Ajustamento e Análise de Variância.

Seja i um vetor ($n\times1$) de 1's, então,

$$\sum x_i = x_1 + x_2 + \ldots + x_n = i'x = n\bar{x}$$

$$i'x = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\bar{x} = \frac{1}{n}\sum x_i = \frac{1}{n} i'x$$

$$i\bar{x} = i\frac{1}{n}i'x = \frac{1}{n}ii'x = \begin{bmatrix} \bar{x} \\ \bar{x} \\ \vdots \\ \bar{x} \end{bmatrix}$$

$$\begin{bmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{bmatrix} = \left[x - i\bar{x}\right] = \left[x - \frac{1}{n} ii'x\right]$$

ou

$$\left[x - \frac{1}{n} ii'x\right] = \left[Ix - \frac{1}{n} ii'x\right] = \left[I - \frac{1}{n} ii'\right]x = M^0 x$$

que é o vetor de desvios com relação a média.

Considere agora,

$$y = Xb + e$$

Colocando todas as variáveis na forma de desvio tem-se,

$$M^0 y = M^0 Xb + M^0 e$$

uma vez que os resíduos tem média zero,

$$M^0 e = e$$

e então,

$$M^0 y = M^0 Xb + e$$

Premultiplicando por $y' = b'X' + e'$

$$y'M^0 y = (b'X' + e')(M^0 Xb + e)$$

$$y'M^0 y = b'X'M^0 Xb + b'X'e + e'M^0 Xb + e'e$$

mas, $\quad e'M^0 X = e'X = 0$

$$\begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix} \begin{bmatrix} X_1 - \overline{X} \\ \vdots \\ X_n - \overline{X} \end{bmatrix} = e_1(X_1 - \overline{X}) + \ldots + e_n(X_n - \overline{X}) = e_1 X_1 + \ldots + e_n X_n - \overline{X}(e_1 + \ldots + e_n) =$$

$$= e_1 X_1 + \ldots + e_n X_n - \frac{\sum X}{n} \cancelto{0}{\sum e} = e'X = 0$$

Portanto,

$$y'M^0y = b'X'M^0Xb + e'e$$

Dividindo - se por $y'M^0y$ tem - se,

$$1 = \frac{b'X'M^0Xb}{y'M^0y} + \frac{e'e}{y'M^0y}$$

ou

$$1 - \frac{SQres}{SQTotal} = \frac{SQregr}{SQTotal} = R^2$$

O coeficiente de determinação R^2 é uma medida de ajuste do modelo.

Mede a proporção da variação de y que é explicada por X.

A prova de que $0 \leq R^2 \leq 1$ requer X conter uma coluna de 1's, isto é, ter constante. Senão, $M^o e \neq e$ e $e'M^oX \neq 0$.

R^2 depende do uso de MQO e se o modelo tem constante.

$R^2 = 0$ se todos os valores previstos de y forem $\bar{y}$, isto é, a regressão é uma linha horizontal. Neste caso, $SQres = SQtotal$

$R^2 = 1$ se a regressão passar exatamente pelos pontos observados. Neste caso, $SQres = 0$.

Cuidado:

Contextos diferentes, pouco pode ser dito sobre o R^2 .

séries temporais x cross-section

A variável dependente deve ser a mesma para se comparar o R^2 em dois modelos diferentes.

Ex: mesmo modelo, mas R^2 diferentes.

Renda = poupança + consumo

$\Rightarrow$ Consumo = $a + b$ renda

Renda – Poupança = $a + b$ renda

$\Rightarrow$ Poupança = $-\,a + (1-b)$ renda

O R^2 nunca diminuirá quando outra variável for adicionada à equação. Devido a isso, geralmente é apresentado o R^2 ajustado para graus de liberdade.

$$\bar{R}^2 = 1 - \frac{e'e/(n-K)}{y'M^0y/(n-1)} = 1 - \frac{(n-1)(1-R^2)}{(n-K)}$$

