Classical Linear Regression Model

- Notation and Assumptions
- Model Estimation
  - Method of Moments
  - Least Squares
  - Partitioned Regression
- Model Interpretation

Notations

- $\mathbf{y}$: Dependent Variable (Regressand)
  - y_i, i = 1,2,…n
- $\mathbf{X}$: Explanatory Variables (Regressors)
  - $\mathbf{x}_i'$, i = 1,2,…,n
  - x_{ij}, i = 1,2,…,n; j = 1,2,…,K

Assumptions

- Assumption 1: Linearity

  $y_i = \ {}_1x_{i1} + \ {}_2x_{i2} + \ldots + \ {}_Kx_{iK} + \ {}_i$

  (i = 1,2,…,n)

- Linearity in Vector Notation

  $y_i = \mathbf{x}_i' \ + \ {}_i$

  (i = 1,2,…,n)

$$\mathbf{x}_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iK} \end{bmatrix}, \qquad \begin{bmatrix} {}_1 \\ {}_2 \\ \vdots \\ {}_K \end{bmatrix}$$

Assumptions

- Linearity in Matrix Notation

  $\mathbf{y} = \mathbf{X} \quad +$

$y x x_{111111}$ $\mathbf{x}'$ L K

å X MM,,MMKM

$y x x_{nnnnnK}$ $\mathbf{x}'$ 1 L

Assumptions

- Assumption 2: Exogeneity
  $E(\varepsilon_i|\mathbf{X}) = E(\varepsilon_i|x_{j1},x_{j2},\ldots,x_{jK}) = 0$
  $(i,j = 1,2,\ldots,n)$
- Implications of Exogeneity
  - $E(\varepsilon_i) = E(E(\varepsilon_i|\mathbf{X})) = 0$
  - $E(x_{jk}\varepsilon_i) = E(E(x_{jk}\varepsilon_i|\mathbf{X})) = E(x_{jk}E(\varepsilon_i|\mathbf{X})) = 0$
  - $Cov(\varepsilon_i,x_{jk}) = E(x_{jk}\varepsilon_i) - E(x_{jk})E(\varepsilon_i) = 0$

Assumptions

- Assumption 3: No Multicolinearity
  rank($\mathbf{X}$) = K, with probability 1.
- Assumption 4: Spherical Error Variance
  $E(\ _i^2|\mathbf{X}) = \ ^2 > 0$, i = 1,2,…,n
  $E(\ _i\ _j|\mathbf{X}) = 0$, i,j = 1,2,…,n; i j
  In matrix notation:
  $Var(\mathbf{e}|\mathbf{X}) = E(\ \ '|\mathbf{X}) = \ ^2\mathbf{I}_n$

Assumptions

- Implication of Spherical Error Variance (Homoscedasticity and No Autocorrelation)
  - $\text{Var}(\ ) = \text{E}(\text{Var}(\ |\mathbf{X}))+\text{Var}(\text{E}(\ |\mathbf{X})) =$ $\text{E}(\text{Var}(\ |\mathbf{X})) = \ ^2\mathbf{I}_n$
  - $\text{Var}(\mathbf{X}'\ |\mathbf{X}) = \text{E}(\mathbf{X}'\ \ '\mathbf{X}|\mathbf{X}) = \ ^2\text{E}(\mathbf{X}'\mathbf{X})$
    Note: $\text{E}(\mathbf{X}'\ |\mathbf{X}) = 0$ followed from Exogeneity Assumption

Discussions

- Exogeneity in Time Series Models
- Random Samples
  The sample ($\mathbf{y}$,$\mathbf{X}$) is a random sample if $\{y_i,\mathbf{x}_i\}$ is i.i.d (independently and identically distributed).
- Fixed Regressors
  $\mathbf{X}$ is fixed or deterministic.

Discussions

- Nonlinearity in Variables
  $g(y_i) = \quad {}_1f_1(x_{i1})+ \quad {}_2f_2(x_{i2})+\ldots+ \quad {}_Kf_K(x_{iK})+ \quad {}_i$
  (i = 1,2,…,n)
  - Linearity in Parameters and Model Errors

$y_{iii}\,eeey^{\hat{a}x\hat{a}}_{iii}$ ln() $\hat{a}$

$yxeyx_{iikiikki}$ $\sum_{k\ 1}^{K}$ $_{ki}$ ln()ln() $_0$ $\sum_{k\ 1}^{K}$

Method-of-Moments Estimator

- From the implication of strict exogeneity assumption, $E(x_{jk}\ \ _{i}) = 0$ for i = 1,2,…,n and k = 1,2,…,K. That is,
- Moment Conditions: $E(\mathbf{X}'\ \ ) = \mathbf{0}$
- $\mathbf{X}'(\mathbf{y}-\mathbf{X}\ \ ) = \mathbf{0}$ or $\mathbf{X}'\mathbf{y}=\mathbf{X}'\mathbf{X}$
- The method-of-moments estimator of :
  $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$

