INTRODUCTION TO TRANSFER FUNCTION NOISE MODEL

Transfer function model is a statistical model describing the relationship between an output variable Y_t and one or more input variables X_t. Let us start with a deterministic case that the dynamic dependence of Y_t on the current and past values of X_t , namely $\{X_{t-j}\}_{j=0}^{\infty}$, can be written as

$$Y_t = \nu_0 X_t + \nu_1 X_{t-1} + \nu_s X_{t-2} + \cdots = \nu(B) X_t, \qquad (1)$$

where $\nu_0, \nu_1, \cdots$ are constant denoting the impact of X_{t-j} on Y_t, and $\nu(B) = \nu_0 + \nu_1 B + \nu_2 B^2 + \cdots$ with B denoting the backshift operator such that $BX_t = X_{t-1}$.

1. The coefficients $\nu_0, \nu_1, \cdots$ are referred to as the impulse response function of the system. For the model in eqn. (1) to be *meaningful*, the impulse responses must be are absolutely summable, i.e., $\sum_{j=0}^{\infty} |\nu_j| < \infty$.
2. In this case, the system is said to be stable. The value

$$g = \sum_{j=0}^{\infty} \nu_j$$

   is called the stead-state gain as it represents the impact on Y when X_{t-j} are held constant over time.

3. Note that the function $\nu(B)$ is infinite order polynomials in B and in practice may be expressed as a rational polynomial in B such as

$$\nu(B) = \frac{\omega(B) B^b}{\delta(B)}, \qquad (2)$$

   where b is a non-negative integer and called the time delay of the system, $w(B) = \omega_0 + \omega_1 B + \cdots + \omega_s B^s$ and $\delta(B) = 1 - \delta_1 B - \cdots - \delta_r B^r$ are finite order polynomials in B, and $\omega_0 \neq 0$.

In practice, the output Y_t is not a deterministic function of X_t. It is often disturbed by some noise, say N_t. The noise may be serially correlated and is usually assumed that N_t follows a stationary and inverible $\mathrm{ARMA(p, q)}$ model, i.e.

$$\phi(B) N_t = \theta(B) a_t, \qquad (3)$$

where $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ and $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ are polynomials in B of degree p and q, respectively, and $\{a_t\}$ is a sequence of NID random variable with mean zero and variance σ_a^2.

Putting together, we have a transfer-function noise model as

$$Y_t - \mu_y = \nu(B)(X_t - \mu_x) + N_t, \qquad (4a)$$

$$= \frac{\omega(B)B^b}{\delta(B)}(X_t - \mu_x) + \frac{\theta(B)}{\phi(B)}a_t, \qquad (4b)$$

where $\mathrm{E}(\mathrm{Y_t}) = \mu_\mathrm{y}$ and $\mathrm{E}(\mathrm{X_t}) = \mu_\mathrm{x}$, $\theta(\mathrm{B}), \phi(\mathrm{B}), \omega(\mathrm{B})$ and $\delta(\mathrm{B})$ are defined as before with degree $\mathrm{q}, \mathrm{p}, \mathrm{s}$ and r, respectively, and $\{\mathrm{a_t}\}$ are Gaussian white noise series.

- <u>Remark 1:</u> The noise term N_t should be independent of X_t. Otherwise, the model is not identifiable.
- <u>Remark 2:</u> If $b > 0$, the transfer function model is useful in predicting the turning points of Y_t given those of X_t.
- <u>Remark 3:</u> When there are multiple input variables, say two, the transfer function noise model becomes

$$Y_t - \mu_y = \frac{\omega_1(B)B^{b_1}}{\delta_1(B)}(X_{1t} - \mu_{x1}) + \frac{\omega_2(B)B^{b_2}}{\delta_2(B)}(X_{2t} - \mu_{x2}) + \frac{\theta(B)}{\phi(B)}a_t,$$

where $\omega_i(B), \delta_i(B), i = 1,2$ are similarly defined as in eqn. (4).

For simplicity and without loss of generality, we shall introduce the model building process using the single input TFN model.