Voltando a igualdade,

$$y'M^0y = b'X'M^0Xb + e'e$$

ou

$$y'M^0y = b'X'M^0(y-e) + e'e$$

$$y'M^0y = b'X'M^0y - b'X'M^0e + e'e$$

Mas, $M^0e = e$ e $b'X'e = 0$, então

$$y'M^0y = b'X'M^0y + e'e$$

$$y'M^0y = [y_1 y_2 \ldots y_n]\begin{bmatrix} y_1 - \bar{y} \\ y_2 - \bar{y} \\ \cdots \\ y_n - \bar{y} \end{bmatrix} = y_1^2 - y_1\bar{y} + y_2^2 - y_2\bar{y} + \ldots + y_n^2 - y_n\bar{y}$$

$$= y_1^2 + y_2^2 + \ldots + y_n^2 - \bar{y}(y_1 + \ldots + y_n)$$

$$= \sum y_i^2 - \bar{y}\sum y_i \cdot \frac{n}{n}$$

$$= \sum y_i^2 - n\bar{y}^2$$

Analogamente,

$$b'X'M^0y = b'X'y - n\bar{y}^2$$

Causa de variação	g. l.	SQ	QM	F
Regressão	$K-1$	$b'X'y-n\bar{y}^2$	$b'X'y-n\bar{y}^2 / K-1$	$QMregr/QMres$
Resíduo	$n-K$	$e'e$	$e'e/n-K$	
Total	$n-1$	$y'y-n\bar{y}^2$		

$SQres = \sum\left(y_i - \hat{y}_i\right)^2$ Mede a variação de y que não é explicada pela regressão

$SQregr = \sum\left(\hat{y}_i - \bar{y}\right)^2$ Mede a variação de y que é explicada pela regressão

$SQtotal = \sum\left(y_i - \bar{y}\right)^2$ Mede a variação de y em relação a média

Teste de significância da regressão.

Queremos testar a hipótese de que todos os coeficientes, exceto o termo constante, são zero.

$$F(K-1, n-K) = \frac{R^2/(K-1)}{(1-R^2)/(n-K)} = \frac{\dfrac{SQ\,\mathrm{Re}\,gr}{SQtotal}\Big/K-1}{\dfrac{SQres}{SQtotal}\Big/n-K} = \frac{\dfrac{SQ\,\mathrm{Re}\,gr}{K-1}}{\dfrac{SQres}{n-K}} = \frac{QMreg}{QMres}$$

EXEMPLO

$$y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

Y	X_1	X_2
-4	0	3
5	1	1
4	2	2
11	3	0

$$SQtotal = \sum y^2 - \frac{\left(\sum y\right)^2}{n} = 178 - \frac{16^2}{4} = 114$$

$$SQtotal = y'y - n\bar{y}^2 = 178 - 4.4^2 = 114$$

$$SQregr = b'X'y - n\bar{y}^2 = \begin{bmatrix} 5{,}5 & 2 & -3 \end{bmatrix} \begin{bmatrix} 16 \\ 46 \\ 1 \end{bmatrix} - 64 = 113$$

$$SQres = SQtotal - SQregr = 114 - 113 = 1$$

	g.l.	SQ	QM	F
regressão	2	113	56,5	56,5
resíduo	1	1	1	
total	3	114		

$$R^2 = \frac{SQregr}{SQtotal} = \frac{113}{114} = 0,99$$

Previsão:

Suponha que desejamos prever o valor y^o associado com o vetor de regressão X^o

$$y^0 = X^0\beta + \varepsilon^0$$

Segue do teorema de Gauss-Markov que

$$\hat{y}^0 = X^0 b$$

é o estimador linear não tendencioso de variância mínima de $E\left(y^0\right)$.

O erro de previsão é:

$$e^0 = \hat{y}^0 - y^0 = X^0 b - X^0\beta - \varepsilon^0 = X^0\left(b-\beta\right) - \varepsilon^0 \Rightarrow E(e^0) = 0$$

$$\text{var}\left(e^o\right) = E\left[\left(\hat{y}^o - y^o\right)\left(\hat{y}^o - y^o\right)'\right] = E\left[\left(X^o(b-\beta) - \varepsilon^o\right)\left((b-\beta)' X^{o'} - \varepsilon^{o'}\right)\right]$$

$$= X^o E\left[(b-\beta)(b-\beta)'\right]X^{o'} + E\left(\varepsilon^o \varepsilon^{o'}\right) = X^o \sigma^2 (X'X)^{-1} X^{o'} + \sigma^2$$

$$X^o E\left[(b-\beta)\varepsilon^{o\prime}\right] = E\left[\varepsilon^o (b-\beta)'\right] X^{o\prime} = 0$$ porque não existe correlação entre b e ε^0

Intervalo de previsão:

$$\hat{y}^o - t\sqrt{s^2\left(1 + X^o (X'X)^{-1} X^{o\prime}\right)} \le y^o \le \hat{y}^o + t\sqrt{s^2\left(1 + X^o (X'X)^{-1} X^{o\prime}\right)}$$

A variância da previsão pode ser estimada utilizando-se s^2 ao invés de σ^2.

