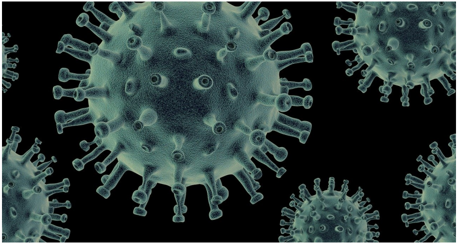
Autocorrelation Function (ACF) for MA(q).

November 17, 2020



Basic formulas on expectation, covariance and variance - repetition

Let $(\Omega, \mathcal{F}, P)$ be a probability space:

- Random variable X is a Borel mapping $X : \Omega \mapsto \mathbb{R}$, i.e.

$$(B \text{ is a Borel set }) \Rightarrow X^{-1}(B) \in \mathcal{F}$$

  call the set of random variables $\mathcal{B}(\Omega)$;
- $\mathcal{L}^1(\Omega)$ is the space of random variables whose moduli has a finite *expectation*:

$$\mathcal{L}^1(\Omega) = \{X \in \mathcal{B}(\Omega) : E|X| < \infty\};$$

- similarly we denote $\mathcal{L}^2(\Omega)$

$$\mathcal{L}^2(\Omega) = \left\{X \in \mathcal{B}(\Omega) : E(X^2) < \infty\right\},$$

- and more generally

$$\mathcal{L}^k(\Omega) = \left\{X \in \mathcal{B}(\Omega) : E|X|^k < \infty\right\},$$

  and the moduli with even k is useless.

Basic formulas on expectation, covariance and variance - repetition

Linearity of expectation

Assume $X, Y \in \mathcal{L}^1(\Omega)$. Then,

- there exists $E(X+Y)$ (i.e. $X+Y \in \mathcal{L}^1(\Omega)$) and

$$E(X+Y) = E(X) + E(Y);$$

- for any $\alpha \in \mathbb{R}$, $E(\alpha X)$ exists (i.e. $\alpha X \in \mathcal{L}^1(\Omega)$) and

$$E(\alpha X) = \alpha E(X).$$

Basic formulas on expectation, covariance and variance - repetition

Remark

The above condition can be alternatively expressed as follows: if $X, Y \in \mathcal{L}^1(\Omega)$, then

- for any pair of constants $\alpha, \beta \in \mathbb{R}$ there exists $\alpha X + \beta Y \in \mathcal{L}^1(\Omega)$ and

$$E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y);$$

- in other words, $\mathcal{L}^1(\Omega)$ is the vector subspace of $\mathcal{B}(\Omega)$ and $E(\cdot)$ is a *linear functional* on it.

Basic formulas on expectation, covariance and variance - repetition

Covariance is bilinear

The set $\mathcal{L}^2(\Omega)$ is a vector subspace of $\mathcal{L}^1(\Omega)$. Moreover, *Cov* is a bilinear operator on $\mathcal{L}^2(\Omega) \times \mathcal{L}^2(\Omega)$. That is, if $X, Y, Z \in \mathcal{L}^2(\Omega)$, then

- For any α, β

  $$Cov(\alpha X + \beta Y, Z) = \alpha Cov(X, Z) + \beta Cov(Y, Z)$$

  and

  $$Cov(Z, \alpha X + \beta Y) = \alpha Cov(Z, X) + \beta Cov(Z, Y)$$

- In other words: for any $Z \in \mathcal{L}^2(\Omega)$, both $Cov(\cdot, Z)$ and $Cov(Z, \cdot)$ are a linear functional on $\mathcal{L}^2(\Omega)$.

