Prediction in a general setup

Suppose Y and $W_n, \ldots, W_1$ are any random variables with finite 2^{nd} order joint moments

$E(Y) = \mu \quad ; \quad E(W_i) = \mu_{W_i}$

$Cov(Y, W_i)\ \forall i \quad Cov(W_i, W_j)\ \forall i, j$ and $V(Y)$ are all finite.

Consider linear predictor of Y based on $(W_n, \ldots, W_1)$

$$P_{(W_n, \ldots, W_1)} Y = P(Y \mid W) = a_0^* + a_1^* W_n + \cdots + a_n^* W_1$$

The above BLP is $\exists$ mean square prediction error of LP is minimum

i.e. $E(Y - P(Y|W))^2$ is min w.r.t. all possible linear predictors.

Derivation of BLP

Let $\underset{\sim}{a} = (a_0, a_1, \ldots, a_n)'$

$$S(\underset{\sim}{a}) = E\left(Y - a_0 - \sum_{i=1}^{n} a_i W_{n+1-i}\right)^2$$

$$\underset{\sim}{a}_{BLP} = \underset{\underset{\sim}{a}}{\arg\min}\ S(\underset{\sim}{a})$$

BLP eq^{n}s

$$\frac{\partial S(\underset{\sim}{a})}{\partial a_j} = 0 \; ; \quad j = 0, 1, \ldots, n$$

$\dfrac{\partial S(\underset{\sim}{a})}{\partial a_0} = 0$ gives

$$E\left(Y - a_0 - \sum_{i=1}^{n} a_i W_{n+1-i}\right) = 0$$

i.e. $a_0 = \mu_Y - \underset{\sim}{a}_n' \underset{\sim}{\mu}_{\underset{\sim}{W}}$

$$\underset{\sim}{\mu}_{\underset{\sim}{W}} = (\mu_{W_n}, \ldots, \mu_{W_1})'$$

$\dfrac{\partial S(\underset{\sim}{a})}{\partial a_j} = 0$ gives

$$E\left(Y - a_0 - \sum_{1}^{n} a_i W_{n+1-i}\right) W_{n+1-j} = 0$$

$$j = 1, \ldots, n$$

i.e. $E(Y W_{n+1-j}) - a_0 \mu_{W_{n+1-j}} - \sum_{1}^{n} a_i E(W_{n+1-i} W_{n+1-j}) = 0$

$$E(Y W_{n+1-j}) - (\mu_Y - \underset{\sim}{a}_n' \underset{\sim}{\mu}_{\underset{\sim}{W}}) \mu_{W_{n+1-j}} - \sum_{1}^{n} a_i E(W_{n+1-i} W_{n+1-j}) = 0$$

$$\left(E(Y W_{n+1-j}) - \mu_Y \mu_{W_{n+1-j}}\right) - \sum_{i=1}^{n} a_i \left(E(W_{n+1-i} W_{n+1-j}) - \mu_{W_{n+1-i}} \mu_{W_{n+1-j}}\right) = 0.$$

i.e. $\operatorname{Cov}(Y, W_{n+1-j}) - \sum_{i=1}^{n} a_i \operatorname{Cov}(W_{n+1-i}, W_{n+1-j}) = 0$

$$j = 1, \ldots, n$$

Applications

(i) Estimation of missing values.

(eg)

Let $X_t = \phi X_{t-1} + \epsilon_t$; $|\phi|<1$, $\epsilon_t \sim WN(0,\sigma^2)$

Suppose X_2 is missing and we wish to use X_1 and X_3 to estimate X_2

Frame BLP of X_2 using X_1, X_3

$$P_{(X_1,X_3)} X_2 = \alpha_1 X_1 + \alpha_2 X_3$$

$$\left(Y = X_2 \ ; \ \underset{\sim}{W} = (X_1, X_3)' \right)$$

$$\mathrm{cov}(\underset{\sim}{W}) = \begin{pmatrix} \gamma_0 & \gamma_2 \\ \gamma_2 & \gamma_0 \end{pmatrix} = \frac{\sigma^2}{1-\phi^2}\begin{pmatrix} 1 & \phi^2 \\ \phi^2 & 1 \end{pmatrix}$$

$$\underset{\sim}{\gamma} = \begin{pmatrix} \mathrm{cov}(Y, X_1) \\ \mathrm{cov}(Y, X_3) \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \gamma_1 \end{pmatrix} = \frac{\sigma^2}{1-\phi^2}\begin{pmatrix} \phi \\ \phi \end{pmatrix}$$

BLP eq^{n}s $\quad \Gamma \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \underset{\sim}{\gamma}$

$$\text{i.e} \quad \begin{pmatrix} 1 & \phi^2 \\ \phi^2 & 1 \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \phi \\ \phi \end{pmatrix}.$$

