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

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

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

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

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

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

Thus, the conditional log likelihood f^n is

$$l(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_1^n (x_t - \mu - \theta\epsilon_{t-1})^2$$

$$l(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_1^n \epsilon_t^2$$

Note that
$$\epsilon_t = (x_t - \mu) - \theta\epsilon_{t-1}$$

$$\epsilon_t = (x_t - \mu) - \theta\left((x_{t-1} - \mu) - \theta\epsilon_{t-2}\right)$$

$$= (x_t - \mu) - \theta(x_{t-1} - \mu) + \theta^2\epsilon_{t-2}$$

$$= (x_t - \mu) - \theta(x_{t-1} - \mu) + \theta^2(x_{t-2} - \mu - \theta\epsilon_{t-3})$$

$$= (x_t - \mu) - \theta(x_{t-1} - \mu) + \theta^2(x_{t-2} - \mu) - \theta^3\epsilon_{t-3}$$

$$\vdots$$

$$\epsilon_t = (x_t - \mu) - \theta(x_{t-1} - \mu) + \theta^2(x_{t-2} - \mu) - \cdots$$
$$\cdots + (-\theta)^{t-1}(x_1 - \mu) + (-\theta)^t \epsilon_0$$

If $\epsilon_0 = 0$ at the initialization,

$$\epsilon_t = \sum_{i=1}^{t} (x_i - \mu)(-\theta)^{t-i}$$ and $l(\underset{\sim}{\eta})$ is

$$l(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log \sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\left(\sum_{i=1}^{t}(x_i - \mu)(-\theta)^{t-i}\right)^2$$

$$\hat{\underset{\sim}{\eta}}_{CMLE} = \underset{\underset{\sim}{\eta}}{\arg\max}\ l(\underset{\sim}{\eta})$$

CMLEs are obtained using some iterative procedure.

Exact MLE formulation

Let's look at the multivariate formulation.

$\underset{\sim}{X} = (X_1, \ldots, X_n)'$ realization from an n-dimensional multivariate normal

i.e. $\underset{\sim}{X} \sim N_n(\mu \underset{\sim}{1}_n, \Omega)$

$$\Omega = E(\underset{\sim}{X} - \mu\underset{\sim}{1}_n)(\underset{\sim}{X} - \mu\underset{\sim}{1}_n)'$$

$$= \sigma^2 \begin{pmatrix} (1+\theta^2) & \theta & 0 & 0 & \cdots & 0 \\ \theta & (1+\theta^2) & \theta & 0 & \cdots & 0 \\ & \ddots & \ddots & \ddots & & \\ & & & & \theta & \theta \\ 0 & & & \cdots & \theta & (1+\theta^2) \end{pmatrix}$$

Likelihood f^n

$$L(\underset{\sim}{\eta}) = (2\pi)^{-n/2} |\Omega|^{-1/2} \exp\left(-\frac{1}{2}(\underset{\sim}{x} - \mu \underset{\sim}{1}_n)' \Omega^{-1} (\underset{\sim}{x} - \mu \underset{\sim}{1}_n)\right)$$

log likelihood f^n

$$l(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{1}{2}\log|\Omega| - \frac{1}{2}(\underset{\sim}{x} - \mu \underset{\sim}{1}_n)' \Omega^{-1} (\underset{\sim}{x} - \mu \underset{\sim}{1}_n)$$

Consider a factorization of Ω as

$$\Omega = A D A' \quad \text{---} \quad (*)$$

$$A = \begin{pmatrix} 1 & & & & \\ \frac{\theta}{1+\theta^2} & 1 & & \bigcirc & \\ 0 & \frac{\theta(1+\theta^2)}{1+\theta^2+\theta^4} & 1 & & \\ & & \ddots & \ddots & \\ & & 0 & \underline{\quad} & 1 \end{pmatrix}$$

$$\uparrow \left(\frac{\theta(1+\theta^2+\cdots+\theta^{2(n-2)})}{1+\theta^2+\cdots+\theta^{2(n-1)}}\right)$$

$$D = \sigma^2 \operatorname{diag}\left(1+\theta^2, \frac{1+\theta^2+\theta^4}{1+\theta^2}, \frac{1+\theta^2+\theta^4+\theta^6}{1+\theta^2+\theta^4}, \ldots\ldots, \frac{1+\theta^2+\theta^4+\cdots+\theta^{2n}}{1+\theta^2+\cdots+\theta^{2(n-1)}}\right)$$

