Specialist Courses

Daniele Zago

December 13, 2021

Contents

Lecture 3: Transfer function models

2021-12-06

Theorem 1 (Wold decomposition)
Let $(X_t)_t$ be a non-deterministic stationary time series with $\mathbb{E}[X_t] = 0$, then

$$X_t = \sum_{j=0}^{\infty} \psi_j a_{t-j} + V_t,$$

where V_t is deterministic and

1. $\psi_0 = 1$ *and* $\sum_{j=1}^{\infty} \psi_j^2 < \infty$.
2. $a_t = WN(0, \sigma^2)$.
3. $\mathbb{E}[a_t V_s]$ *for all* $s, t = 0, \pm 1, \pm 2, \dots$

With this decomposition we can approximate any stationary time series using a linear process of the form

$$X_t = \mu + \sum_{j=0}^{\infty} \psi_j a_{t-j},$$

where $\sum_{j=1}^{\infty} |\psi_j| < \infty$.

3.1 Transfer function models

Transfer function models are models where an output series is related to one or more input series. We link the output series y_t to the predictor series x_t via

$$y_i = \nu(B) x_t + \eta_t, \tag{1}$$

where $\nu(B) = \sum_{j=-\infty}^{\infty} \nu_j B^j$ is the *transfer function of the linear filter* that transform x_t into y_t, and η_t is a noise series independent of x_t. The weights ν_j are called ***impulse response weights*** and the TFM is called ***stable*** if

$$\sum_{j=-\infty}^{\infty} |\nu_j| < \infty,$$

and in particular we are in a BIBO setting (Bounded Input Bounded Output). The TFM is said to be ***causal*** if $\nu_j = 0$ for $j < 0$, since the present output is affected only by the system current and past values,

$$y_t = \sum_{j=0}^{\infty} \nu_j B^j.$$

The purpose of TF models is to identify the TF $\nu(B)$ and the noise model, possibly using a simpler representation which is similar to an ARIMA model

$$\delta(B) y_t = \omega(B) B^b x_t,$$

where

$$\delta(B) = 1 - \delta_1 B - \delta_2 B^2 - \dots \delta_r B^r$$

$$\omega(B) = \omega_0 - \omega_1 B - \dots - \omega_s B^s$$

and b is a delay parameter that tells us the lag that elapses before the impulse of the input variable produces an effect on the output variable.

With the above representation, we can rearrange the terms so that y_t has an explicit representation in terms of x_t,

$$y_t = \frac{\omega(B)}{\delta(B)} x_{t-b} + \eta_t, \tag{2}$$

and by equating Equation (2) to Equation (1) we can write the transfer function $\nu(B)$ as

$$\nu(B) = \frac{\omega(B)B^b}{\delta(B)}. \tag{3}$$

and the orders s, r, b of the model in Equation (2) can be found by equating the coefficients of B^j to both sides in Equation (3)

$$\delta(B)\nu(B) = \omega(B)B^b,$$

which yields the following equation

$$(1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r)(\nu_0 + \nu_1 B + \dots) = (\omega - \omega_1 B - \dots - \omega_s B^s)B^b,$$

and we obtain the following relationships between the components of the model

$$\begin{aligned}
&\nu_j = 0 && \text{if } j < b \\
&\nu_j = \delta_1 \nu_{j-1} + \delta_2 \nu_{j-2} + \dots + \delta_r \nu_{j-r} + \omega_0 && \text{if } j = b \\
&\nu_j = \delta_1 \nu_{j-1} + \delta_2 \nu_{j-2} + \dots + \delta_r \nu_{j-r} - \omega_{j-b} && \text{if } j = b+1, \dots, b+s \\
&\nu_j = \delta_1 \nu_{j-1} + \delta_2 \nu_{j-2} + \dots + \delta_r \nu_{j-r} && \text{if } j > b+s
\end{aligned}$$

By observing the behaviour of the cross-correlation function between x_t and y_t – similarly to what we do with ACF and PACF for estimating p, d, q in an ARIMA model) – we can find the appropriate values of s, r, b.

Def. (Cross-correlation function)
We say that X_t and Y_t are ***jointly stationary*** if they are univariate stationary and $\text{Cov}(X_t, Y_s) = f(|s-t|)$, and in this case we define the ***cross-correlation function*** between X_t and Y_t as the function

$$\gamma_{XY}(k) = \mathbb{E}[(X_t - \mu_X)(Y_{t+k} - \mu_Y)],$$

Marginals By definition we have that $\rho_{XX}(k) = \rho_X(k)$.

