Handout 10

Transfer Function Models

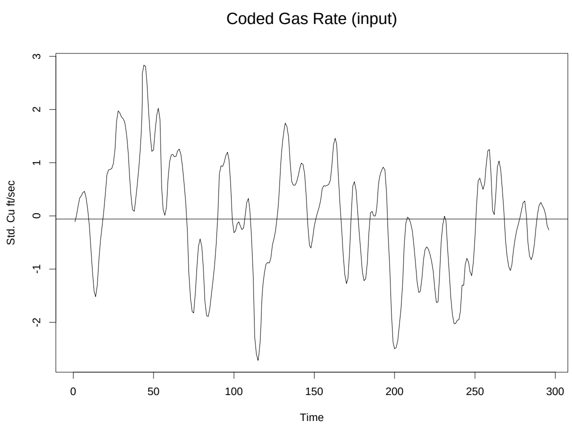
Class notes for Statistics 451: Applied Time Series
Iowa State University

March 29, 2006
18h 20min

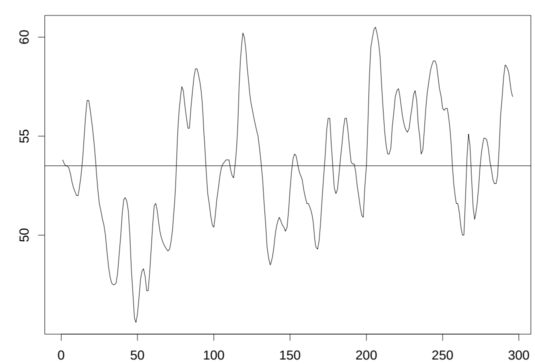
Transfer Function Models

- Single "response" time series and one or more "explanatory" (or "input") time series.
- Allows the use of explanatory variables to explain variability in the "response" time series.
- Similar to regression analysis (sometimes called "dynamic regression analysis").
- For forecasting, most effective when the explanatory time series is a "leading indicator" for the response time series
- For multiple response with feedback, needed generalization is "multiple time series."

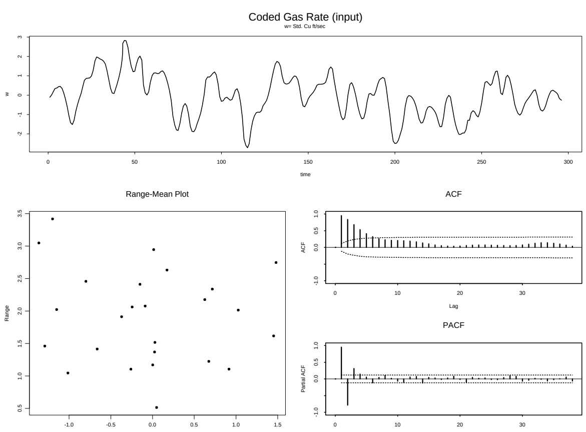
Gas Furnace Input Rate (cu. ft/min)



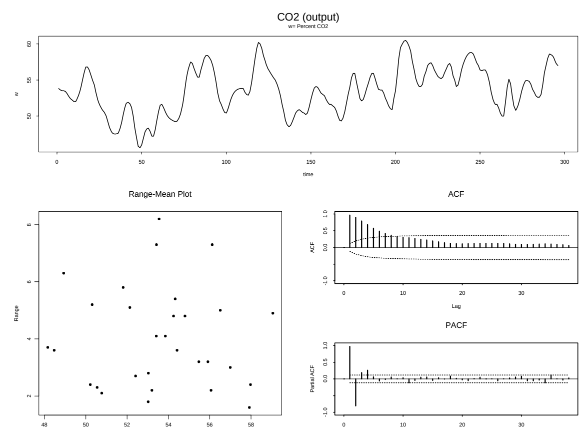
Gas Furnace Percent CO_2 in Outlet Gas



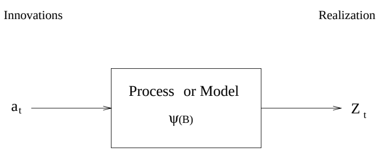
Gas Furnace Input Rate (cu. ft/min)