Using $(*)$ likelihood f^n is

$$L(\underset{\sim}{\eta}) = (2\pi)^{-n/2} |ADA'|^{-1/2} \exp\left(-\frac{1}{2}(\underset{\sim}{x} - \mu \underset{\sim}{1}_n)'(A')^{-1} D^{-1} A^{-1} (\underset{\sim}{x} - \mu \underset{\sim}{1}_n)\right)$$

Note that $|\Omega| = |ADA'| = |A||D||A'| = |D| = \prod_{t=1}^{n} d_{tt}$

$$d_{tt} = \sigma^2 \frac{1+\theta^2+\cdots+\theta^{2t}}{1+\theta^2+\cdots+\theta^{2(t-1)}}$$

Let $\underset{\sim}{x}^0 = A^{-1}(\underset{\sim}{x} - \mu \underset{\sim}{1}_n)$

i.e. $A \underset{\sim}{x}_0 = \underset{\sim}{x} - \underset{\sim}{\mu}$

$$x_1^0 = x_1 - \mu$$

$$\vdots$$

$$x_t^0 = (x_t - \mu) - \frac{\theta(1+\theta^2+\cdots+\theta^{2(t-2)})}{1+\theta^2+\cdots+\theta^{2(t-1)}} x_{t-1}^0$$

$$L(\underset{\sim}{\eta}) = (2\pi)^{-n/2} \left(\prod_1^n d_{tt}\right)^{-1/2} \exp\left(-\frac{1}{2}\sum_{t=1}^{n} x_t^{0^2}/d_{tt}\right)$$

$$\hat{\eta}_{EMLE} = \underset{\underset{\sim}{\eta}}{\arg\max}\, l(\underset{\sim}{\eta})$$; $l(\underset{\sim}{\eta})$ is the log likelihood

Once again Iterative optimization techniques are used to obtain EMLE using the data $(x_1, \ldots, x_n)$

MLE of Gaussian MA(q)

$$X_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$$

$$\epsilon_t \overset{i.i.d.}{\sim} N(0, \sigma^2)$$

$$\underset{\sim}{\eta} = (\mu, \theta_1, \ldots, \theta_q, \sigma^2)' \text{ – parameter vector}$$

Conditional MLE formulation

Consider the likelihood conditional on the assumption that

$$\epsilon_0 = 0, \quad \epsilon_{-1} = \epsilon_{-2} = \cdots = \epsilon_{-(q-1)} = 0$$

(at the expected value of $\epsilon_0, \epsilon_{-1}, \ldots, \epsilon_{-(q-1)}$)

Conditional likelihood (conditioned on $\underset{\sim}{\epsilon}_0 = (\epsilon_0, \epsilon_{-1}, \ldots, \epsilon_{-(q-1)})' = \underset{\sim}{0}$

$$L_c(\underset{\sim}{\eta}) = f_{X_n, \ldots, X_1 | \underset{\sim}{\epsilon}_0 = \underset{\sim}{0}}(x_n, \ldots, x_1; \underset{\sim}{\eta} \,|\, \underset{\sim}{\epsilon}_0 = \underset{\sim}{0})$$

$$= f_{X_n | X_{n-1}, \ldots X_1, \underset{\sim}{\epsilon}_0 = \underset{\sim}{0}} \; f_{X_{n-1}, \ldots X_1 | \underset{\sim}{\epsilon}_0 = \underset{\sim}{0}}$$

$$\vdots$$

$$= f_{X_1 | \underset{\sim}{\epsilon}_0 = \underset{\sim}{0}} \left(\prod_{t=2}^{n} f_{X_t | X_{t-1}, \ldots, X_1, \underset{\sim}{\epsilon}_0 = \underset{\sim}{0}} \right)$$

Note that $X_1 | \underset{\sim}{\epsilon}_0 = \underset{\sim}{0} \sim N(\mu, \sigma^2)$

ϵ_t given $x_t, x_{t-1}, \ldots \underset{\sim}{\epsilon}_0$ can be expressed as

$$\epsilon_t = x_t - \mu - \theta_1 \epsilon_{t-1} - \cdots - \theta_q \epsilon_{t-q}$$

$\forall t \geq 2$

$$X_t | X_{t-1}, \ldots, X_1, \underset{\sim}{\epsilon}_0 = \underset{\sim}{0} \sim N(\mu + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}, \sigma^2)$$

Conditional log likelihood

$$l_c(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\left(x_t - \mu - \theta_1\epsilon_{t-1} - \cdots - \theta_q\epsilon_{t-q}\right)^2$$

$$l_c(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\epsilon_t^2$$

$$\hat{\underset{\sim}{\eta}}_{CMLE} = \underset{\underset{\sim}{\eta}}{\arg\max}\ l_c(\underset{\sim}{\eta})$$

Iterative methods used to obtain CMLEs.

