VE472 Lecture 10

Jing Liu

UM-SJTU Joint Institute

Summer

- If the random variables which make up $\{Y_t\}$ are uncorrelated, and

$$\mu(t) = 0 \qquad \text{and} \qquad \mathbb{E}\left[Y_t^2\right] = \sigma^2$$

then Y_t is stationary with autocovariance function

$$\gamma(u) = \begin{cases} \sigma^2 & u = 0, \\ 0 & \text{otherwise.} \end{cases}$$

- This type of series is so-called white noise. Again not all series are stationary, however, many are related in some ways to series which are stationary.

- For example,

Simple Linear trend	$Y_t = \beta_0 + \beta_1 t + \varepsilon_t$
Seasonal	$Y_t = \alpha \sin(2\pi\omega t + \phi) + \varepsilon_t$
Integrated	$Y_{t+1} - Y_t = \varepsilon_{t+1}$

not stationary.
mean is changing

where Y_t in each of cases is not stationary, but the component ε_t may be.

Autoregressive model

- If Y_t satisfies

linear combination of previous values + noise

$$Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$$

where ε_t is white noise and ϕ_u are constants, then Y_t is called an

autoregressive series of order p,

denoted by $\mathrm{AR}(p)$.

- Autoregressive series are important because:

1. They have a natural interpretation, that is, the next value observed is a linear combination of a simple function of the most recent observations.
2. It is easy to estimate their parameters. eg: least sqaures estimate
3. They are easy to forecast.

- However, notice $\mathrm{AR}(p)$ for a given p is not necessarily stationary!

- A natural question to ask, which I have delayed discussing, is

  Why do we care whether to model a times series as stationary or not?

- Let me illustrate its importance using $\mathrm{AR}(1)$,

$$Y_t = \phi Y_{t-1} + \varepsilon_t$$

- Because Y_{t-1} and ε_t are uncorrelated since ε_t is white noise, we have

$$\mathrm{Var}[Y_t] = \phi^2 \mathrm{Var}[Y_{t-1}] + \sigma_\varepsilon^2$$

- Suppose Y_t is stationary, then

$$\mathrm{Var}[Y_t] = \mathrm{Var}[Y_{t-1}] = \sigma_Y^2 \implies \sigma_Y^2 = \phi^2\sigma_Y^2 + \sigma_\varepsilon^2 \implies \sigma_Y^2 > \phi^2\sigma_Y^2$$

  mean/variance constant

$$\implies 1 > \phi^2$$

  which means $\mathrm{AR}(1)$ is not stationary if $\phi^2 \geq 1$.

- Using more rigorous formulation, we can rewrite

$$Y_t = \phi Y_{t-1} + \varepsilon_t$$

as the following

$$(1 - \phi\mathcal{L})\, Y_t = \varepsilon_t$$

where $\mathcal{L}$ is the lag operator, which is defined as

$$\underline{\mathcal{L}Y_t = Y_{t-1}}$$

- It is possible to show formally the inverse of the operator defining $\mathrm{AR}(1)$ is

$$(1 - \phi\mathcal{L})^{-1} = \sum_{u=0}^{\infty} \phi^u \mathcal{L}^u$$

which allows us to consider Y_t in terms of ε_t,

$$(1 - \phi\mathcal{L})^{-1}(1 - \phi\mathcal{L})\, Y_t = (1 - \phi\mathcal{L})^{-1} \varepsilon_t \implies Y_t = \sum_{u=0}^{\infty} \phi^u \varepsilon_{t-u}$$

- Recall from basic calculus, not all infinite series are convergent,

$$Y_t = \sum_{u=0}^{\infty} \phi^u \varepsilon_{t-u}$$

something similar holds here, the series is only convergent if

$$|\phi| < 1 \iff 1 - \phi z = 0, \qquad z \in \mathbb{C}$$

has its solutions lie outside the unit circle in the complex plane.

- If $|\phi| > 1$ the above summation does not converge,

$$Y_t = \sum_{u=0}^{\infty} \phi^u \varepsilon_{t-u}$$

in other words, no such Y_t satisfies

$$Y_t = \phi Y_{t-1} + \varepsilon_t$$

- However, if we rearrange

$$Y_t = \phi Y_{t-1} + \varepsilon_t \iff Y_{t+1} = \phi Y_t + \varepsilon_{t+1} \iff Y_t = \frac{1}{\phi} Y_{t+1} - \frac{1}{\phi} \varepsilon_{t+1}$$

- Then a similar argument to that above produces the representation

$$Y_t = -\sum_{u=0}^{\infty} \phi^{-u} \varepsilon_{t+u}$$

which is convergent when $|\phi| > 1$, but of course it is not useful in practice.