Symmetry It's relevant the order in which we compute the cross-correlation function, since unlike the ACF the CCF is not symmetric around the origin,

$$\rho_{XY}(k) \neq \rho_{XY}(-k),$$

instead we have that

$$\rho_{XY}(k) \neq \rho_{YX}(-k).$$

However, we have a way of obtaining the direction of association between the time series by inspecting the graph of the ACF. The direction depends on the software implementation of the function.

Example (AR(1) model)
Let $Y_t \sim \mathrm{AR}(1)$, then we have $(1-\varphi B)Y_t = X_t$ and for time $t+k$ we can write

$$Y_{t+k} = \frac{1}{1-\varphi B} X_{t+k} = X_{t+k} + \varphi X_{t+k-1} + \varphi^2 X_{t+k-2} + \dots,$$

therefore the cross-covariance function between X_t and Y_t are

$$\gamma_{XY}(k) = \mathbb{E}[X_t Y_{t+k}] = \begin{cases} \varphi^k \sigma_k^2 & \text{if } k \geq 0 \\ 0 & \text{if } k \leq 0 \end{cases}$$

In general the ARMA(p,q) model can be written as a transfer function model without the white noise term η_t, and where X_t is a white noise itself.

3.2 Cross-correlation function and TF models

Let x_t and y_t be stationary series with $\mu_x = \mu_y = 0$, then the transfer function at time $t+k$ is

$$y_{t+k} = \nu_0 x_{t+k} + \nu_1 x_{t+k-1} + \nu_2 x_{t+k-2} + \dots + \eta_{t+k},$$

therefore if we multiply both left and right by x_t and take expectations we have

$$\gamma_{xy}(k) = \nu_0 \gamma_x(k) + \nu_1 \gamma_x(k-1) + \nu_2 \gamma_x(k-2) + \dots,$$

hence the CCF in the doubly stationary case has the following simple representation:

$$\rho_{xy}(k) = \frac{\sigma_x}{\sigma_y} \left[\nu_0 \rho_x(k) + \nu_1 \rho_x(k-1) + \nu_2 \rho_x(k-2) + \dots \right]. \tag{4}$$

Therefore, by Equation (4) we observe that the relationship between the CCF and IRF ν_j is contaminated by the fact that they are not white noise, and therefore display the correlations at previous times. However for a ***white noise model*** we would see $\rho_x(k) = 0$ for all $k \neq 0$ and therefore we would have a direct way of estimating ν_k by letting

$$\gamma_{xy}(k) = \nu_k \sigma_k^2,$$

hence we can estimate the covariance function nd obtain an impulse response function which is directly proportional to the CCF,

$$\rho_{xy}(k) = \frac{\sigma_x}{\sigma_y}\nu_k \implies \nu_k = \frac{\sigma_y}{\sigma_x}\rho_{xy}(k). \tag{5}$$

Idea Therefore, our goal for estimating a TF model is to reduce the problem to a whitened series for x_t, and then apply the estimation procedure above.

In the general TF model given by

$$y_t = \nu(B)x_t + \eta_t,$$

if we assume $x_t \sim \text{ARMA}(p, q)$ we can calculate the ***pre-whitened input series***

$$\alpha_t = \frac{\varphi_x(B)}{\vartheta_x(B)}x_t,$$

and applying this transformation to both y_t and η_t we can obtain the ***filtered series***

$$\begin{cases} \beta_t = \frac{\varphi_x(B)}{\vartheta_x(B)}y_t \\ \varepsilon_t = \frac{\varphi_x(B)}{\vartheta_x(B)}\eta_t \end{cases}$$

Finally, the TF model becomes

$$\beta_t = \nu(B)\alpha_t + \varepsilon_t,$$

where the input series is $\alpha_t \sim \text{WN}(0, \sigma^2)$ and we can estimate the transfer function using Equation (5) between β_t and α_t.

3.2.1 General procedure for the identification of a TF model

1. Identify an ARMA(p, q) model for the input x_t,

$$\varphi_x(B)x_t = \vartheta_x(B)\alpha_t$$

2. Prewhiten $x_t \to \alpha_t = \frac{\varphi_x(B)}{\vartheta_x(B)}x_t$ and apply the same filter to $y_t \to \beta_t = \frac{\varphi_x(B)}{\vartheta_x(B)}y_t$.
3. Calculate the CCF between the whitened input series and the residuals of the model for y_t,

$$\widehat{\nu}_k = \frac{\widehat{\sigma}_\beta}{\widehat{\sigma}_\alpha}\ldots,$$

   to get a preliminary estimation of the transfer function ν_k.
4. Identify the order b, r, s of the TF model by inspecting the estimated TF (or equivalently the CCF) and estimate the transfer function using the fact that

$$\widehat{\nu}_j = \frac{\widehat{\omega}(B)}{\widehat{\delta}(B)}B^b,$$

   which is of course done by nonlinear least squares or other methods.