Exact MLE formulation

$\underset{\sim}{X} = (X_1, \ldots, X_n)'$ — $\{X_t\}$ is Gaussian MA(q)

$$\Rightarrow \underset{\sim}{X} \sim N_n(\underset{\sim}{\mu}, \Omega)$$

$$\underset{\sim}{\mu} = E(\underset{\sim}{X}) = \mu\,\underset{\sim}{1}_n$$

$$\Omega = Cov(\underset{\sim}{X}) = E(\underset{\sim}{X} - \underset{\sim}{\mu})(\underset{\sim}{X} - \underset{\sim}{\mu})'$$

$$= \begin{pmatrix} \gamma_0 & \gamma_1 & & \gamma_q & & 0 \\ \gamma_1 & \gamma_0 & \gamma_1 & & \ddots & \\ & \ddots & \ddots & \ddots & & \\ \gamma_q & & \ddots & \ddots & \ddots & \\ & \ddots & & \ddots & \ddots & \gamma_1 \\ 0 & & \gamma_q & & \gamma_1 & \gamma_0 \end{pmatrix}$$

$$\gamma_K = \begin{cases} \sigma^2(\theta_K + \theta_{K+1}\theta_1 + \cdots + \theta_q\theta_{q-K}), & K = 0, \ldots, q \\ 0, & o/w. \end{cases}$$

Exact log likelihood f^n

$$l(\underset{\sim}{\eta}) = -\frac{n}{2}\log 2\pi - \frac{1}{2}\log|\Omega| - \frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{\mu})'\Omega^{-1}(\underset{\sim}{x}-\underset{\sim}{\mu}).$$

$$\hat{\underset{\sim}{\eta}}_{EMLE} = \underset{\underset{\sim}{\eta}}{\arg\max}\ l(\underset{\sim}{\eta})$$

Note: Similar to MA(1), decomposition of Ω as $\Omega = ADA'$ leads to simplification of $l(\underset{\sim}{\eta})$, expressing it explicitly in terms of the parameters $\mu, \theta_1, .., \theta_q$ & σ^2.

Note: Iterative optimization techniques are used to obtain the EMLEs using $l(\underset{\sim}{\eta})$.

Remark: Conditional LSE for MA model

Consider an invertible MA(1) $(|\theta|<1)$

$$X_t = \mu + \epsilon_t + \theta\epsilon_{t-1} \ ; \ \underset{\sim}{\eta} = (\mu, \theta)'$$

Ordinary LSE of μ, θ which is defined as

$$\hat{\eta}_{OLS} = \underset{\underset{\sim}{\eta}}{\arg\min} \underbrace{\sum_{t=2}^{n}(x_t - \mu - \theta.\epsilon_{t-1})^2}_{\sum_{t=2}^{n}\epsilon_t^2}$$

is not feasible since ϵ_{t-1} is not observable

Let us put a condition that ϵ_0 is given (e.g $\epsilon_0 = 0$ It's expected value)

$$x_1 = \mu + \epsilon_1 + \theta \epsilon_0 ; \quad \epsilon_1 = (x_1 - \mu) - \theta \epsilon_0$$

$$x_2 = \mu + \epsilon_2 + \theta \epsilon_1 ; \quad \epsilon_2 = (x_2 - \mu) - \theta \epsilon_1$$

$$\epsilon_2 = (x_2 - \mu) - \theta\left((x_1 - \mu) - \theta \epsilon_0\right)$$

$$\text{i.e. } \epsilon_2 = (x_2 - \mu) - \theta(x_1 - \mu) + \theta^2 \epsilon_0$$

$$\vdots$$

$$x_{t-1} = \mu + \epsilon_{t-1} + \theta \epsilon_{t-2}$$

$$\epsilon_{t-1} = (x_{t-1} - \mu) - \theta \epsilon_{t-2}$$
$$= (x_{t-1} - \mu) - \theta\left((x_{t-2} - \mu) - \theta \epsilon_{t-3}\right)$$

$$\vdots$$

$$\epsilon_{t-1} = \sum_{i=1}^{t-1} (x_i - \mu)(-\theta)^{t-1-i} \quad (\epsilon_0 = 0).$$

Conditional LSE of μ & θ is defined as

$$\hat{\underset{\sim}{\eta}}_{CLSE} = \arg\min_{\underset{\sim}{\eta}} \sum_{t=2}^{n} \left(x_t - \mu - \theta \sum_{i=1}^{t-1} (x_i - \mu)(-\theta)^{t-1-i}\right)^2$$

<u>Remark</u>: CLSE for MA(q) can be framed in a similar manner.