Least Squares Estimator

- OLS: $\mathbf{b}$ = argmin SSR( )
- SSR( ) =   '
  = ($\mathbf{y}$-$\mathbf{X}$)'($\mathbf{y}$-$\mathbf{X}$)
  = $\quad_{i=1,2\ldots,n}(y_i - \mathbf{x}_i'\mathbf{b})^2$
- $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$

The Algebra of Least Squares

$$\frac{\mathrm{SSR}(\hat{\mathbf{a}})}{\hat{\mathbf{a}}} \quad 2\mathbf{X}'\mathbf{y} \quad 2\mathbf{X}'\mathbf{X}\hat{\mathbf{a}}$$

$$\frac{{}^{2}\mathrm{SSR}(\hat{\mathbf{a}})}{\hat{\mathbf{a}}\ \ \hat{\mathbf{a}}'} \quad 2\mathbf{X}'\mathbf{X}$$

- Normal Equations: $\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{y}$
- $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$

The Algebra of Least Squares

- $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \quad + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$
- $\mathbf{b} = (\mathbf{X}'\mathbf{X}/n)^{-1}(\mathbf{X}'\mathbf{y}/n) = \mathbf{S}_{xx}^{-1}\mathbf{s}_{xy}$
  $\mathbf{S}_{xx}$ is the sample average of $\mathbf{x}_i\mathbf{x}_i'$, and $\mathbf{s}_{xy}$ is the sample average of $\mathbf{x}_i y_i$.
- OLS Residuals: $\mathbf{e} = \mathbf{y} – \mathbf{X}\mathbf{b}$
  $\mathbf{X}'\mathbf{e} = \mathbf{0}$
- $s^2 = SSR/(n-K) = \mathbf{e}'\mathbf{e}/(n-K)$

The Algebra of Least Squares

- Projection Matrix **P** and "Residual Maker" **M**:
  - **P** = **X**(**X'X**)$^{-1}$**X'**, **M** = $\mathbf{I}_n$ **- P**
  - **PX** = **X**, **MX** = **0**
  - **Py** = **Xb**, **My** = **y-Py** = **e**, **M** = **e**
- SSR = **e'e** = '**M**
- **h** = Vector of diagonal elements in **P** can be used to check for the **influential observation**: $h_i > K/n$, with 0 h_i 1.

The Algebra of Least Squares

- Fitted Value: $\hat{\mathbf{y}} \quad \mathbf{Xb} \quad \mathbf{Py}, \quad \mathbf{e} \quad \mathbf{y} \quad \hat{\mathbf{y}}$
- Uncentered R^2: $R_{uc}^2 = 1 - \mathbf{e'e}/\mathbf{y'y} = \hat{\mathbf{y}}'\hat{\mathbf{y}} / \mathbf{y'y}$
- Centered R^2: $R^2 = 1 \quad \mathbf{e'e}/(\mathbf{y} \quad \bar{\mathbf{y}})'(\mathbf{y} \quad \bar{\mathbf{y}})$

$$(\mathbf{y} \quad \bar{\mathbf{y}})'(\mathbf{y} \quad \bar{\mathbf{y}}) \quad (\hat{\mathbf{y}} \quad \bar{\mathbf{y}})'(\hat{\mathbf{y}} \quad \bar{\mathbf{y}}) \quad \mathbf{e'e}$$

with $\bar{\mathbf{y}} \quad \frac{1}{n} \, {}_{i \ 1}^{\ n} \, y_i$

Analysis of Variance

- TSS = $(\mathbf{y} \quad \bar{\mathbf{y}})'(\mathbf{y} \quad \bar{\mathbf{y}})$
- ESS = $(\hat{\mathbf{y}} \quad \bar{\mathbf{y}})'(\hat{\mathbf{y}} \quad \bar{\mathbf{y}})$, where $\bar{\mathbf{y}} \quad \bar{\hat{\mathbf{y}}}$
- RSS = $\mathbf{e}'\mathbf{e}$ (or SSR)
- TSS = ESS + RSS
- Degrees of Freedom
  - TSS: n-1
  - ESS: K-1
  - RSS: n-K

Analysis of Variance

- R^2

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\mathbf{e}'\mathbf{e}}{(\mathbf{y}-\bar{\mathbf{y}})'(\mathbf{y}-\bar{\mathbf{y}})}$$

- Adjusted R^2

$$\bar{R}^2 = 1 - \frac{\frac{RSS}{n-K}}{\frac{TSS}{n-1}} = 1 - \frac{\frac{\mathbf{e}'\mathbf{e}}{n-K}}{\frac{(\mathbf{y}-\bar{\mathbf{y}})'(\mathbf{y}-\bar{\mathbf{y}})}{n-1}}$$

Discussions

- Examples
  - Constant or Mean Regression

    $y_i = \quad + \quad {}_i$ (i=1,2,…,n)

    $$a = \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

  - Simple Regression

    $y_i = \quad + \quad x_i + \quad {}_i$ (i=1,2,…,n)

$$b = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \quad a = \bar{y} - b\bar{x}$$