O intervalo de confiança para y^0 é

$$\hat{y}^o - t\sqrt{s^2\left(X^o (X'X)^{-1} X^{o\prime}\right)} \le y^o \le \hat{y}^o + t\sqrt{s^2\left(X^o (X'X)^{-1} X^{o\prime}\right)}$$

Comparação entre intervalo de previsão e intervalo de confiança

1) O intervalo de previsão prevê observação

$$y^0 = \beta' X^0 + \varepsilon^0$$

Ex: estamos interessados em prever os gastos com alimentos <u>de uma família</u> que recebe 2 s.m./mês.

2) O intervalo de confiança prevê média

$$E\left(y^0\right) = \beta' X^0$$

Ex: estamos interessados em prever os gastos com alimentos de <u>todas as famílias</u> que recebem 2 s.m./mês.

O mesmo previsor é usado em ambos os casos

$$\hat{y}^0 = E\left(\hat{y}^0\right) = b' X^0$$

Entretanto a variabilidade amostral é diferente. Para prever uma observação individual nós consideramos a variância do erro de previsão

$$\operatorname{var}\left(\hat{y}^o - y^o\right) = \sigma^2\left(1 + X^{o'}\left(X'X\right)^{-1} X^o\right)$$

Enquanto a variância da média do previsor é

$$\operatorname{var}\left(\hat{y}^o\right) = \sigma^2\left(X^{o'}\left(X'X\right)^{-1} X^o\right)$$

Portanto,

$$\operatorname{var}\left(\hat{y}^o - y^o\right) = \operatorname{var}(\hat{y}^0) + \sigma^2$$

Ao se prever uma observação individual, adiciona-se a incerteza de se prever o erro futuro ε^0.

Medindo a Precisão das Previsões

Estatística U de Theil: $$U = \sqrt{\frac{\left(\frac{1}{n^0}\right)\sum_i \left(y_i - \hat{y}_i^0\right)^2}{\left(\frac{1}{n^0}\right)\sum_i y_i^2}}$$

onde n^0 é o número de períodos sendo previsto

i é o número de previsões

Valores altos indicam má previsão

EXEMPLO

$$y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

Y	X_1	X_2
-4	0	3
5	1	1
4	2	2
11	3	0

Obter o valor estimado de y^0 para $X_1 = X_2 = 2{,}5$, e obter o intervalo de previsão a 90%.

$$\hat{y}^o - t\sqrt{s^2\left(1 + X^o\left(X'X\right)^{-1} X^{o'}\right)} \leq y^o \leq \hat{y}^o + t\sqrt{s^2\left(1 + X^o\left(X'X\right)^{-1} X^{o'}\right)}$$

$$\hat{y}^0 = 5{,}5 + 2.(2{,}5) - 3.(2{,}5) = 3$$

$$X^0 = \begin{bmatrix} 1 & 2{,}5 & 2{,}5 \end{bmatrix} \qquad X^o\left(X'X\right)^{-1} X^{o'} = 2{,}25$$

$$3 - 6{,}31\sqrt{1.(1 + 2{,}25)} \leq y^0 \leq 3 + 6{,}31\sqrt{1.(1 + 2{,}25)}$$

$$-8{,}38 \leq y^0 \leq 14{,}38$$

Restrições da demanda

1) Homogênea de grau zero

$$e_{i1} + e_{i2} + \ldots + e_{in} = -e_{iY}$$

Elasticidade preço direta e cruzada = - elasticidade renda

2) Agregação de Engel

$$e_{1y}W_1 + e_{2y}W_2 + \ldots + e_{ny}W_n = 1$$

Proporção da despesa total gasta com o bem 1

3) Simetria

$$e_{ij} = e_{ji}$$

Elasticidades preço cruzada são iguais

Restrição na função de produção:

$$\ln Y = \gamma + \alpha \ln L + \beta \ln K + \varepsilon$$

impondo retornos constantes à escala

$$\alpha + \beta = 1$$

Regressão com Restrições

Exemplo:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

$$\beta_1 + \beta_2 = 1$$

A restrição é imposta ao modelo.