Gas Furnace Percent CO_2 in Outlet Gas



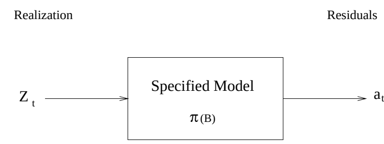
Innovation to Realization Filter



Model:

$$\begin{aligned} Z_t &= \psi(\mathsf{B})a_t \\ &= (1 + \mathsf{B}\psi_1 + \mathsf{B}^2\psi_2 + \cdots)a_t \\ Z_t &= a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots \end{aligned}$$

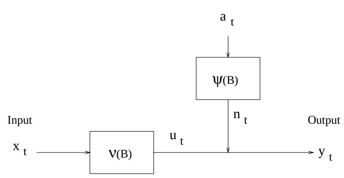
Realization to Residual Filter



Model:

$$\begin{aligned} a_t = \frac{1}{\psi(\mathsf{B})} &= \pi(\mathsf{B})Z_t \\ &= (1 - \mathsf{B}\pi_1 - \mathsf{B}^2\pi_2 - \cdots)Z_t \\ a_t &= Z_t - \pi_1 Z_{t-1} - \pi_2 Z_{t-2} - \cdots \end{aligned}$$

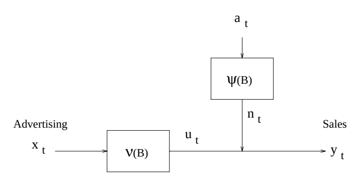
Transfer Function (Dynamic Regression) Model



Model:

$$\begin{aligned} y_t &= u_t + n_t \\ &= \nu(\mathsf{B})x_t + \psi(\mathsf{B})a_t \\ &= (\nu_0 + \nu_1\mathsf{B} + \nu_2\mathsf{B}^2 + \cdots)x_t + (1 + \psi_1\mathsf{B} + \psi_2\mathsf{B}^2 + \cdots)a_t \\ &= \nu_0 x_t + \nu_1 x_{t-1} + \nu_2 x_{t-2} + \cdots \\ &\quad + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots + a_t \end{aligned}$$

Dynamic Regression Model for Effect of Advertising on Sales



Model:

$$y_t = \nu(\mathsf{B})x_t + \psi(\mathsf{B})a_t$$

Notes on the Transfer Function Model

- Input x_t can, depending on the application, be either controlled or stochastic.
- Input x_t can, depending on the application, be continuous, discrete, or binary (as in the intervention models).
- There can be more than one input variable.

Some Dynamic Regression Models for CO_2 in Outlet Gas

AR(9):

$$y_t = \theta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_9 y_{t-9} + a_t$$

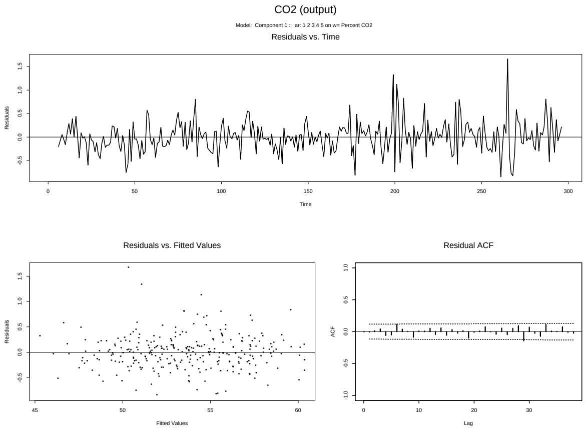
Regression on present and past values of input

$$y_t = \theta_0 + \nu_0 x_t + \nu_1 x_{t-1} + \cdots + \nu_k x_{t-9} + a_t$$

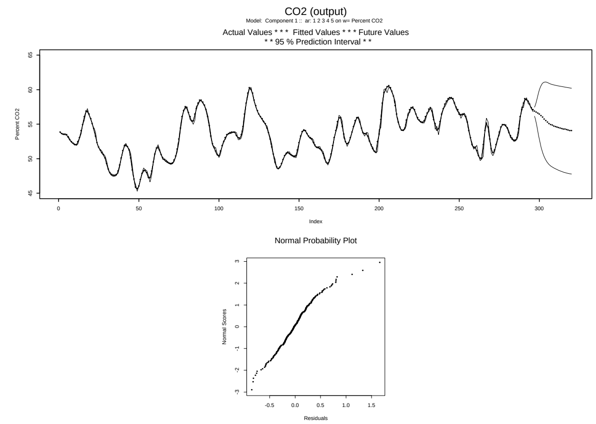
Regression on past output as well as present and past values of input:

$$\begin{aligned} y_t = \; & \theta_0 + \nu_0 x_t + \nu_1 x_{t-1} + \cdots + \nu_9 x_{t-9} \\ & + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_9 y_{t-9} + a_t \end{aligned}$$

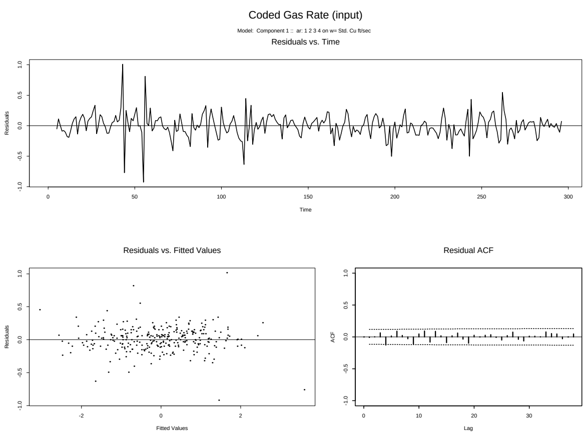
Univariate AR(5) Model for Gas Furnace CO_2 in Outlet Gas, Part 1



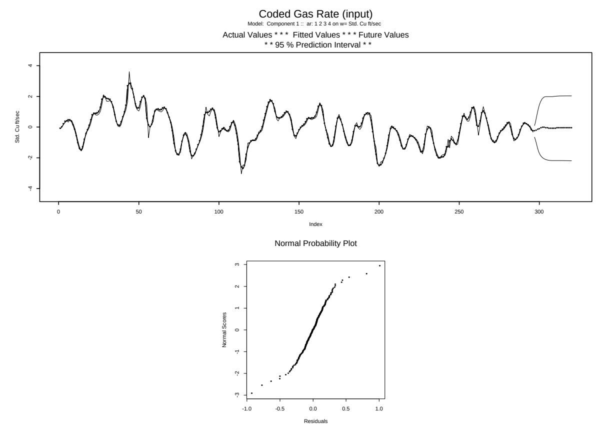
Univariate AR(5) Model for Gas Furnace CO_2 in Outlet Gas, Part 2



Univariate AR(4) Model for Gas Furnace Input Rate, Part 1



Univariate AR(4) Model for Gas Furnace Input Rate, Part 2



Cross Correlation Function

The cross covariance function

$$\gamma_{xy}(k) = \mathsf{E}[(x_t - \mu_x)(y_{t+k} - \mu_y)]$$

describes, for $k > 0$, the relationship between future values of y and current values of x.

The cross correlation function

$$\rho_{xy}(k) = \frac{\gamma_{xy}(k)}{\sigma_x \sigma_y}$$

gives the correlation between x_t and y_{t+k} (or, equivalently, between x_{t-k} and y_t).

To see how present values of x might be related to past values of y, use

$$\rho_{xy}(-k) = \rho_{yx}(k)$$

Data for Computing Sample Cross CorrelationsBetween x_t and y_{t+k}

t	x_t	y_t	y_{t+1}	y_{t+2}	y_{t+3}	y_{t+4}	y_{t+5}
1	-0.109	53.8	53.6	53.5	53.5	53.4	53.1
2	0.000	53.6	53.5	53.5	53.4	53.1	52.7
3	0.178	53.5	53.5	53.4	53.1	52.7	52.4
4	0.339	53.5	53.4	53.1	52.7	52.4	52.2
5	0.373	53.4	53.1	52.7	52.4	52.2	52.0
6	0.441	53.1	52.7	52.4	52.2	52.0	52.0
7	0.461	52.7	52.4	52.2	52.0	52.0	52.4
8	0.348	52.4	52.2	52.0	52.0	52.4	53.0
9	0.127	52.2	52.0	52.0	52.4	53.0	54.0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
294	0.017	57.8	57.3	57.0	NA	NA	NA
295	-0.182	57.3	57.0	NA	NA	NA	NA
296	-0.262	57.0	NA	NA	NA	NA	NA