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}_{BLP} = \frac{1}{1-\phi^4}\begin{pmatrix} 1 & -\phi^2 \\ -\phi^2 & 1 \end{pmatrix}\begin{pmatrix} \phi \\ \phi \end{pmatrix}.$$

$$\text{i.e} \quad \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}_{BLP} = \frac{1}{1+\phi^2}\begin{pmatrix} \phi \\ \phi \end{pmatrix}$$

$$P_{(X_1, X_3)} X_2 = \frac{\phi}{1+\phi^2}(X_1 + X_3)$$

missing value estimate of X_2 using BLP

minimum mean square prediction error

$$E\left(X_2 - P_{(X_1, X_3)} X_2\right)^2 = V(X_2) - \underline{\alpha}' \underset{\sim}{\gamma}$$

$$= \frac{\sigma^2}{1-\phi^2} - \frac{1}{1+\phi^2}(\phi, \phi) \frac{\sigma^2}{1-\phi^2}\begin{pmatrix}\phi \\ \phi\end{pmatrix}$$

$$= \frac{\sigma^2}{1+\phi^2}$$

(ii) Back forecasting

use $(X_1, \ldots X_n)$ to back forecast $X_0, X_{-1}, \ldots$

required for conditional MLE initialization of AR / ARMA models.

$$\underbrace{P_{(X_1, \ldots X_n)}}_{\underset{\sim}{W}} X_0 \leftarrow Y$$

(iii) Defining partial auto correlation function (PACF) - a major tool in model identification

Partial Auto Correlation function (PACF)

PACF, $\alpha(k)$, at lag k, is the correlation between X_1 and X_{k+1}, adjusting for the intervening observations $(X_2, \ldots, X_k)$.

Def^n: PACF $\alpha(\cdot)$ of a stationary process $\{X_t\}$ is defined by

$$\alpha(1) = \text{Corrl}^n (X_2, X_1) = \rho(1)$$

and $\alpha(k) = \text{Corrl}^n \left(X_{k+1} - P_{(X_k, \ldots, X_2)} X_{k+1} , \right.$

$k \geq 2$ $\left. X_1 - P_{(X_k, \ldots, X_2)} X_1 \right)$

$\alpha(k)$: PACF at lag k

$P_{(\,)} X_{k+1}$ & $P_{(\,)} X_1$ can be found using the BLP approach

PACF of standard prob models

PACF of AR(1)

$\{X_t\}$ covariance stationary AR(1)

$$X_t = \phi X_{t-1} + \epsilon_t \;;\; |\phi| < 1 , \; \epsilon_t \sim WN(0, \sigma^2)$$

$$\alpha(1) = \rho(1) = \phi$$

$$\alpha(2) = \text{Corr}\left(X_3 - P_{X_2} X_3 , \; X_1 - P_{X_2} X_1 \right)$$

$$f(\beta) = E\left(X_3 - \beta X_2\right)^2$$

BLP eq^n : $E\left(X_3 - \beta X_2\right) X_2 = 0$

$$\beta_{BLP} = \rho_1 = \phi$$

$$\tilde{f}(\alpha) = E\left(X_1 - \alpha X_2\right)^2$$

BLP eq^n : $E\left(X_1 - \alpha X_2\right) X_2 = 0$

$$\alpha_{BLP} = \rho_1 = \phi$$

$\Rightarrow$ $P_{X_2} X_3 = \phi X_2$ & $P_{X_2} X_1 = \phi X_2$

$$\alpha_{\cancel{2}}(2) = \mathrm{Corr}\left(X_3 - \phi X_2, \; X_1 - \phi X_2\right)$$

$$= \mathrm{Corr}\left(\epsilon_3, \; X_1 - \phi X_2\right)$$

$$= \frac{\mathrm{Cov}\left(\epsilon_3, X_1 - \phi X_2\right)}{\left[V(\epsilon_3) \; V(X_1 - \phi X_2)\right]^{1/2}}$$

$= 0$ as $\mathrm{Cov}\left(\epsilon_t, X_{t-j}\right) = 0 \quad \forall j > 0$

$$\alpha(3) = \mathrm{Corr}\left(X_4 - P_{(X_2, X_3)} X_4, \; X_1 - P_{(X_2, X_3)} X_1\right).$$

$$= \mathrm{Corr}\left(X_4 - \phi X_3, \; X_1 - \alpha_{1(BLP)} X_2 - \alpha_{2(BLP)} X_3\right)$$

$$= \mathrm{Corr}\left(\epsilon_4, \; X_1 - \alpha_{1(BLP)} X_2 - \alpha_{2(BLP)} X_3\right)$$

$$= 0$$

$\forall K \geq 2 \quad \alpha(K) = 0.$

Remark: PACF of AR(1) cuts off after lag 1.