Estimating the autocorrelation

Andreas Jakobsson

Estimating the autocorrelation

The two standard ways to estimate the ACF can be expressed as

$$\hat{r}_y^u(k) = \frac{1}{N-k} \sum_{t=k+1}^{N} \left(y_t - \hat{m}_y\right)\left(y_{t-k} - \hat{m}_y\right)^*$$

$$\hat{r}_y^b(k) = \frac{1}{N} \sum_{t=k+1}^{N} \left(y_t - \hat{m}_y\right)\left(y_{t-k} - \hat{m}_y\right)^* = \frac{1}{N}\psi_k$$

for $0 \le k \le N-1$. Which is the better estimate?

Estimating the autocorrelation

The two standard ways to estimate the ACF can be expressed as

$$\hat{r}_y^u(k) = \frac{1}{N-k} \sum_{t=k+1}^{N} \left(y_t - \hat{m}_y\right)\left(y_{t-k} - \hat{m}_y\right)^*$$

$$\hat{r}_y^b(k) = \frac{1}{N} \sum_{t=k+1}^{N} \left(y_t - \hat{m}_y\right)\left(y_{t-k} - \hat{m}_y\right)^* = \frac{1}{N}\psi_k$$

for $0 \le k \le N-1$. Which is the better estimate?

It can be shown that $E\{\psi_k\} \approx (N-k)\left(r_y(k) - V\{\hat{m}_y\}\right)$, which implies that

$$E\{\hat{r}_y^u(k)\} = \frac{1}{N-k} E\{\psi_k\} = r_y(k) - V\{\hat{m}_y\}$$

$$E\{\hat{r}_y^b(k)\} = \frac{1}{N} E\{\psi_k\} = r_y(k) - \frac{k}{N} r_y(k) - \frac{N-k}{N} V\{\hat{m}_y\}$$

Both estimators are thus biased.

Estimating the autocorrelation

The two standard ways to estimate the ACF can be expressed as

$$\hat{r}_y^u(k) = \frac{1}{N-k} \sum_{t=k+1}^{N} \left(y_t - \hat{m}_y\right)\left(y_{t-k} - \hat{m}_y\right)^*$$

$$\hat{r}_y^b(k) = \frac{1}{N} \sum_{t=k+1}^{N} \left(y_t - \hat{m}_y\right)\left(y_{t-k} - \hat{m}_y\right)^* = \frac{1}{N}\psi_k$$

for $0 \le k \le N-1$. Which is the better estimate?

It can be shown that $E\{\psi_k\} \approx (N-k)\left(r_y(k) - V\{\hat{m}_y\}\right)$, which implies that

$$E\{\hat{r}_y^u(k)\} = \frac{1}{N-k} E\{\psi_k\} = r_y(k) - V\{\hat{m}_y\}$$

$$E\{\hat{r}_y^b(k)\} = \frac{1}{N} E\{\psi_k\} = r_y(k) - \frac{k}{N} r_y(k) - \frac{N-k}{N} V\{\hat{m}_y\}$$

Both estimators are thus biased.

For a *zero-mean* process

$$E\{\psi_k\} = \sum_{t=k+1}^{N} r_y(k) = (N-k) r_y(k)$$

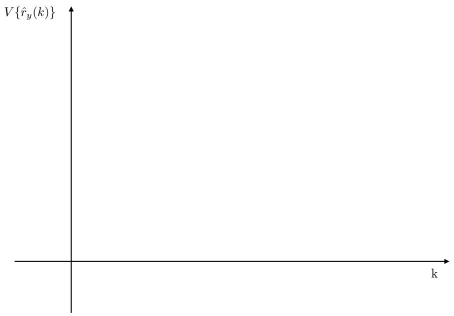
In this case, this implies that

$$E\{\hat{r}_y^u(k)\} = \frac{1}{N-k} E\{\psi_k\} = r_y(k)$$

$$E\{\hat{r}_y^b(k)\} = \frac{1}{N} E\{\psi_k\} = \frac{N}{N-k} r_y(k)$$

In this case, $r_y^u(k)$ is unbiased, whereas $r_y^b(k)$ is only *asymptotically* unbiased.

Estimating the autocorrelation



Estimating the autocorrelation

For a zero-mean white Gaussian process with variance σ_e^2, and

$$\hat{\rho}_e(k) = \frac{\hat{r}_e(k)}{\hat{r}_e(0)}$$

where $\hat{r}_e(k)$ is the *biased* estimate of the ACF, then, for $k \neq 0$,

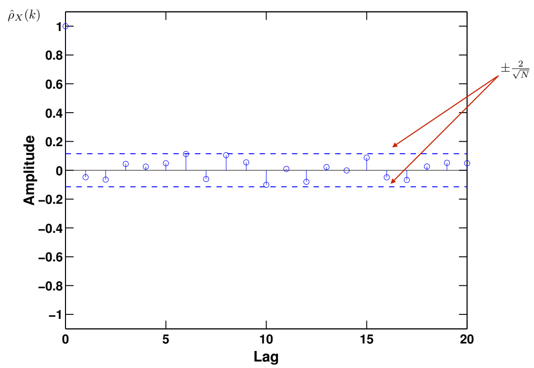
$$E\{\hat{\rho}_e(k)\} = 0$$
$$V\{\hat{\rho}_e(k)\} = \frac{1}{N}$$

Furthermore, $\hat{\rho}_e(k)$ is asymptotically Normal distributed for $k > 0$.

This implies that a 95% confidence interval can be formed as

$$\hat{\rho}_e(k) \approx 0 \pm \frac{2}{\sqrt{N}}$$

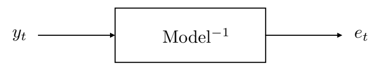
Estimating the autocorrelation



As far as we can tell, this is a *white* noise. This assumes that $\hat{\rho}_e(k) \in \mathcal{N}$.

About 5% of the ACF to be (slightly) outside the confidence intervall. As a rule of thumb, only estimate $r_y(k)$ for lags up to (at most) $N/4$.

Testing the model



Is the residual white?