \sum_{j=0}^{\infty} \theta_1^{j} \, Y_{t-j}$$

but we only have $\{Y_1, \ldots, Y_n\}$ so

$$\hat{Y}_t(1) = -\sum_{j=0}^{t-1} \theta_1^{j+1} Y_{t-j} \qquad \text{———} \ \star$$

Is this a MSE minimizer?

MSE minimizer $a_1 Y_n + \cdots + a_n Y_1$ should be a solution to eq'n

$$\begin{bmatrix} \gamma(0) & \cdots & \gamma(n-1) \\ \vdots & \ddots & \vdots \\ \gamma(n-1) & \cdots & \gamma(0) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \gamma(1) \\ \vdots \\ \gamma(n) \end{bmatrix}$$

Divide both side by $\gamma(0)$ and get

$$\begin{bmatrix} \rho(0) & \cdots & \rho(n-1) \\ \vdots & \ddots & \vdots \\ \rho(n-1) & \cdots & \rho(0) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \rho(1) \\ \vdots \\ \rho(n) \end{bmatrix}$$

$$\begin{bmatrix} 1 & \frac{-\theta}{1+\theta^2} & 0 & 0 & & \\ \frac{-\theta}{1+\theta^2} & 1 & \frac{-\theta}{1+\theta^2} & 0 & & 0 \\ 0 & \frac{-\theta}{1+\theta^2} & 1 & \frac{-\theta}{1+\theta^2} & & \\ & A & 1 & A & & \\ & & \ddots & \ddots & \ddots & A \\ & 0 & & \ddots & A & 1 \end{bmatrix} \begin{bmatrix} -\theta_1 \\ -\theta_1^2 \\ -\theta_1^3 \\ \vdots \\ -\theta_1^n \end{bmatrix} = \begin{bmatrix} \frac{-\theta_1}{1+\theta_1^2} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$A = -\frac{\theta}{1+\theta^2}$$

1st row

$$1\cdot(-\theta_1) + \left(\frac{-\theta_1}{1+\theta_1^2}\right)\cdot\left(-\theta_1^2\right)$$

$$= \frac{-\theta_1\left(1+\theta_1^2\right) + \theta_1^3}{1+\theta^2} = \frac{-\theta_1}{1+\theta_1^2}$$

2nd row

$$\left(\frac{-\theta_1}{1+\theta_1^2}\right)\cdot(-\theta_1) + (1)\cdot\left(-\theta_1^2\right) + \left(\frac{-\theta_1}{1+\theta_1^2}\right)\left(-\theta_1^3\right)$$

$$= \frac{\theta_1^2}{1+\theta_1^2} + \frac{-\theta_1^2\left(1+\theta_1^2\right)}{1+\theta_1^2} + \frac{\theta_1^4}{1+\theta_1^2}$$

$$= 0.$$

3rd row $= 0$

$\vdots$

n th row $\left(\frac{-\theta_1}{1+\theta_1^2}\right)\cdot\left(-\theta_1^{n-1}\right) + 1\cdot\left(-\theta_1^{n}\right)$

$$= \frac{\theta_1^n}{1+\theta_1^2} + \frac{-\theta_1^n(1+\theta_2^2)}{1+\theta_1^2}$$

$$= \frac{-\theta_1^{n+2}}{1+\theta_1^2} \qquad |\theta_1|<1$$

$$\approx 0 .$$

One - step predictor for MA(1)

$$\hat{Y}_t(1) \doteq - \sum_{j=0}^{t-1} \theta_1^{j+1} Y_{t-j} \quad \circledast.$$

~~for~~

or use innovation algorithm numerically.