- If $|\phi| = 1$, then there is no stationary solution whatsoever for $\mathrm{AR}(1)$.
- In general, the $\mathrm{AR}(p)$ model

$$Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$$

generate a stationary solution that is useful in terms of prediction if

$$1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p = 0, \qquad z \in \mathbb{C}$$

has its solutions lie outside the unit circle.

Moving Average Model

- If Y_t satisfies

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

where ε_t is white noise and θ_u are constants, then Y_t is called an

moving average series of order q,

denote by $\mathrm{MA}(q)$.

- No additional conditions are required to ensure stationarity for $\mathrm{MA}(q)$.
- However, we usually require the following equation

$$1 + \theta_1 z + \cdots + \theta_q z^q = 0, \qquad z \in \mathbb{C}$$

has its solutions lie outside the unit circle.

garantee unique set of theta

- This requirement is due to unidentifiable issue, e.g., for $\mathrm{MA}(1)$

$$\gamma(u) = \begin{cases} (1+\theta_1^2)\sigma^2 & u = 0 \\ \theta_1\sigma^2 & u = 1 \\ 0 & \text{otherwise} \end{cases}$$

- Notice if we replace θ_1 by $1/\theta_1$ and σ^2 by $\theta_1^2\sigma^2$ the autocovariance function is unchanged. So there are two sets of parameters for the same model.
- Thus it is impossible to obtain a unique solution without the requirement.
- Formally, without the additional requirement the operator for $\mathrm{MA}(q)$

$$1 + \theta_1\mathcal{L} + \cdots + \theta_q\mathcal{L}^q$$

is not invertible, to avoid which we require the following

$$1 + \theta_1 z + \cdots + \theta_q z^q = 0, \qquad z \in \mathbb{C}$$

has its solutions lie outside the unit circle.

Autoregressive Moving Average Model

- If Y_t satisfies

$$Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

  where ε_t is white noise, and ϕ_u and θ_v are constants, then Y_t is called an

  autoregressive moving average series of order (p, q),

  denote by $\mathrm{ARMA}(p, q)$.

- An $\mathrm{ARMA}(p, q)$ series is stationary if the roots of the following

$$1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p = 0, \qquad z \in \mathbb{C}$$

  has its solutions lie outside the unit circle.

Q: What is the mean of a stationary $\mathrm{ARMA}(p, q)$ series?

$$\underbrace{Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p}}_{\text{mean = 0 (stationary)}} + \underbrace{\varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}}_{\text{mean = 0}}$$

- When a series has a non-zero constant mean

$$\mu$$

then we can model the deviation from the mean using $\mathrm{ARMA}(p, q)$ series

$$\boxed{Y_t - \mu} = \phi_1 (Y_{t-1} - \mu) + \cdots + \phi_p (Y_{t-p} - \mu) + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

- Alternatively, the model can be adjusted by introducing a constant directly,

$$Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

- Let $f(t)$ be a function of t and Z_t is a stationary $\mathrm{ARMA}(p, q)$ series, then

$$\boxed{Y_t = f(t) + Z_t}$$

can be modelled by first detrending.

- A random walk is defined by the equation

$$Y_t = Y_{t-1} + \varepsilon_t$$

where ε_t is white noise.

Q: Have you seen a random walk before? AR

- The following can be considered as an integrated moving average model

$$Y_t = Y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$$

which is also an $\mathrm{ARMA}(1,1)$ model, but again it is non-stationary.

- Both models above can be transformed to stationarity by differencing

$$\nabla Y_t = Y_t - Y_{t-1}$$

- In the first case, the difference follows the white noise model

$$\nabla Y_t = \varepsilon_t$$

- In the second, the difference follows the $\mathrm{MA}(1)$ model

$$\nabla Y_t = \varepsilon_t + \theta\varepsilon_{t-1}$$

The effect of Differencing I

- Suppose that Y_t has a linear trend,

$$Y_t = \beta_0 + \beta_1 t + Z_t$$

where Z_t is stationary with $\mathbb{E}\left[Z_t\right] = 0$, then differencing removes the trend,

$$\nabla Y_t = \beta_1 + \nabla Z_t$$

- Now suppose Y_t has quadratic trend,

$$\begin{aligned} Y_t &= \beta_0 + \beta_1 t + Z_t \\ \implies \nabla Y_t &= \left(\beta_0 + \beta_1 t + \beta_2 t^2\right) - \left(\beta_0 + \beta_1 \left(t-1\right) + \beta_2 t \left(t-1\right)^2\right) + \nabla Z_t \\ &= \left(\beta_1 - \beta_2\right) + 2\beta_2 t + \nabla Z_t \\ \nabla^2 Y &= 2\beta_2 + \nabla^2 Z_t \end{aligned}$$

- A polynomial trend of order k can be removed by the operator ∇^k in general.

