Suppose $\hat{X}_t = \alpha_1 X_{t-1} + \cdots + \alpha_{k-1} X_{t-k+1}$

where $\{\alpha_1, \ldots, \alpha_{k-1}\}$ are the mean squared linear regression coefficients obtained by minimizing $E[(X_t - \hat{X}_t)^2]$.

The routine minimization method through differentiation gives the following linear system of equation:

$$\gamma_i = \alpha_1 \gamma_{i-1} + \alpha_2 \gamma_{i-2} + \cdots + \alpha_{k-1} \gamma_{i-k+1}, \quad (1 \le i \le k-1) \qquad (*)$$

Hence

$$\rho_i = \alpha_1 \rho_{i-1} + \alpha_2 \rho_{i-2} + \cdots + \alpha_{k-1} \rho_{i-k+1}, \quad (1 \le i \le k-1)$$

In term of matrix notation, the above system becomes

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_{k-1} \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} \\ \rho_1 & 1 & \rho_1 & & \rho_{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-2} & \rho_{k-3} & \rho_{k-4} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{k-1} \end{bmatrix} \quad \cdots (**)$$

Similarly $\hat{X}_{t-k} = \eta_1 X_{t-1} + \cdots + \eta_{k-1} X_{t-k+1}$

where $\{\eta_1, \ldots, \eta_{k-1}\}$ are the mean squared linear regression coefficients obtained by minimizing $E[(X_{t-k} - \hat{X}_{t-k})^2]$.

Hence

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_{k-1} \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} \\ \rho_1 & 1 & \rho_1 & & \rho_{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-2} & \rho_{k-3} & \rho_{k-4} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_{k-1} \end{bmatrix}$$

which implies that $\alpha_i = \eta_i \ (1 \le i \le k-1)$.

The formula for PACF at lag k is:

$$\phi_{kk} = \operatorname{corr}\left((X_t - \hat{X}_t), (X_{t-k} - \hat{X}_{t-k}) \right)$$

Now,

$$\operatorname{Var}(X_t - \hat{X}_t) = E\left[(X_t - \alpha_1 X_{t-1} - \cdots - \alpha_{k-1} X_{t-k+1})^2 \right]$$

$$= E\left[X_t (X_t - \alpha_1 X_{t-1} - \cdots - \alpha_{k-1} X_{t-k+1}) \right]$$

equals to 0 $\because$ (*)

$$\left.\begin{array}{l} -\alpha_1 E\left[X_{t-1} (X_t - \alpha_1 X_{t-1} - \cdots - \alpha_{k-1} X_{t-k+1}) \right] \\ - \cdots \\ -\alpha_{k-1} E\left[X_{t-k+1} (X_t - \alpha_1 X_{t-1} - \cdots - \alpha_{k-1} X_{t-k+1}) \right] \end{array}\right.$$

$$= E\left[X_t (X_t - \alpha_1 X_{t-1} - \cdots - \alpha_{k-1} X_{t-k+1}) \right]$$

$$= \gamma_0 - \alpha_1 \gamma_1 - \cdots - \alpha_{k-1} \gamma_{k-1}$$

Hence:

$$\operatorname{Var}(X_{t-k} - \hat{X}_{t-k}) = \operatorname{Var}(X_t - \hat{X}_t)$$

$$= \gamma_0 - \alpha_1 \gamma_1 - \cdots - \alpha_{k-1} \gamma_{k-1}$$

※ Next, using $\alpha_i = \eta_i$ $(1 \le i \le k-1)$, we have

$$\operatorname{Cov}\left((X_t - \hat{X}_t), (X_{t-k} - \hat{X}_{t-k}) \right)$$

$$= E\left[(X_t - \alpha_1 X_{t-1} - \cdots - \alpha_{k-1} X_{t-k+1})(X_{t-k} - \alpha_1 X_{t-1} - \cdots - \alpha_{k-1} X_{t-k+1}) \right]$$

$$= E\left[(X_t - \alpha_1 X_{t-1} - \cdots - \alpha_{k-1} X_{t-k+1})(X_{t-k}) \right]$$

$$= \gamma_k - \alpha_1 \gamma_{k-1} - \cdots - \alpha_{k-1} \gamma_1$$

Therefore :

$$\phi_{kk} = \frac{\gamma_k - \alpha_1 \gamma_{k-1} - \cdots - \alpha_{k-1}\gamma_1}{\gamma_0 - \alpha_1\gamma_1 - \cdots - \alpha_{k-1}\gamma_{k-1}}$$

$$= \frac{\rho_k - \alpha_1 \rho_{k-1} - \cdots - \alpha_{k-1}\rho_1}{1 - \alpha_1\rho_1 - \cdots - \alpha_{k-1}\rho_{k-1}} \quad \cdots\cdots (***)$$

By Cramer's Rule, we can calculate α_i as following:

$$\alpha_i = \frac{\begin{vmatrix} 1 & \rho_1 & \cdots & \rho_{i-2} & \rho_1 & \rho_i & \cdots & \rho_{k-2} \\ \rho_1 & 1 & & \rho_{i-3} & \rho_2 & \rho_{i-1} & & \rho_{k-3} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \rho_{k-2} & \rho_{k-3} & & \rho_{k-i} & \rho_{k-1} & \rho_{k-i-2} & & 1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \cdots & \rho_{i-2} & \rho_{i-1} & \rho_i & \cdots & \rho_{k-2} \\ \rho_1 & 1 & & \rho_{i-3} & \rho_{i-2} & \rho_{i-1} & & \rho_{k-3} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \rho_{k-2} & \rho_{k-3} & & \rho_{k-i} & \rho_{k-i-1} & \rho_{k-i-2} & & 1 \end{vmatrix}}$$

Subtituting α_i above to (***) and simplify the form, we have

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{vmatrix}}$$

For MA(1) Model, as we have $\rho_1 = \dfrac{\theta_1}{1+\theta_1^2}$ and $\rho_h = 0$ for any $h \geq 2$, therefore

$$\phi_{KK} = \frac{\begin{vmatrix} 1 & \rho_1 & 0 & \cdots & 0 & \rho_1 \\ \rho_1 & 1 & \rho_1 & & 0 & 0 \\ 0 & \rho_1 & 1 & & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & & \rho_1 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & 0 & \cdots & 0 & 0 \\ \rho_1 & 1 & \rho_1 & & 0 & 0 \\ 0 & \rho_1 & 1 & & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & & \rho_1 & 1 \end{vmatrix}}$$

$$\phi_{11} = \frac{|\rho_1|}{|1|} = \rho_1 = \frac{\theta_1}{1+\theta_1^2}$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{-\rho_1}{1-\rho_1^2} = \frac{-\theta_1^2}{1+\theta_1^2+\theta_1^4}$$

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & 0 \\ 0 & \rho_1 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & 0 \\ \rho_1 & 1 & \rho_1 \\ 0 & \rho_1 & 1 \end{vmatrix}} = \frac{\rho_1^3}{1-2\rho_1^2} = \frac{\theta_1^3}{1+\theta_1^2+\theta_1^4+\theta_1^6}$$

There is a clear pattern that indicates:

$$\phi_{KK} = \frac{(-1)^{K+1}\,\theta_1^K}{\sum_{i=0}^{K} \theta_1^{2i}}$$