Sample Cross Correlation Function (CCF)

The sample cross correlation function

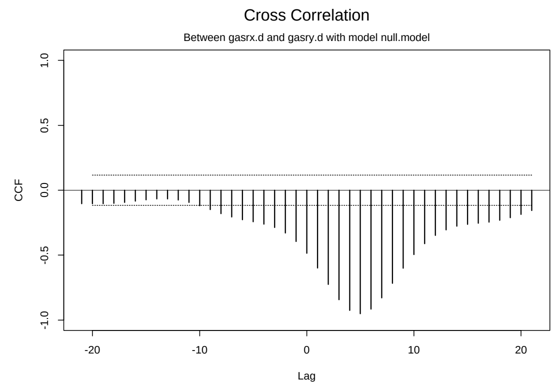
$$\widehat{\rho}_{xy}(k) = \frac{\widehat{\gamma}_{xy}(k)}{S_x S_y}$$

gives the sample correlation between x_t and y_{t+k} (or, equivalently, between x_{t-k} and y_t).

To see how x_t is correlated with past values y_t, use

$$\widehat{\rho}_{xy}(-k) = \widehat{\rho}_{yx}(k)$$

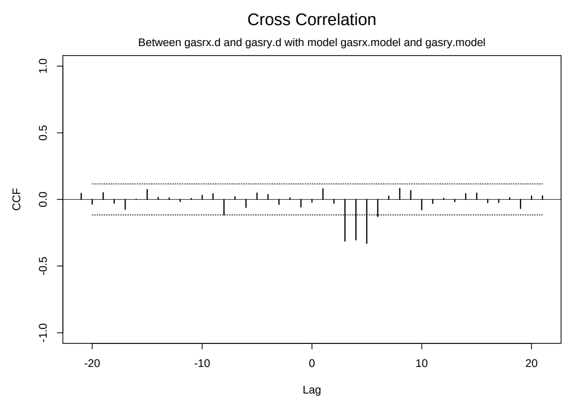
Cross Correlation Function Between Gas Furnace Input Rate and CO_2 in Outlet Gas



Notes on Cross Correlation Function

- Both x_t and y_t must be stationary for cross correlations to be defined.
- To judge whether a sample CCF value is significantly different from 0, use
$$t = \frac{\widehat{\rho}_{xy}(k)}{\mathsf{S}_{\widehat{\rho}_{xy}(k)}}$$
where
$$\mathsf{S}_{\widehat{\rho}_{xy}(k)} = \sqrt{\frac{1}{n-k}} \approx \sqrt{\frac{1}{n}}$$
Compare with standard normal quantiles (signal if outside ± 2).
- As with ACF and PACF, commonly used tests and standard errors for the CCF are only approximate.
- The sample CCF is difficult to interpret unless either x_t or y_t is white noise.

Cross Correlation Function Between Residuals of Gas Furnace Input Rate and Residuals of CO_2 in Outlet Gas



Interpreting the Cross Covariance Function

The transfer function model can be written as

$$y_{t+k} = \nu_0 x_{t+k} + \nu_1 x_{t+k-1} + \nu_2 x_{t+k-2} + \cdots + \nu_k x_t + n_{t+k}$$