The effect of Differencing II

- Now let us consider the case of a stochastic trend, that is

$$Y_t = M_t + \varepsilon_t$$

  where M_t is a random process which vary slowing but randomly over time.

- For example, we can assume M_t is generated by random walk model,

$$M_t = M_{t-1} + \eta_t$$

  where η_t and ε_t are independent, then differencing also gives stationarity

$$\nabla Y_t = \eta_t + \varepsilon_t - \varepsilon_{t-1}$$

  which is very much like a $\mathrm{MA}(1)$ model but with additional white noise.

Integrated autoregressive moving average model

- If $W_t = \nabla^d Y_t$ is $\mathrm{ARMA}(p, q)$ series then Y_t is said to be

  integrated autoregressive moving average series of order (p, d, q),

  denoted by $\mathrm{ARIMA}\,(p, d, q)$. In terms of operator formulation, we have

$$\phi\left(\mathcal{L}\right)\nabla^d Y_t = \theta\left(\mathcal{L}\right)\varepsilon_t$$

  where ϕ and θ above denote the polynomial operator of the lag operator $\mathcal{L}$,

$$\begin{aligned}\phi\left(\mathcal{L}\right) &= 1 - \phi_1\mathcal{L} - \cdots - \phi_p\mathcal{L}^p \\ \theta\left(\mathcal{L}\right) &= 1 + \theta_1\mathcal{L} + \cdots + \theta_q\mathcal{L}^q\end{aligned}$$

- If the same requirement for $\mathrm{MA}(q)$ is used, then the inverse of $\theta\left(\mathcal{L}\right)$ exists

$$\theta\left(\mathcal{L}\right)^{-1}\phi\left(\mathcal{L}\right)\nabla^d Y_t = \varepsilon_t$$

  and the three operators collectively will bring Y_t to white noise.

- Given the inverse of $\theta(\mathcal{L})$ exists, we will be able to rewrite

$$Y_t = \sum_{u=1}^{\infty} \pi_u Y_{t-u} + \varepsilon_t$$

where π_u are functions of ϕ_u and θ_v.

- This is very much like the existence of inverse operator in $\mathrm{AR}(1)$

$$Y_t = \phi Y_{t-1} + \varepsilon_t \iff (1-\phi\mathcal{L}) Y_t = \varepsilon_t$$

and power series representation of the inverse

$$(1-\phi\mathcal{L})^{-1} = \sum_{u=0}^{\infty} \phi^u \mathcal{L}^u$$

ensures the following representation of Y_t,

$$(1-\phi\mathcal{L})^{-1}(1-\phi\mathcal{L}) Y_t = (1-\phi\mathcal{L})^{-1} \varepsilon_t \implies Y_t = \sum_{u=0}^{\infty} \phi^u \varepsilon_{t-u}$$

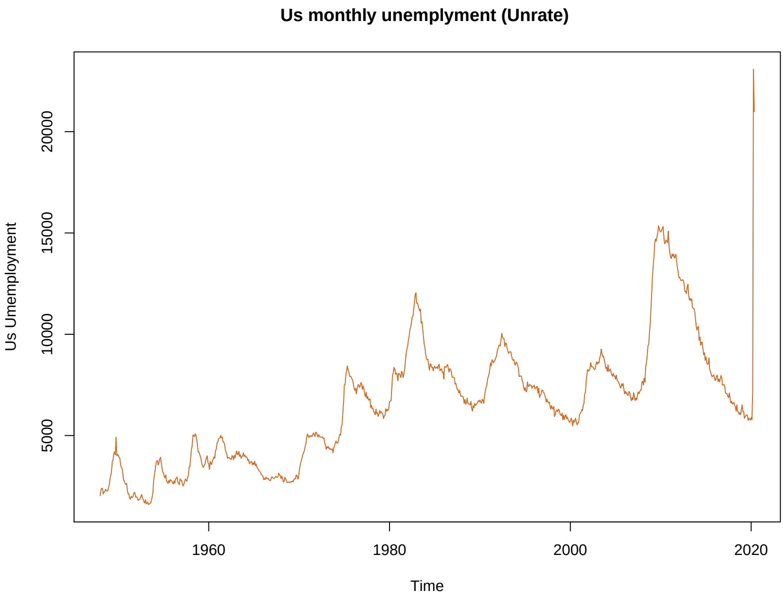
- The $\mathrm{ARIMA}(p, d, q)$ is usually constructed using Maximum likelihood by assuming the white noise has a normal distribution.
- Other methods of estimating the parameters include
  - Method of moments,
  - Nonlinear least squares
  - Kalman filtering (Bayesian)