<u>MODEL BUILDING PROCESS</u>

The procedure of building the single input TFN model includes

1. Preliminary identification of the impulse response coefficients ν_i's
2. Specification of the noise term $\mathrm{N_t}$
3. Specification of transfer function
4. Estimation of the TFN model specified in Step 2 and 3
5. Model diagnostic checks

1. Preliminary estimation of $\nu(\mathrm{B})$

For a single input TFN model, the regression

$$Y_t = c + \nu_0 X_t + \nu_1 X_{t-1} + \cdots + \nu_h X_{t-h} + e_t,$$

would not provide consistent estimates of ν_i's in general, where h is a large positive integer, since X_t and N_t might be serially dependent. Therefore, Box and Jenkins (1970) suggests the following procedure to identify (B) .

1) Suppose that X_t follows an $ARMA$ model

$$\phi_x(B)(X_t - \mu_x) = \theta_x(B)\alpha_t,$$

where $\{\alpha_t\}$ is a sequence of white noises. Applying the operator $\phi_x(B)/\theta_x(B)$ on both sides of eqn. (4a), we have

$$\beta_t = \nu(B)\underbrace{\frac{\phi_x(B)}{\theta_x(B)}(X_t - \mu_x)}_{\alpha_t} + \frac{\phi_x(B)}{\theta_x(B)}N_t = \nu(B)\alpha_t + n_t, \quad (5)$$

where $\beta_t = \phi_x(B)\theta^{-1}(B)(Y_t - \mu_y)$ and $n_t = \phi_x(B)\theta^{-1}(B)N_t$. It is worth noting that $\{n_t\}$ is independent of $\{\eta_t\}$ and η_t is a white noise series.

2) Multiplying both sides of eqn. (5) by α_{t-j} for $j \geq 0$, we have

$$\beta_t\alpha_{t-j} = \nu(B)\alpha_t\alpha_{t-j} + n_t\alpha_{t-j}.$$

Taking expectation, we have

$$cov(\beta_t, \alpha_{t-j}) = \nu_j \cdot var(\alpha_{t-j}).$$

Consequently, we have

$$\nu_j = \frac{cov(\beta_t, \alpha_{t-j})}{var(\alpha_t)} = corr(\beta_t, \alpha_{t-j}) \cdot \frac{se(\beta_t)}{se(\alpha_t)}.$$

2. Specification of the noise term N_t

N_t can be estimated using

$$\widehat{N}_t = (Y_t - \bar{Y}) - \hat{\nu}(B)(X_t - \bar{X}),$$

where $\bar{Y}$ and $\bar{X}$ are the sample means for μ_y and μ_x, respectively. By examining graphs such as the sample ACF and the sample of PACF of $\widehat{N}_t$, identify the $ARMA$ model needed to fit the noise series.

3. Specification of transfer function

The goal is to find the rational polynomials $\omega(B)$ and $\delta(B)$ as well as the delay parameter b to best approximate $\nu(B)$. From

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

we have

$$\nu_0 + \nu_1 B + \nu_2 B^2 + \cdots = \frac{\omega_0 B^b + \omega_1 B^{b+1} + \cdots + \omega_s B^{b+s}}{1 - \delta_1 B - \cdots - \delta_r B^r}.$$

By equating the coefficients of B^j, it is easy to show that

1) $\nu_j = 0$ for $j < b$ if b is positive;
2) $\nu_b, \nu_{b+1}, \cdots, \nu_{b+s-r+1}$ follow no fixed pattern (if $s < r$)
3) ν_j with $j \geq b + s - r + 1$ follows a rth order difference equation (if $s \geq r$)

$$\nu_j = \delta_1 \nu_{j-1} + \cdots + \delta_r \nu_{j-r} \;\; or \;\; \delta(B)\nu_j = 0, \qquad (6)$$

with starting values $\nu_{b+s}, \cdots, \nu_{b+s-r+1}$.

? • Remark 4: We could also use the Corner method, which is based on the Pade approximation of a polynomial.