Partitioned Regression

- $\mathbf{y} = \mathbf{X} \quad + \qquad \mathbf{X}_1 \;\; 1 + \mathbf{X}_2 \;\; 2 +$
  $\mathbf{X} = \; \mathbf{X}_1 \; \mathbf{X}_2],$ $\qquad 1' \quad 2' \; '$
- Normal Equations
  $\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{y}$, or

  $\mathbf{X}_1'\mathbf{X}_1 \; \mathbf{X}_1'\mathbf{X}_2 \; \mathbf{b}_1 \; \mathbf{X}_1'\mathbf{y}$

  $\mathbf{X}_2'\mathbf{X}_1 \; \mathbf{X}_2'\mathbf{X}_2 \; \mathbf{b}_2 \; \mathbf{X}_2'\mathbf{y}$

Partitioned Regression

- $\mathbf{b}1 = (\mathbf{X}1'\mathbf{X}1)^{-1}\mathbf{X}1'\mathbf{y} - (\mathbf{X}1'\mathbf{X}1)^{-1}(\mathbf{X}1'\mathbf{X}2)\mathbf{b}2$
  $= (\mathbf{X}1'\mathbf{X}1)^{-1}\mathbf{X}1'(\mathbf{y}-\mathbf{X}2\mathbf{b}2)$
- $\mathbf{b}2 = (\mathbf{X}2'\mathbf{M}1\mathbf{X}2)^{-1}\mathbf{X}2'\mathbf{M}1\mathbf{y} = (\mathbf{X}2^{*\prime}\mathbf{X}2^{*})^{-1}\mathbf{X}2^{*\prime}\mathbf{y}^{*}$
  $\mathbf{M}1 = [\mathbf{I}-\mathbf{X}1(\mathbf{X}1'\mathbf{X}1)^{-1}\mathbf{X}1']$
  $\mathbf{X}2^{*} = \mathbf{M}1\mathbf{X}2$ (matrix of residuals obtained from each column of **X**2 regressed on **X**1)
  $\mathbf{y}^{*} = \mathbf{M}1\mathbf{y}$ (residuals of **y** regressed on **X**1)
- **b**2 is LS estimator of $\mathbf{y}^{*}$ on $\mathbf{X}2^{*}$

Partitioned Regression

- **Frisch-Waugh Theorem**
  If $\mathbf{y} = \mathbf{X}1 \;\; 1 + \mathbf{X}2 \;\; 2 +$, then $\mathbf{b}2$ is the LS estimator for the regression of residuals *from a regression of **y** on **X1** alone* are regressed on the set of residuals *obtained when each column of **X**2 is regressed on **X1***.
- If $\mathbf{X}1$ and $\mathbf{X}2$ are independent, $\mathbf{X}1'\mathbf{X}2 = 0$.
  $\mathbf{b}1 = (\mathbf{X}1'\mathbf{X}1)^{-1}\mathbf{X}1'\mathbf{y}$
  $\mathbf{b}2 = (\mathbf{X}2'\mathbf{X}2)^{-1}\,\mathbf{X}2'\mathbf{y}$

Partitioned Regression

- **Applications of Frisch-Waugh Theorem**
  - Partial Regression Coefficient
  - Trend Regression

    $y_t = \quad + \quad t + \quad {}_t$

Model Interpretation

- Marginal Effects (*Ceteris Paribus* Interpretation): **y = Xb + e**, where $b_k \quad \frac{\mathbf{y}}{\mathbf{x}_k}$
- Elasticity Interpretation

$$yeeeyb_{iiiik} \quad \ln() \quad \hat{a} \quad \frac{\ln()\mathbf{y}\mathbf{y}1}{\mathbf{x}\mathbf{y}\mathbf{x}_{ik}}$$

$$yxeyx_{iikiikki} \quad {}^{K}_{k\ 1} \quad {}_{ki} \quad \ln()\ln() \quad {}_{0} \quad {}^{K}_{k\ 1}$$

$$b_k \quad \frac{\ln()\mathbf{y}\mathbf{y}}{\ln()\mathbf{x}\mathbf{y}\mathbf{x}_{ik}} \quad \frac{\mathbf{x}_k}{}$$

Example

- U. S. Gasoline Market, 1953-2004
  - EXPG = Total U.S. gasoline expenditure
  - PG = Price index for gasoline
  - Y = Per capita disposable income
  - Pnc = Price index for new cars
  - Puc = Price index for used cars
  - Ppt = Price index for public transportation
  - Pd = Aggregate price index for consumer durables
  - Pn = Aggregate price index for consumer nondurables
  - Ps = Aggregate price index for consumer services
  - Pop = U.S. total population in thousands

Example

- **y** = **X**   +
- **y** = G; **X** = [1 PG Y]
  where G = (EXPG/PG)/POP
- **y** = *ln*(G); **X** = [1 *ln*(PG) *ln*(Y)]
- Elasticity Interpretation

Example (Continued)

- **y** = **X**1 + **X**2 +
- **y** = *ln*(G); **X1** = [1 YEAR];
  X2 = [*ln*(PG) *ln*(Y)]
  where G = (EXPG/PG)/POP
- **Frisch-Waugh Theorem**