Note: For a general p, $(i,j)^{th}$ element of V_p^{-1}, $v^{ij}(p)$ is given by (result is due to Galbraith, 1974, Jr of Applied Probability paper):

$$v^{ij}(p) = \left(\sum_{k=0}^{i-1} \phi_k \phi_{k+j-i} - \sum_{k=p+1-j}^{p+i-j} \phi_k \phi_{k+j-i} \right)$$

$1 \leq i \leq j \leq p$ with $\phi_0 = -1$ above

The above is called the Galbraith's formula and can be used to write $L(\underset{\sim}{\theta})$ or $l(\underset{\sim}{\theta})$ explicitly in terms of ϕ_i's.

Conditional MLE formulation : AR(p)

Regard the values of the first p observations as deterministic and maximize the likelihood conditional on the first p observations.

Now, realize that the joint conditional p.d.f of $X_n, \ldots, X_{p+1}$ given $X_p, \ldots X_1$ is given by:

$$f_{X_n, \ldots X_{p+1} | X_p, \ldots X_1} = f_{X_n | X_{n-1} \ldots X_1} f_{X_{n-1}|X_{n-2}, \ldots X_1} \cdots f_{X_{p+1}|X_p, \ldots, X_1}$$

$$= f_{X_n | X_{n-1}, \ldots X_{n-p}} f_{X_{n-1}|X_{n-2}, \ldots X_{n-1-p}} \cdots f_{X_{p+1}|X_p, \ldots, X_1}$$

i.e. $f_{X_n,\ldots X_{p+1}|X_p,\ldots,X_1} = \prod_{t=p+1}^{n} f_{X_t|X_{t-1},\ldots,X_{t-p}}$

$\forall\, t > p$, we have

$$X_t | X_{t-1},\ldots,X_{t-p} \equiv X_t | X_{t-1},\ldots X_1$$

$$\sim N(c + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p}, \sigma^2)$$

Hence conditional log likelihood

$$l_c(\underset{\sim}{\theta}) = -\frac{n-p}{2}\log 2\pi - \frac{n-p}{2}\log \sigma^2 - \frac{1}{2\sigma^2}\sum_{t=p+1}^{n}\left(x_t - c - \sum_{i=1}^{p}\phi_i x_{t-i}\right)^2 \quad - *$$

Note that maximization of $(*)$ w.r.t. $(c, \phi_1, \ldots, \phi_p)$ is equivalent to ~~maximization of~~ minimization of

$$\sum_{t=p+1}^{n}\left(x_t - c - \sum_{i=1}^{p}\phi_i x_{t-i}\right)^2 \quad \text{w.r.t } (c, \phi_1, \ldots \phi_p)$$

$\Rightarrow$ CMLEs of $(c, \phi_1, \ldots, \phi_p)$ are same as the OLS estimates.

<u>Note</u>: CMLE & EMLE have the same asymptotic distⁿ.

<u>Remark</u>: Estimation of AR parameters using Yule-Walker eqn

$$X_t = c + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \epsilon_t$$

Yule-Walker eqn

$$\gamma_h = \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2} + \cdots + \phi_p \gamma_{h-p} \; ; \; h > 0$$

Using p Yule-Walker eqns and estimated $\hat{\gamma}_i$s one can obtain Y-W eqn based estimates of AR model parameters.

e.g. consider an AR(3) setup

$$\hat{\gamma}_1 = \phi_1 \hat{\gamma}_0 + \phi_2 \hat{\gamma}_1 + \phi_3 \hat{\gamma}_2$$

$$\hat{\gamma}_2 = \phi_1 \hat{\gamma}_1 + \phi_2 \hat{\gamma}_0 + \phi_3 \hat{\gamma}_1$$

$$\hat{\gamma}_3 = \phi_1 \hat{\gamma}_2 + \phi_2 \hat{\gamma}_1 + \phi_3 \hat{\gamma}_0$$

i.e.
$$\begin{pmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \hat{\phi}_3 \end{pmatrix} = \begin{pmatrix} \hat{\gamma}_0 & \hat{\gamma}_1 & \hat{\gamma}_2 \\ & \hat{\gamma}_0 & \hat{\gamma}_1 \\ & & \hat{\gamma}_0 \end{pmatrix}^{-1} \begin{pmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \\ \hat{\gamma}_3 \end{pmatrix}.$$