Basic formulas on expectation, covariance and variance - repetition

Covariance is bilinear- distributive properties

Because of the bilinearity of the covariance we can observe the *distributive properties* (similar to those in standard arithmetic) between adding random variables and multiplying by scalars:

$$Cov(\alpha_1 X_1 + \beta_1 Y_1, \alpha_2 X_2 + \beta_2 Y_2) =$$

$$= \alpha_1 Cov(X_1, \alpha_2 X_2 + \beta_2 Y_2) + \beta_1 Cov(Y_1, \alpha_2 X_2 + \beta_2 Y_2)$$

$$= \alpha_1(\alpha_2 Cov(X_1, X_2) + \beta_2 Cov(X_1, Y_2)) +$$

$$+\beta_1(\alpha_2 Cov(Y_1, X_2) + \beta_2 Cov(Y_1, Y_2)) =$$

$$= \alpha_1\alpha_2 Cov(X_1, X_2) + \alpha_1\beta_2 Cov(X_1, Y_2))$$

$$+\beta_1\alpha_2 Cov(Y_1, X_2) + \beta_1\beta_2 Cov(Y_1, Y_2).$$

Basic formulas on expectation, covariance and variance - repetition

Covariance is bilinear- distributive properties

More generally, if $X_1, X_2, \dots, X_n$, and $Y_1, Y_2, \dots, Y_n$ are random variables $\in \mathcal{L}^2(\Omega)$ and $\alpha_1, \alpha_2, \dots, \alpha_n$ and $\beta_1, \beta_2, \dots, \beta_n$ are scalars from $\mathbb{R}$ then using bilinearity twice, we obtain

$$
\begin{aligned}
Cov\left(\sum_{k=1}^{n} \alpha_k X_k, \sum_{k=1}^{n} \beta_k Y_k\right) &= \sum_{i=1}^{n} \alpha_i Cov\left(X_i, \sum_{j=1}^{n} \beta_j Y_j\right) \\
&= \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{n} \beta_j Cov(X_i, Y_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j Cov(X_i, Y_j).
\end{aligned}
$$

Compare with distributive properties in arithmetic.

Basic formulas on expectation, covariance and variance - repetition

Properties of variance

Remembering that for any $X, Y \in \mathcal{L}^2(\Omega)$, $Cov(X,X) = Var(X)$ and $Cov(X,Y) = Cov(Y,X)$, we have the following formula: let $X_1, X_2, \ldots, X_n \in \mathcal{L}^2(\Omega)$. Then

$$Var\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} Var(X_k) + 2\sum_{i<j} Cov(X_i, X_j).$$

In particular, if $X_1, X_2, \ldots, X_n$ are pairwise uncorrelated then

$$Var\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} Var(X_k).$$

VERTE-for diriving this formula.

Basic formulas on expectation, covariance and variance - repetition

Deriving the formula on variance

We have

$$
\begin{aligned}
Var\left(\sum_{k=1}^{n} X_k\right) &= Cov\left(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} X_k\right) \\
&= \sum_{i=1}^{n} Cov\left(X_i, \sum_{j=1}^{n} X_j\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j) \\
&= \sum_{i=1}^{n} Cov(X_i, X_i) + \sum_{i \neq j} Cov(X_i, X_j).(VERTE)
\end{aligned}
$$

Basic formulas on expectation, covariance and variance - repetition

Deriving the formula on variance - continue

Furhermore, by $Cov(X_i, X_i) = Var(X_i)$ and $Cov(X_i, X_j) = Cov(X_j, X_i)$ we have

$$Var\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} Var(X_k) + 2\sum_{i<j} Cov(X_i, X_j).$$

If $X_1, X_2, ..., X_n$ are pairwise uncorrelated then $Cov(X_i, X_j) = 0$ for $i \neq 0$, hence

$$Var\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} Var(X_k).$$

Moving average MA(q)

We consider the time series MA(q) in the form:

$$X_t = \epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} - \ldots - \theta_q\epsilon_{t-q},$$

where $\theta_1, \theta_2, \ldots, \theta_p$ are unknown parameters. Here ϵ_t is a white noise. That is for all $t \in \mathbb{N}$:

- the expectation of ϵ_t is 0, i.e.

$$E(\epsilon_t) = 0$$

- the variables ϵ_t are pairwise uncorrelated, and the variance of ϵ_t is constant, i.e.

$$Cov(\epsilon_t, \epsilon_{t+h}) = \begin{cases} 0 & if \quad h \neq 0 \\ \sigma^2 & if \quad h = 0 \end{cases}$$

  for some σ^2.

Stationarity of MA(q)

The time series X_t in the form of MA(q) is always stationary, i.e.

- the expectation is a constant, i.e.
- the covariance function $Cov(X_t, X_{t+h}) = \gamma(h)$ depends only on h;
- in particalar, $Var(X_t) = \gamma(0) = \sigma^2$ is a constant value.