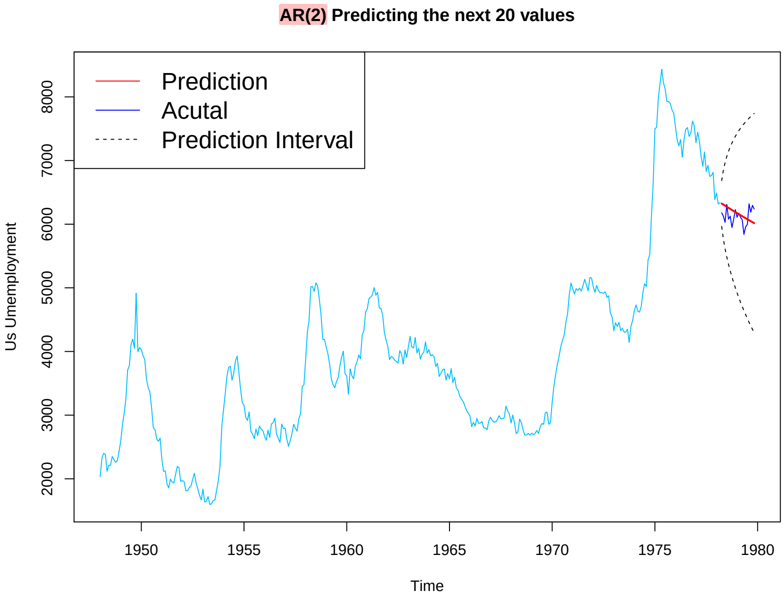
  none of which is particularly quick in comparison to QR in linear model!

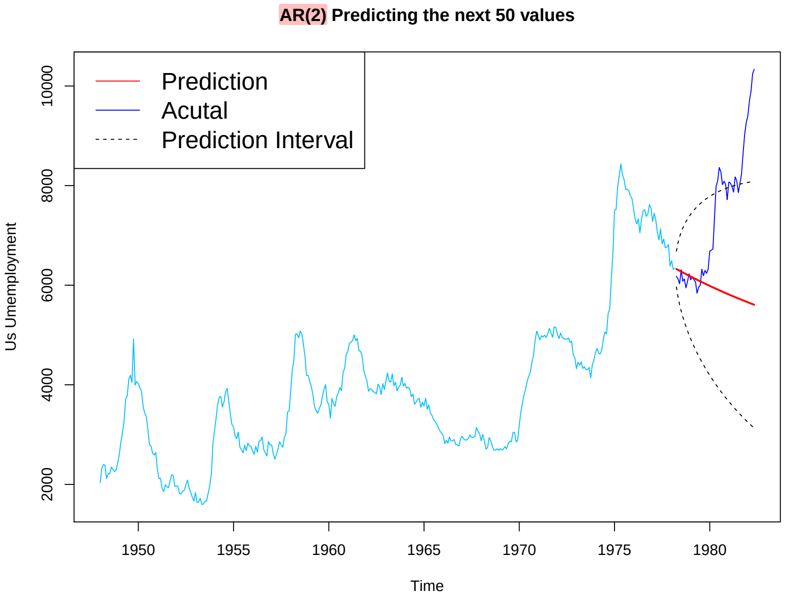
```
> # R uses kalman filtering in arima and predict
> z.arima = arima(umemp.ts, order = c(2, 0, 0))
> (pred = predict(z.arima, n.ahead = 4))
```

```
$pred
          Apr      May      Jun      Jul
1978 6324.327 6307.630 6290.516 6273.454

$se
          Apr      May      Jun      Jul
1978 180.3189 271.1904 339.4592 395.5865
```


Us monthly unemplyment (Unrate)
Us Umemployment
5000
10000
15000
20000
1960
1980
2000
2020
Time


AR(2) Predicting the next 20 values
Prediction
Acutal
Prediction Interval
Us Umemployment
2000
3000
4000
5000
6000
7000
8000
1950
1955
1960
1965
1970
1975
1980
Time



AR(2) Predicting the next 50 values
Prediction
Acutal
Prediction Interval
Us Umemployment
Time
2000
4000
6000
8000
10000
1950
1955
1960
1965
1970
1975
1980