MA(1)

$$y_t = \mu + \epsilon_t + \theta\epsilon_{t-1}$$

Last time, we computed the mean $\mathrm{E}(y_t)$, $\mathrm{Var}(y_t)$ and $Corr\left(y_t, y_{t-k}\right)$ for $k > 0$.

The forecasting formula is the conditional expectation.

$$\begin{aligned} E\left(y_{t+1}|y_t, \epsilon_t, \epsilon_{t-1}\right) &= E\left(\mu + \epsilon_{t+1} + \theta\epsilon_t|y_t, \epsilon_t, \epsilon_{t-1}\right) \\ &= \mu + \theta\epsilon_t \\ E\left(y_{t+2}|y_t, \epsilon_t, \epsilon_{t-1}\right) &= E\left(\mu + \epsilon_{t+2} + \theta\epsilon_{t+1}|y_t, \epsilon_t, \epsilon_{t-1}\right) = \mu \end{aligned}$$

The impulse response function tells us how y_t responds to a one-time shock. Shock today's ϵ and then shut them down forever. Trace out the response of y.Let's say the sequence is ...,$\epsilon_{t-1} = 0, \epsilon_t = 1, \epsilon_{t+1} = 0, ...$

$$\begin{aligned} y_t &= \mu + 1 \\ y_{t+1} &= \mu + \theta \\ y_{t+2} &= \mu \\ y_{t+3} &= \mu \end{aligned}$$

How to estimate? We observe only the $y's$. Can't run a regression.

$$y_t = \mu + \epsilon_t + \theta\epsilon_{t-1}$$

We do a think called maximum likelihood estimation. Assume a pdf for the $\epsilon's$. Then express the $\epsilon's$ in terms of the model (becomes a function of the $y's$ or the data). This transforms the pdf into the likelihood function. Ask computer to choose the parameters $\left(\mu, \theta, \sigma_\epsilon^2\right)$ to maximize the likelihood function.

Assume $\epsilon's$ are normal iid.

$$f_1\left(\epsilon_t\right) = \frac{1}{\sigma_\epsilon\sqrt{2\pi}} e^{-\frac{\epsilon_t^2}{2\sigma_\epsilon^2}}$$

The joint distribution is the product of these $f_1's$ because of independence. The joint pdf of the $\epsilon's$

$$\begin{aligned} f\left(\epsilon_T, \epsilon_{T-1}, ..., \epsilon_1\right) &= f_1\left(\epsilon_T\right) f_1\left(\epsilon_{T-1}\right) \cdots f_1\left(\epsilon_1\right) \\ &= \left(\frac{1}{\sigma_\epsilon\sqrt{2\pi}}\right)^T e^{-\frac{1}{2\sigma_\epsilon^2}\sum_{t=1}^T \epsilon_t^2} \end{aligned}$$

Write the $\epsilon's$ in terms of the $y's$

$$\begin{aligned} \epsilon_t &= y_t - \mu - \theta\epsilon_{t-1} \\ &= y_t - \mu - \theta\left[y_{t-1} - \mu - \theta\epsilon_{t-2}\right] \\ &= y_t - \mu - \theta\left[y_{t-1} - \mu - \theta\left(y_{t-2} - \mu - \theta\epsilon_{t-3}\right)\right] \end{aligned}$$

Maybe this is more clear

$$\begin{aligned}\epsilon_1 &= y_1 - \mu \\ \epsilon_2 &= y_2 - \mu - \theta\left(y_1 - \mu\right) \\ \epsilon_3 &= y_3 - \mu - \theta\left(y_2 - \mu - \theta\left(y_1 - \mu\right)\right)\end{aligned}$$

Substitute this back into the joint pdf

$$f\left(y_T, y_{T-1}, ..., y_1 | \mu, \theta, \sigma_\epsilon^2\right)$$

Now it's a function of the data and we call it the likelihood function. Computer searches for $\mu, \theta, \sigma_\epsilon^2$ to find the maximum. Well, not quite. The likelihood is nonlinear.

$$f\left(y_T, y_{T-1}, ..., y_1 | \mu, \theta, \sigma_\epsilon^2\right) = \left(\frac{1}{\sigma_\epsilon \sqrt{2\pi}}\right)^T e^{-\frac{1}{2\sigma_\epsilon^2}\sum_{t=1}^T \epsilon_t(y_t, y_{t-1}, ... y_1)^2}$$

Take the log of the likelihood. It's called the log-likelihood function.

$$\begin{aligned}\ln\left(f\left(\right)\right) &= -T\ln\left[\left(\sigma_\epsilon^2\right)^{\frac{1}{2}}\right] - T\ln\left[\left(2\pi\right)^{\frac{1}{2}}\right] \\ &\quad - \frac{1}{2\sigma_\epsilon^2}\sum_{t=1}^T \epsilon_t\left(y_t, y_{t-1}, ... y_1\right)^2\end{aligned}$$

The choices of $\mu, \theta, \sigma_\epsilon^2$ that maximize the log likelihood also maximize the likelihood. What if we divide the log likelihood by T? Also, put hats on the $\sigma_\epsilon^{2\prime}s$ to signify that they are functions of the data.

$$\begin{aligned}\frac{\ln\left(f\left(\right)\right)}{T} &= -\ln\left[\left(\hat{\sigma}_\epsilon^2\right)^{\frac{1}{2}}\right] - \ln\left[\left(2\pi\right)^{\frac{1}{2}}\right] \\ &\quad - \frac{1}{2\hat{\sigma}_\epsilon^2}\left[\frac{1}{T}\sum_{t=1}^T \epsilon_t\left(y_t, y_{t-1}, ... y_1\right)^2\right] \\ &= -\ln\left[\left(\hat{\sigma}_\epsilon^2\right)^{\frac{1}{2}}\right] \underbrace{- \ln\left[\left(2\pi\right)^{\frac{1}{2}}\right] - \frac{1}{2}}_{\text{Irrelevant}}\end{aligned}$$

Write the part of the Log likelihood function that matters as LL

$$\frac{LL}{T} = -\ln\left[\left(\hat{\sigma}_\epsilon^2\right)^{\frac{1}{2}}\right]$$

MA(2) model

$$y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}$$

I showed you how to get $E\left(y_t\right)$,Var(y_t),corr(y_t, y_{t-k}) for MA(1). Mimic the steps and you can do it for MA(2). But for now, I tell you a joke.