Maximum likelihood estimation for MA models

MA(1): Conditional MLE formulation

$$X_t = \mu + \epsilon_t + \theta \epsilon_{t-1}; \quad \epsilon_t \overset{i.i.d.}{\sim} N(0, \sigma^2)$$

$$\underset{\sim}{\eta} = (\mu, \theta, \sigma^2)' \leftarrow \text{parameter vector}$$

Note that $X_t | \epsilon_{t-1} \sim N(\mu + \theta \epsilon_{t-1}, \sigma^2)$

Suppose, we assume that $\epsilon_0 = 0$ (it's expected value) is given, then

$$X_1 | \epsilon_0 \sim N(\mu, \sigma^2)$$

$\epsilon_1 = X_1 - \mu - \theta \epsilon_0$; hence ϵ_1 given $X_1 = x_1$ & $\epsilon_0 = 0$ is

$\epsilon_1 = x_1 - \mu$

$$X_2 | X_1, \epsilon_0 = 0 \sim N(\mu + \theta \epsilon_1, \sigma^2)$$
$$\text{i.e. } N(\mu + \theta(x - \mu), \sigma^2)$$

$$X_3 | X_2, X_1, \epsilon_0 \sim N(\mu + \theta \epsilon_2, \sigma^2)$$

ϵ_2 given $X_2, X_1, \epsilon_0 = 0$ is $\epsilon_2 = x_2 - \mu - \theta \epsilon_1$

i.e. $\epsilon_2 = x_2 - \mu - \theta(x_1 - \mu)$

i.e. $X_3 | X_2, X_1, \epsilon_0 \sim N\left(\mu + \theta\left((x_2 - \mu) - \theta(x_1 - \mu)\right), \sigma^2\right)$

Thus given $\epsilon_0 = 0$, the full sequence $\epsilon_1, \ldots \epsilon_n$ can be expressed in terms of $(x_1, \ldots, x_n)$, μ & θ through the relationship

$$\epsilon_t = x_t - \mu - \theta \epsilon_{t-1}$$
$$\epsilon_{t-1} = x_{t-1} - \mu - \theta \epsilon_{t-2} \; \ldots$$

$\forall\, t \geq 2;\; X_t | X_{t-1}, \ldots X_1, \epsilon_0 = 0 \equiv X_t | \epsilon_{t-1} \sim N(\mu + \theta \epsilon_{t-1}, \sigma^2)$

Conditional likelihood f^n, conditioned on $\epsilon_0 = 0$, is

$$L(\underset{\sim}{\eta}) = f_{X_n, \ldots, X_1 | \epsilon_0 = 0}(x_n, \ldots, x_1; \underset{\sim}{\eta} | \epsilon_0 = 0)$$

$$= f_{X_n | X_{n-1}, \ldots X_1, \epsilon_0 = 0}\; f_{X_{n-1}, \ldots, X_1 | \epsilon_0 = 0}$$

$$= f_{X_n | X_{n-1} \ldots X_1, \epsilon_0 = 0}\; f_{X_{n-1} | X_{n-2}, \ldots X_1, \epsilon_0 = 0}\; f_{X_{n-2} \ldots X_1 | \epsilon_0 = 0}$$

$$= f_{X_1 | \epsilon_0 = 0} \prod_{t=2}^{n} f_{X_t | X_{t-1}, \ldots, X_1, \epsilon_0 = 0}$$

$$= f_{X_1 | \epsilon_0 = 0} \prod_{t=2}^{n} f_{X_t | \epsilon_{t-1}}$$