Let $h > 0$. We write $\gamma(h) = Cov(X_t, X_{t+h})$. Write

$$X_t = \epsilon_t - \sum_{k=1}^{q} \theta_k \epsilon_{t-k}.$$

We have

$$X_{t+h} = \epsilon_{t+h} - \sum_{k=1}^{q} \theta_k \epsilon_{t+h-k}.$$

Covariance of MA(q) - deriving the formula (CONTINUE)

By bilinearity of covariance and its distributive properties and the properties of ϵ_t, for $h \geq 0$ we have

$$
\begin{aligned}
Cov(X_t, X_{t+h}) &= Cov\left(\epsilon_t - \sum_{k=1}^{q} \theta_k \epsilon_{t-k}, \epsilon_{t+h} - \sum_{k=1}^{q} \theta_k \epsilon_{t+h-k}\right) \\
&= Cov(\epsilon_t, \epsilon_{t+h}) - \underbrace{\sum_{k=1}^{q} \theta_k Cov(\epsilon_{t+h}, \epsilon_{t-k})}_{\text{all components are } 0} \\
&- \underbrace{\sum_{k=1}^{q} \theta_k Cov(\epsilon_t, \epsilon_{t+h-k})}_{\text{all but } k=h-th \text{ component is } 0.} \\
&+ \sum_{i=1}^{q} \sum_{j=1}^{q} \theta_i \theta_j Cov(\epsilon_{t-i}, \epsilon_{t+h-j}).
\end{aligned}
$$

Having that for $h = 1, 2, ..., q-1$ we have

$$Cov(X_t, X_{t+h}) = Cov(\epsilon_t, \epsilon_{t+h}) - \theta_h Cov(\epsilon_t, \epsilon_t) + \underbrace{\sum_{i,j} \theta_i \theta_j Cov(\epsilon_{t-i}, \epsilon_{t+h-j})}_{\text{all but indexes (i,j) satisfying } j-i=h \text{ are } 0.}.$$

Hence for such h we have

$$Cov(X_t, X_{t+h}) = \sigma^2 \left(-\theta_h + \sum_{k=1}^{q-h} \theta_k \theta_{k+h} \right) = \gamma(h).$$

Now we derive the formula for $h = q$:

$$Cov(X_t, X_{t+q}) = -\sigma^2\theta_q + \underbrace{\sum_{(i,j):i-j=q} \theta_i\theta_j Cov(\epsilon_{t-i}, \epsilon_{t+q-j})}_{\text{suming over empty set of (i,j).}}.$$

Hence for such $h = q$ we have

$$Cov(X_t, X_{t+q}) = -\sigma^2\theta_q = \gamma(q).$$

Covariance of MA(q) - deriving the formula (CONTINUE)

Now we derive the formula for $h > q$:

$$Cov(X_t, X_{t+h}) = -\sigma^2 \underbrace{\sum_{k=1}^{q} \theta_k Cov(\epsilon_t, \epsilon_{t+h-k})}_{\text{all components are } 0.}$$

$$+ \underbrace{\sum_{(i,j):i-j=h} \theta_i \theta_j Cov(\epsilon_{t-i}, \epsilon_{t+h-j})}_{\text{suming over empty set of (i,j).}}.$$

Hence for such $h > q$ we have

$$Cov(X_t, X_{t+h}) = 0.$$

Covariance of MA(q) - deriving the formula (CONTINUE)

Now we derive the formula for $h = 0$. We have

$$Cov(X_t, X_t) = Var(X_t)$$

$$= \underbrace{Var\left(\epsilon_t - \sum_{k=1}^{q} \theta_k \epsilon_{t-k}\right)}_{\text{all components inside are uncorrelated random variables}}$$

$$= Var(\epsilon_t) + \sum_{k=1}^{q} Var(\theta_k \epsilon_{t-k})$$

$$= \sigma^2 + \sum_{k=1}^{q} \underbrace{\theta_k^2 Var(\epsilon_{t-k})}_{\text{By the formula } Var(\alpha X) = \alpha^2 Var(X), \alpha \in \mathbb{R}.}$$

$$= \sigma^2 \left(1 + \sum_{k=1}^{q} \theta_k^2\right) = \gamma(0).$$

Covariance of MA(q) - deriving the formula (CONTINUE)