ESTIMATION

Our goal is to estimate the following equation

$$y_t - \mu_y = \frac{\omega(B)}{\delta(B)} B^b (X_t - \mu_x) + N_t.$$

1) The theoretical means of y_t and x_t can be estimated as $\hat{\mu}_y = \sum y_t / n$ and $\hat{\mu}_x = \sum x_t / n$
2) Let $x_t = X_t - \hat{\mu}_x$ and consider the output d_t from the dynamic component as

$$d_t = \frac{\omega(B)}{\delta(B)} B^b x_t = \frac{\omega(B)}{\delta(B)} x_{t-b}$$

or more specifically,

$$d_t = \delta_1 d_{t-1} + \cdots + \delta_r d_{t-r} + \omega_0 x_{t-b} - \omega_1 x_{t-b-1} - \cdots \omega_m x_{t-b-m},$$

and the noise is determined using

$$N_t = \left(y_t - \hat{y}_y\right) - d_t$$

3) Then, the a′s can be obtained from

$$a_t = \frac{\phi(B)}{\theta(B)} N_t,$$

or more specifically,

$$a_t = \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q} + N_t - \phi_1 N_{t-1} - \phi_2 N_{t-2} - \cdots \phi_p N_{t-p}.$$

4) Finally, assumed that $a_t \sim NID(0, \sigma^2)$ and one can employ MLE or other ARMA filters, such as McLeod (1977) algorithm for the estimation of model parameters.

Diagnostic checking

1) Cross-correlation check: to check whether the noise $\{a_t\}$ and the input series $\{x_t\}$ are uncorrelated.

$$Q_0 = m(m+2)\sum_{j=0}^{K}(m-j)^{-1}\,\hat{\rho}_{\alpha a}^2(j) \sim \chi_{K+1-M}^2$$

where $m = n - t_0 + 1$ is the number of residuals $\hat{a}_t$ calculated, and M is the number of parameters δ_i and ω_i estimated in the transfer function $\nu(B) = \omega(B)/\delta(B)$. Note that the number of degrees of freedom for Q_0 is independent of the number of parameters estimated in the noise model.

2) Autocorrelation check: to check whether the noise model is adequate. A portmanteau test for $ARMA$ models can be used.

$$Q_1 = m(m+2)\sum_{j=1}^{K}(m-j)^{-1}\,\hat{\rho}_a^2(j).$$

The Q_1 statistic approximately follows a χ^2 distribution with $(k-p-q)$ degrees of freedom depending only on the number of parameters in the noise model.

PRACTICE QUESTIONS

1. Box, Jenkins, and Reinsel (1994) fit a transfer function model to data from a gas furnace. The input variable x_t is the volume of methane entering the chamber in cubic feet per minute and the output is the concentration of carbon dioxide emitted y_t. The transfer function model is

$$y_t = \frac{-(0.53 + 0.37B + 0.51B^2)}{1 - 0.57B} x_t + \frac{1}{1 - 0.53B + 0.63B^2} \varepsilon_t$$

where the input and output variables are measured every nine seconds.

a) What are the value of b, s, and r for this model?

Answer: $b = 0$ (no lag), $r = 1$ (one term in the denominator) and $s = 2$ (two terms in the numerator)

b) What is the form of the *ARIMA* model for the errors?

Answer: There are no terms in the numerator and two in the denominator so the errors follow an $AR(2)$ model.

c) If the methane input was increased, how long would it take before the carbon dioxide concentration in the output is impacted?

Answer: Since there is no lag, the carbon dioxide concentration in the output is impacted immediately.

2. An input and output time series consists of 2500 observations. The prewhitened input series is modeled by an AR(2) model $y_t = 0.4y_{t-1} + 0.2y_{t-2} + a_t$. Suppose that we have estimated $\hat{\sigma}_\alpha = 0.3$ and $\hat{\sigma}_\beta = 0.35$. The estimated cross-correlation function between the prewhitened input and output time series is shown below.

Lag	0	1	2	3	4	5	6	7	8	9	10
$\rho_{\alpha\beta}$	0.01	0.03	-0.03	0.5	-0.4	-0.09	-0.05	-0.03	-0.02	0.09	-0.01

a) Find the approximate standard error of the cross-correlation function. [5%]

Answer: $1/\sqrt{2500} = 0.02$

b) Which spikes on the cross-correlation function appear to be significant? [5%]

Answer: Lag 3,4,5,7 and 7 are significant

Lag	1	2	3	4	5	6	7	8	9
$\gamma_{\alpha\beta}(j)$	0.01	0.03	-0.03	0.5	-0.4	-0.09	-0.05	-0.03	-0.02
v_j	0.02	0.035	-0.035	0.583	-0.467	-0.105	-0.058	-0.035	-0.0233
lowerCI	-0.03	-0.01	-0.07	0.46	-0.44	-0.13	-0.09	-0.07	-0.06
UpperCI	0.05	0.07	0.01	0.54	-0.36	-0.05	-0.01	0.01	0.02

c) Estimate the impulse response function. Tentatively identify the form of the transfer function models—i.e. the values of b, r, s in a transfer function model.