5. Identify a model for the estimated residuals $\widehat{\eta}_t$ given by

$$\widehat{\eta}_t = y_t - \widehat{\nu}(B)x_t.$$

6. Estimate the model and check goodness-of-fit, generally by checking both $\widehat{\varepsilon}_t$ and $\widehat{\alpha}_t$ are white noise. Moreover, since we assume that $\varepsilon_t \sim \mathrm{WN}$ and $\eta_t \perp\!\!\!\perp x_t$. we need to check that $\widehat{\rho}_{\alpha,\widehat{\varepsilon}}(k)$ is non significant.

For checking the last step, there are test statistics which are based on Portmanteau tests.

LECTURE 4: SPECTRAL ANALYSIS

2021-12-13

In this lecture we introduce spectral analysis, which transforms the data from the time domain to the frequency domain by decomposing the time series into a Fourier basis of coefficients. The idea is to decompose X_t in terms of combination of sinusoids with random and uncorrelated coefficients.

4.1 Periodicity

Consider a periodic process of the form

$$X_t = C \cdot \cos(2\pi\omega t + \varphi), \quad t = \pm 1, \pm 2, \dots,$$

where ω is a *frequency* index, C the *amplitude* and φ the *phase* of the process. We can introduce random variation in X_t by allowing the amplitude and phase to vary, since

$$X_t = A\cos(2\pi\omega t) + B\sin(2\pi\omega t),$$

where $A = C\cos\varphi$ and $B = -C\sin\varphi$ and we can choose them $A, B \sim \mathcal{N}(0, \sigma^2)$. We have that

1. $C = \sqrt{A^2 + B^2}$
2. $\varphi = \tan^{-1}(-B/A)$
3. X_t is a stationary process with $\mu_t = 0$ and

$$\gamma(h) = \mathrm{Cov}(X_t, X_{t+h}) = \sigma^2 \cos(2\pi\omega h).$$

We consider a generalization as a mixture of periodic series with multiple frequencies and amplitudes,

$$X_t = A_0 + \sum_{i=1}^{q} \left\{A_i \cos(2\pi\omega_i t) + B_i \sin(2\pi\omega_i t)\right\}, \tag{6}$$

where $A_i, B_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$ and the ω_i are distinct frequencies. In this case,

$$\gamma(0) = \sum_{i=1}^{q} \sigma_i^2, \quad \gamma(h) = \sum_{i=1}^{q} \sigma_i^2 \cos(2\pi\omega_i t).$$

The main objective of time series is to sort out the essential frequency components of a time series, including their relative contribution.
For a sample $x_1, \dots, x_n$ from X_t we can write the following representation

$$X_t = A_0 + \sum_{j=1}^{\frac{n-1}{2}} \left\{A_j \cos(2\pi tj/n) + B_j \sin(2\pi tj/n)\right\}, \tag{7}$$

for $t = 1, 2, \dots, n$ and suitably chosen coefficients. If n is even we can modify the above equation and an additional component. Equation (7) holds for any sample and can be interpreted as an approximation to (6) with some coefficients possibly close to zero.

Our problem is now to estimate the A_j's and B_j's using a linear model using the frequencies which are relevant to the observed model. We do so by plotting the ***periodogram***, i.e. the estimates of the variance explained by the j^{th} component $P(j/n) = \frac{1}{2}(\widehat{A}_j^2 + \widehat{B}_j^2)$.

By inspecting the periodogram we can observe which frequencies $\omega_j = j/n$ are predominant over the others and eventually observe frequencies which are "hidden" inside the time series.

Theorem 2 (Parseval's theorem)
The sample variance is the sum of the contribution of the observed periodogram

$$\frac{1}{n}\sum_{t=1}^{n}(X_t - \overline{X})^2 = \frac{1}{2}\sum_{j=1}^{\frac{n-1}{2}}(A_j^2 + B_j^2) = \sum_{j=1}^{\frac{n-1}{2}} P(j/n)$$