$$Y = \alpha + (1 - \beta_2) X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

$$Y - X_1 = \alpha + \beta_2 (X_2 - X_1) + \beta_3 X_3 + \varepsilon$$

$$Z = \alpha + \beta_2 W + \beta_3 X_3 + \varepsilon$$

Considere o modelo

$$y = X\beta + \varepsilon$$

onde $\quad R_{(J \times K)} \beta = q$

O estimador de mínimos quadrados restrito é obtido da solução:

Minimizar

$$S(\beta) = (y - X\beta)'(y - X\beta) \quad \text{sujeito a} \quad R\beta = q$$

A solução Lagrangeana para este problema pode ser escrita como:

$$b_* = \underset{\tilde{\beta},\lambda}{Min}\ S^*(\beta) = (y - X\beta)'(y - X\beta) + 2\lambda'(R\beta - q)$$

$$= y'y - y'X\beta - \beta'X'y + \beta'X'X\beta + 2\lambda'(R\beta - q)$$

As condições necessárias para a minimização são:

$$\frac{\partial S^*}{\partial \beta} = -2X'(y - Xb_*) + 2R'\lambda = 0$$

$$\frac{\partial S^*}{\partial \lambda} = 2(Rb_* - q) = 0$$

ou

$$\begin{aligned} -X'y + X'Xb_* + R'\lambda &= 0 \\ Rb_* &= q \end{aligned}$$

ou

$$X'Xb_* + R'\lambda = X'y$$
$$Rb_* \qquad\qquad = q$$

ou em forma de matriz,

$$\begin{bmatrix} X'X & R' \\ R & 0 \end{bmatrix}\begin{bmatrix} b_* \\ \lambda \end{bmatrix} = \begin{bmatrix} X'y \\ q \end{bmatrix}$$

Matriz Inversa Particionada

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1}\left(I + BFCA^{-1}\right) & -A^{-1}BF \\ -FCA^{-1} & F \end{bmatrix}$$

$$F = \left(D - CA^{-1}B\right)^{-1}$$

$$\begin{bmatrix} X'X & R' \\ R & 0 \end{bmatrix}^{-1} = \begin{bmatrix} (X'X)^{-1}\left(I + R'\left(0 - R\left(X'X\right)^{-1}R'\right)^{-1}R\left(X'X\right)^{-1}\right) & -(X'X)^{-1}R'(0 - R(X'X)^{-1}R')^{-1} \\ & \end{bmatrix}$$

Para obter b_* multiplica-se a 1ª linha da matriz inversa pelo vetor $\begin{bmatrix} X'y \\ q \end{bmatrix}$

$$(X'X)^{-1}\left(I + R'\left(0 - R(X'X)^{-1}R'\right)^{-1}R(X'X)^{-1}\right) \quad -(X'X)^{-1}R'(0 - R(X'X)^{-1}R')^{-1}\begin{bmatrix} X'y \\ q \end{bmatrix}$$

$$b_* = (X'X)^{-1}X'y - (X'X)^{-1}R'\left(R(X'X)^{-1}R'\right)^{-1}R(X'X)^{-1}X'y + (X'X)^{-1}R'\left(R(X'X)^{-1}R'\right)^{-1}q$$

$$\boxed{b_* = b - (X'X)^{-1}R'\left(R(X'X)^{-1}R'\right)^{-1}(Rb - q)}$$

No exemplo:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

$$\beta_1 + \beta_2 = 1$$

para obter b_* utiliza-se a fórmula acima e $R = \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix}$ e $q = 1$

$$b_* = b - (X'X)^{-1} R' \left(R (X'X)^{-1} R'\right)^{-1} (Rb - q)$$

$$e_* = y - Xb_* + Xb - Xb$$

$$e_* = y - Xb - X(b_* - b) = e - X(b_* - b)$$

$$e_*' e_* = \left(e - X(b_* - b)\right)' \left(e - X(b_* - b)\right)$$

$$e_*' e_* = e'e - e'X(b_* - b) - (b_* - b)' X'e + (b_* - b)' X'X(b_* - b)$$

$$e_*' e_* - e'e = (Rb - q)' \left(R (X'X)^{-1} R'\right)^{-1} R (X'X)^{-1} X'X (X'X)^{-1} R' \left(R (X'X)^{-1} R'\right)^{-1} (Rb - q)$$