Finally we can express $Cov(X_t, X_{t+h}) = \gamma(h)$ with the following form:

$$\gamma(h) = \begin{cases} \sigma^2 \left(1 + \sum_{k=1}^{q} \theta_k^2\right) & \text{if} \quad h = 0 \\ \sigma^2 \left(-\theta_h + \sum_{k=1}^{q-h} \theta_k \theta_{k+h}\right) & \text{if} \quad 0 < h < q, \\ -\sigma^2 \theta_q & \text{if} \quad h = q \\ 0 & \text{if} \quad h \geq q. \end{cases}$$

and for negative $h < 0$ we put $\gamma(h) := \gamma(-h)$. Hence $Cov(X_t, X_{t+h})$ depends only on h (not t).

We have obtained that the model MA(q) is stationary. Indeed,

- the expectation is always 0, since by linearity of E and the properties of ϵ_t we have

$$E\left(\epsilon_t - \sum_{k=1}^{q} \theta_k \epsilon_{t-k}\right) = E\epsilon_t - \sum_{k=1}^{q} \theta_k E(\epsilon_{t-k}) = 0;$$

- as we have derived, $Cov(X_t, X_{t+h}) = \gamma(h)$:

$$\gamma(h) = \begin{cases} \sigma^2 \left(1 + \sum_{k=1}^{q} \theta_k^2\right) & \text{if} \quad h = 0, \\ \sigma^2 \left(-\theta_h + \sum_{k=1}^{q-h} \theta_k \theta_{k+h}\right) & \text{if} \quad 0 < |h| < q, \\ -\sigma^2 \theta_q & \text{if} \quad |h| = q, \\ 0 & \text{if} \quad |h| \geq q. \end{cases}$$

The derived formula of ACF for MA(q)

By derived formula of $\gamma(h)$ and by $\rho(h) = \frac{\gamma(h)}{\gamma(0)}$ we can easily derive the formula for the **autocorrelation function (ACF)**:

$$\rho(h) = \begin{cases} 1 & \text{if} \quad h = 0, \\ \dfrac{-\theta_h + \sum\limits_{k=1}^{q-h} \theta_k \theta_{k+h}}{1 + \sum\limits_{k=1}^{q} \theta_k^2} & \text{if} \quad 0 < |h| < q, \\ \dfrac{-\theta_q}{1 + \sum\limits_{k=1}^{q} \theta_k^2} & \text{if} \quad |h| = q, \\ 0 & \text{if} \quad |h| \geq q. \end{cases}$$

Dress rehearsal for estimation of parameters of MA(q)

- Similarly as in case of AR(p) we can try to estimate the parameters $\theta_1, \theta_2, \ldots, \theta_q$ MA(q) in the form

$$X_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \ldots - \theta_q \epsilon_{t-q}$$

- Substituting unknown $\rho(h) = \rho_h$ by its empirical counterpart :

$$\hat{\rho}_h = \frac{\frac{1}{n} \sum_{t=1}^{n-h} (X_t - \bar{X})(X_{t+h} - \bar{X})}{\frac{1}{n} \sum_{t=1}^{n} (X_t - \bar{X})^2}$$

we can compute $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_q$ as a solution of the following system of nonlinear equations (VERTE)

Dress rehearsal for estimation of parameters of MA(q)

$$
\begin{cases}
\hat{\rho}_1 & = & \frac{-\theta_1+\theta_1\theta_2+\theta_2\theta_3+\ldots+\theta_{q-1}\theta_q}{1+\theta_1^2+\theta_2^2+\ldots+\theta_q^2} \\
\hat{\rho}_2 & = & \frac{-\theta_2+\theta_1\theta_3+\theta_2\theta_4+\ldots+\theta_{q-2}\theta_q}{1+\theta_1^2+\theta_2^2+\ldots+\theta_q^2} \\
\vdots & \vdots & \vdots \\
\hat{\rho}_{q-1} & = & \frac{-\theta_{q-1}+\theta_1\theta_q}{1+\theta_1^2+\theta_2^2+\ldots+\theta_q^2} \\
\hat{\rho}_q & = & \frac{-\theta_q}{1+\theta_1^2+\theta_2^2+\ldots+\theta_q^2}
\end{cases}
$$