For illustration, assume that $\mu_y = 0$ and $\mu_x = 0$. Multiplying the model by x_t,taking expectations, and assuming x_t and n_t are independent gives the lag-k cross covariance function as

$$\begin{aligned}
\gamma_{xy}(k) &= \mathsf{E}(y_{t+k}x_t) \\
&= \nu_0 \mathsf{E}(x_{t+k}x_t) + \nu_1 \mathsf{E}(x_{t+k-1}x_t) + \nu_2 \mathsf{E}(x_{t+k-2}x_t) \\
&\quad + \cdots + \nu_k \mathsf{E}(x_t x_t) + \mathsf{E}(n_{t+k}x_t) \\
&= \nu_0\gamma_k + \nu_1\gamma_{k-1} + \nu_2\gamma_{k-2} + \cdots + \nu_k\gamma_0
\end{aligned}$$

where the γ_k values here are the autocovariances of the x_t time series. If x_t is <u>white noise</u>, then $\gamma_k = 0$ for $k \neq 0$, and the cross covariance function at lag k simplifies to

$$\gamma_{xy}(k) = \mathsf{E}(y_{t+k}x_t) = \nu_k\gamma_0 = \nu_k\sigma_x^2$$

Prewhitening Transfer Function Inputs

$$y_t = \nu(\mathsf{B})x_t + n_t$$

If α_t is a "white-noise" process, the model for x_t is

$$x_t = \psi_x(\mathsf{B})\alpha_t = \frac{\theta_x(\mathsf{B})}{\phi_x(\mathsf{B})}\alpha_t \quad \text{or} \quad \alpha_t = \pi_x(\mathsf{B})x_t = \frac{\phi_x(\mathsf{B})}{\theta_x(\mathsf{B})}x_t$$

Filter y_t (i.e., find the "residuals" of y_t using the x_t model)

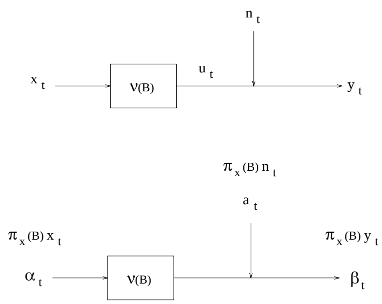
$$\beta_t = \pi_x(\mathsf{B})y_t = \frac{\phi_x(\mathsf{B})}{\theta_x(\mathsf{B})}y_t$$

Because $\pi_x(\mathsf{B})$ is the "wrong" model for y_t, we do not expect the β_t series to be white noise. Multiplying the original model by $\pi_x(\mathsf{B})$ gives

$$\begin{aligned}
\pi_x(\mathsf{B})y_t &= \pi_x(\mathsf{B})\nu(\mathsf{B})x_t + \pi_x(\mathsf{B})n_t \\
\beta_t &= \nu(\mathsf{B})\alpha_t + \pi_x(\mathsf{B})n_t
\end{aligned}$$

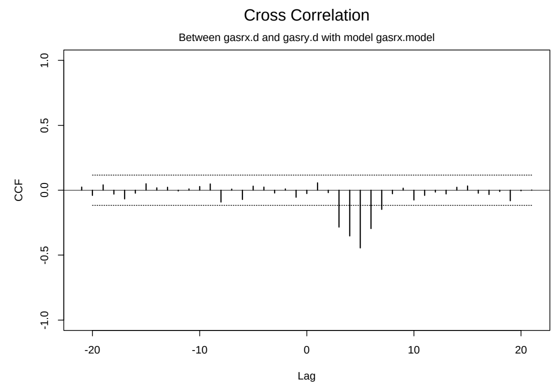
which is a white-noise input process, allowing easy identification of $\nu(\mathsf{B})$.

Use of a Prewhitening Filter to Identify ν(B)

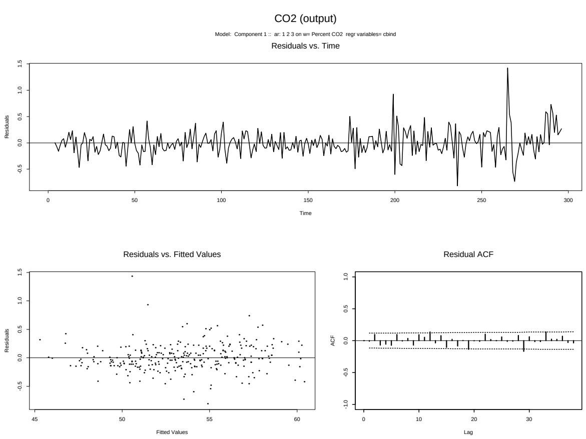


$$\begin{aligned} y_t &= \nu(\text{B})x_t + n_t \\ \pi_x(\text{B})y_t &= \pi_x(\text{B})\nu(\text{B})x_t + \pi_x(\text{B})n_t \\ \beta_t &= \nu(\text{B})\alpha_t + \pi_x(\text{B})n_t \end{aligned}$$