Answer: $b = 3$, $s = 4$, and $r = 0$ (a mistake in the cross-correlation)

3. An input and output time series consists of 300 observations. The prewhitened input series is modeled by an $AR(2)$ model $y_t = 0.5y_{t-1} + 0.2y_{t-2} + \alpha_t$. Suppose that we have estimated $\hat{\sigma}_\alpha = 0.2$ and $\hat{\sigma}_\beta = 0.4$. The estimated cross-correlation function between the prewhitened input and output time series is shown below.

Lag	0	1	2	3	4	5	6	7	8	9	10
$\rho_{\alpha\beta}$	.01	.03	-.03	-.25	-.35	-.51	-.3	-.15	-.02	.07	-.02

a) [5 %] Which spikes on the cross-correlation function appear to be significant?

b) [10%] Suppose that $r = s = 2$. Provide preliminary estimates on w_0, w_1, w_2, δ_1 and δ_2 for the impulse response function as follows:

$$v(B) = \sum_{i=0}^{\infty} v_i B^i = \frac{w_0 - w_1 B - w_2 B^2}{1 - \delta_1 B - \delta_2 B^2} B^b.$$

(Hint: the parameter b should be determined first)

a) **Answer:** [Marking scheme: one point for each lag]

The variance of the sample cross correlation function is approximately $\frac{1}{\sqrt{300}} \cong 0.0577$ and use the result $\hat{v}_{\alpha\beta}(j) = corr(\beta_t, \alpha_{t-j})\frac{se(\beta_t)}{se(\alpha_t)}$. We have the following table. It is shown in the table that the CIs of lags 3,4,5,6 and 7 [one point each] do not contain zero so these spikes are significant statistically.

lag	0	1	2	3	4	5	6	7	8	9	10
$\rho_{\alpha\beta}$	0.01	0.03	-0.03	-0.25	-0.35	-0.51	-0.3	-0.15	-0.02	0.07	-0.02
$v_{\alpha\beta}$	0.02	0.06	-0.06	**-0.5**	**-0.7**	**-1.02**	**-0.6**	**-0.3**	-0.04	0.14	-0.04
LowerCI	-0.11	-0.09	-0.15	-0.37	-0.47	-0.63	-0.42	-0.27	-0.14	-0.05	-0.14
UpperCI	0.13	0.15	0.09	-0.13	-0.23	-0.39	-0.18	-0.03	0.10	0.19	0.10

b) **Answer:** [Marking scheme: 2 point for answering the correct value of w_0, w_1, w_2, δ_1 and δ_2]

The first nonzero $\rho_{\alpha\beta}$ starts at lag $3 > 0$ so $b = 3$. Thus, the transfer function is

$$y_t = \frac{w_0 - w_1 B - w_2 B^2}{1 - \delta_1 B - \delta_2 B^2} x_{t-3} = v(B)x_t$$

By matching the coefficients of $\{v_i\}$, $\{w_i\}$, and $\{\delta_i\}$, we have

$v_0 = v_1 = v_2 = 0$

$v_3 = w_0$ ➔ $w_0 = -0.5$

$v_4 = \delta_1 v_3 + \delta_2 v_2 - w_1$ ➔ $-0.7 = -0.5\delta_1 - w_1,$ (1)

$v_5 = \delta_1 v_4 + \delta_2 v_3 - w_2$ ➔ $-1.02 = -0.7\delta_1 - 0.5\delta_2 - w_2,$ (2)

$v_6 = \delta_1 v_5 + \delta_2 v_4$ ➔ $-0.6 = -1.02\delta_1 - 0.7\delta_2,$ (3)

$v_7 = \delta_1 v_6 + \delta_2 v_5$ ➔ $-0.3 = -0.6\delta_1 - 1.02\delta_2,$ (4)

Using the above result, we can solve for $w_1, w_2, \delta_1, \delta_2$ (four equations and four unknown), and obtain the following results

$w_0 = -0.5, w_1 = 0.376, w_2 = 0.610, \delta_1 = 0.648, \delta_2 = -0.087.$