Dress rehearsal for estimation of parameters of MA(q)

- The solution of the system of equations with respect to $\theta_1, \theta_2, \ldots, \theta_q$ yields the estimator;
- The solution is provided by numerical methods only (for example innovation algorithm, for the details see Brockwell, P.J. and Davis, R.A., *Introduction to Time Series and Forecasting, Second Edition*, page 71 and 150);
- The solution obtained by innovative algorithm is a preliminary estimator of $\theta_1, \theta_2, \ldots, \theta_q$;
- Unlike in case of AR(p) where we solve Yule-Walker formula, there is no analytical solution of $\theta_1, \theta_2, \ldots, \theta_q$;

ACF function for MA(q)

Properties of ACF for MA(q)

If a time series X_t is in the form of MA(q), then ACF for it obeys the following conditions:

- Only $\rho_1, \rho_2, \ldots, \rho_{q-1}$ and ρ_q may be nonzero (see slide 21 in this presentation);
- ρ_q must be nonzero, otherwise the model is MA(q-1), or MA(q-2), or even with the lower delay:

$$\rho_q \neq 0, \text{ but } \rho_{q+1} = \rho_{q+2} = \ldots = 0$$

- In other words, ρ_h falls into 0 from $h = q$ into $h = q + 1$ and stays 0 forever;

ACF function for MA(q)

If X_t is a time series in the form of MA(q), then

- $\hat{\rho}_q$ should be significant (the significance thresholds will be define later), that is

$$|\hat{\rho}_q| \geq \text{significance threshold}$$

- no more than 5 per cent observations $\hat{\rho}_{q+1}, \hat{\rho}_{q+2}, \dots, \hat{\rho}_n$ are insignificant;
- for $h > q$, even if $|\hat{\rho}_h|$ exceeds the significance threshold, the extension must not be to large.

Illustration of ACF function for MA(q)

The plot in the next page illustrates possible values of $\hat{\rho}_h$ for $h = 0, 1, 2, ..., 105$ for time series $MA(5)$ based on a time series $X_1, X_2, \ldots, X_{106}$. We accept the hypothesis that the series is MA(5) if the following conditions hold:

- obviously, $\hat{\rho}_0 = 1$;
- $\hat{\rho}_5$ is significant, i.e. the 5-th bar exceeds the significance level (crosses one red line) (since $q = 5$);
- all but at most $5(= 100 * 20\%)$ bars from $\hat{\rho}_6, \hat{\rho}_7, \ldots, \hat{\rho}_{105}$, i.e. all but 5 of these bars stay in the insignificance level (bars between the red lines);
- any significant value $\hat{\rho}_6, \hat{\rho}_7, \ldots, \hat{\rho}_{105}$ are near significance level (exceeds slightly red line, not to much);

Illustration of ACF function for MA(q)

The *textbook example* of ACF function for MA(5) is plotted below:

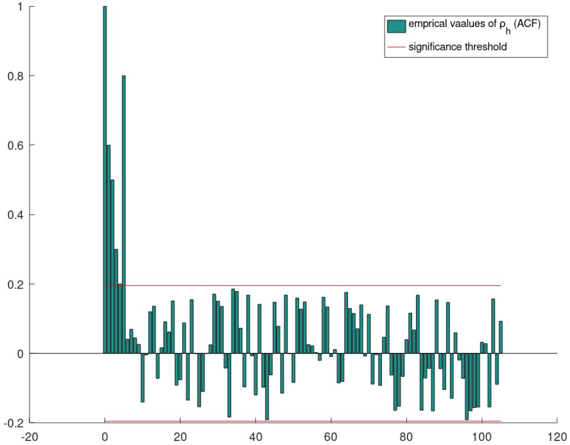


Illustration of ACF function for MA(q)

The *textbook example* of ACF function for MA(5) is plotted below:

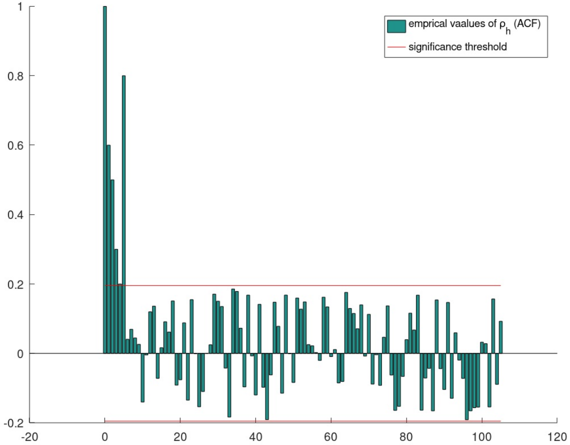
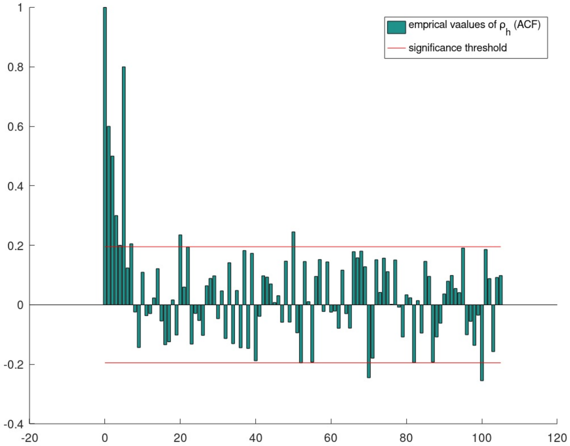


Illustration of ACF function for MA(q)

We may suppose this model to be MA(5). Only $\hat{\rho}_7$, $\hat{\rho}_{20}$, $\hat{\rho}_{50}$, $\hat{\rho}_{70}$ and $\hat{\rho}_{100}$ cross red line, but these values slightly exceed the threshold of significance.



We see that many bars $\hat{\rho}_h$ with $h > 5$ cross the red line. In fact, all crossing are slightly, but to many to accept this model to be in form MA(5).

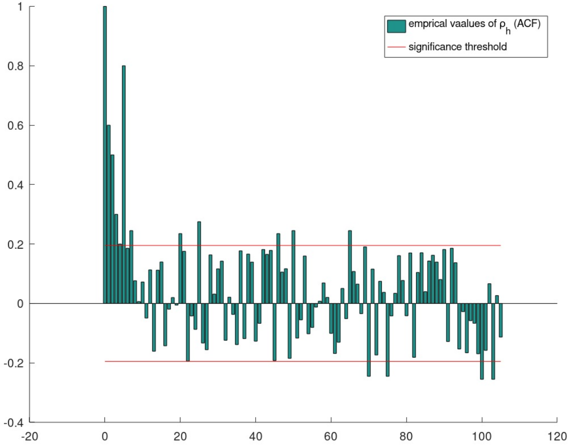


Illustration of ACF function for MA(q)

One radical crossing red line by $\hat{\rho}_{50}$ excludes the possibility this model to be MA(5), perhaps MA(50).

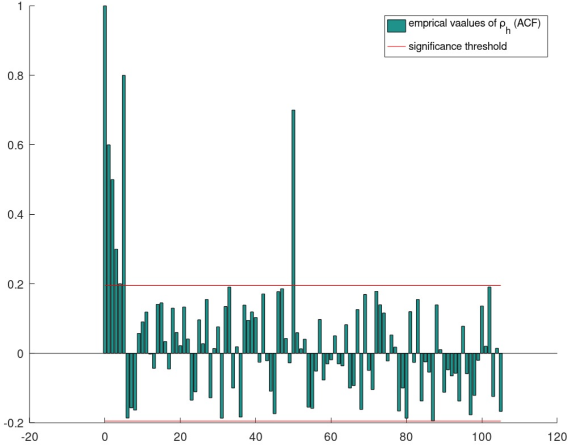


Illustration of ACF function for MA(q)

The interpretation of the bar plot may be loose, especially if we have numerous bar:

- we rather accept the hypothesis if the number of crossing the red line (after $h > q$) is sparse at a glance;
- and no of the crossing is radical at a glance.