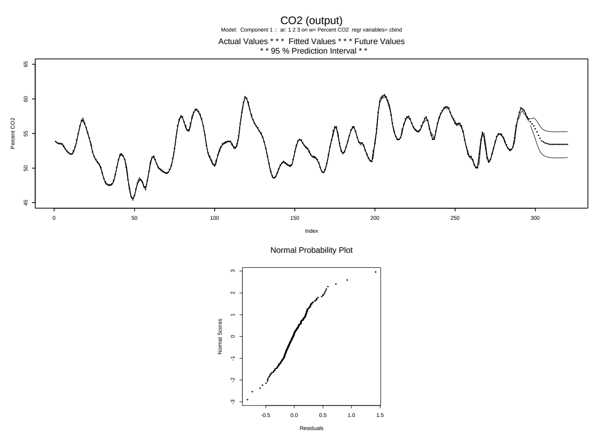
Cross Correlation Function Between Prewhitened Gas Furnace Input Rate and CO_2 in Outlet Gas



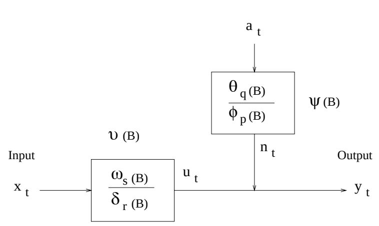
Transfer Function Model for Gas Furnace CO_2 in Outlet Gas, Part 1



Transfer Function Model for Gas Furnace CO_2 in Outlet Gas, Part 2



Parsimonious Transfer Function Models



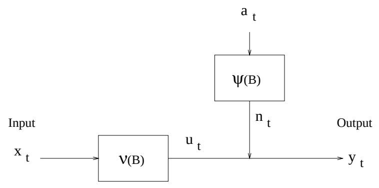
$$\begin{aligned} y_t &= \nu(\text{B})x_t + n_t \\ &\approx \frac{\omega_s(\text{B})}{\delta_r(\text{B})}x_t + \frac{\theta_q(\text{B})}{\phi_p(\text{B})}a_t \end{aligned}$$

where s, r, q, and p are small (0, 1 or 2) and $s + r + q + p$ is typically 3 or 4.

Strategy for Identifying a Transfer Function Model

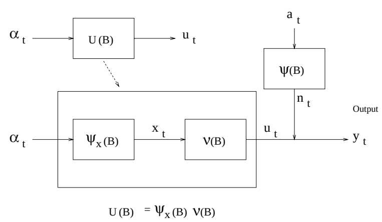
- Model y_t alone as a base-line for comparison
- Model x_t alone
- Use the x_t model to "filter" both x_t and y_t; examine the CCF
- Identify tentative dynamic regression model from CCF
- Fit and check the tentative dynamic regression model
- For a multiple input transfer function model, add new variables by using new x_t to help explain the residuals from the current model.

Forecasting Transfer Function Output with Fixed (Known) Input x_t (so that u_t is deterministic)



- Compute u_t from known x_t and $\nu(\text{B})$ from $u_t = \nu(\text{B})x_t$.
- Use the univariate method to find a prediction interval for n_t.
- Because $y_t = u_t + n_t$, forecasts and intervals for y_t can be obtained by adding u_t to the forecasts and intervals for n_t.

Forecasting Transfer Function Output with Stochastic Input x_t (so that u_t is stochastic)



$$
\begin{aligned}
y_t &= u_t + n_t \\
&= u(\text{B})\alpha_t + \psi(\text{B})a_t \\
\text{Var}[e_n(l)] &= \sigma_\alpha^2 \sum_{j=1}^{l-1} u_j^2 + \sigma_a^2 \sum_{j=1}^{l-1} \psi_j^2
\end{aligned}
$$