$$e_*' e_* - e'e = (Rb - q)' \left(R (X'X)^{-1} R'\right)^{-1} (Rb - q)$$

$$F(J / n - K) = \frac{e_*' e_* - e'e / J}{e'e / (n - K)}$$

Nós vimos que:

$$F = (Rb - q)' \left[s^2 R (X'X)^{-1} R'\right]^{-1} (Rb - q) / J = \frac{(Rb - q)' \left[R (X'X)^{-1} R'\right]^{-1} (Rb - q) / J}{e'e / n - K}$$

O R^2 aumenta ao se aumentar o número de variáveis, portanto o R^2 do modelo restrito é menor que o R^2 do modelo não restrito.

Vimos que,

$$e_*'e_* = e'e - (b_* - b)' X'e - e'X(b_* - b) + (b_* - b)' X'X(b_* - b)$$

$$e_*'e_* = e'e + (b_* - b)' X'X(b_* - b) \geq e'e$$

Como $$R^2 = 1 - \frac{e'e}{SQtotal}$$

$$R_*^2 < R^2$$

Modelo restrito (menor número de variáveis)

Na tabela de análise de variância o teste F é obtido da seguinte forma:

Causa de variação	g. l.	SQ	QM	F
Regressão	$K-1$	$b'X'y - n\bar{y}^2$	$b'X'y - n\bar{y}^2 / K-1$	$QMregr/QMres$
Resíduo	$n-K$	$e'e$	$e'e/n-K$	
Total	$n-1$	$y'y - n\bar{y}^2$		

A hipótese testada é a de que todos os coeficientes do modelo de regressão, exceto a constante, são zero:

$$H_0 : \beta_2 = \beta_3 = \cdots = \beta_K = 0$$

Neste caso temos:

$$F = \frac{SQregr / g.l.regr}{SQres / g.l.res} = \frac{SQtotal - SQres / J}{SQres / n-K}$$

$$= \frac{SQres.restrito - SQres / K-1}{SQres / n-K} = \frac{(e_*'e_* - e'e) / J}{e'e / n-K}$$

Se a hipótese sendo testada for imposta ao modelo teríamos:

$$y = \beta_1 + \varepsilon$$

Neste caso é fácil mostrar que,

$$\hat{\beta}_1 = \frac{\sum y_i}{n} = \bar{y}$$

e, então, a *SQtotal* coincide com a *SQres.restrito*.

Derivação de Matrizes

$$y = \beta_1 + \beta_2 x + \varepsilon$$

Considere 2 parâmetros e 4 observações :

$$b'x'y = \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} \Sigma y \\ \Sigma xy \end{bmatrix}$$

$$A = b_1 \Sigma y + b_2 \Sigma xy$$

$$\frac{\partial A}{\partial b_1} = \Sigma y$$

$$\begin{bmatrix} \Sigma y \\ \Sigma xy \end{bmatrix} = x'y$$

$$\frac{\partial A}{\partial b_2} = \Sigma xy$$

$$R' = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} \qquad \text{J=2 K=3}$$

$$A = \lambda' R b \longrightarrow \frac{\partial A}{\partial b} = R'\lambda$$

$$A = \begin{bmatrix} \lambda_1 & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 - \lambda_2 & \lambda_2 & \lambda_1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \lambda_1 b_1 - \lambda_2 b_1 + \lambda_2 b_2 + \lambda_1 b_3$$

$$A = \lambda_1 b_1 - \lambda_2 b_1 + \lambda_2 b_2 + \lambda_1 b_3$$

$$\frac{\partial A}{\partial b_1} = \lambda_1 - \lambda_2$$

$$\frac{\partial A}{\partial b_2} = \lambda_2$$

$$\frac{\partial A}{\partial b_3} = \lambda_1$$

$$\frac{\partial A}{\partial b} = \begin{bmatrix} \lambda_1 - \lambda_2 \\ \lambda_2 \\ \lambda_1 \end{bmatrix} = R'\lambda$$

****