Lista 4

MI406-Regressão

Caio Gomes Alves

1 Questão 1

1.1 Pergunta

Uma matriz A, de dimensões $n \times n$, é dita **diagonal** se A pode ser escrita como

$$A = \begin{bmatrix} a_1 & 0 & 0 & \dots & 0 \\ 0 & a_2 & 0 & \dots & 0 \\ 0 & 0 & a_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_n \end{bmatrix}$$

Mostre que se A é uma matriz diagonal, então A^{-1} também é diagonal.

1.2 Resposta

Sabemos que para que uma matriz quadrada M qualquer tenha inversa, ela deve poder resolver o seguinte sistema:

$$M = \begin{bmatrix} m_{11} & m_{12} & \dots & m_{1n} \\ m_{21} & m_{22} & \dots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \dots & m_{nn} \end{bmatrix} \quad M^{-1} = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1n} \\ z_{21} & z_{22} & \dots & z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \dots & z_{nn} \end{bmatrix}$$

$$M \times M^{-1} = I \Rightarrow \begin{bmatrix} m_{11} & m_{12} & \dots & m_{1n} \\ m_{21} & m_{22} & \dots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \dots & m_{nn} \end{bmatrix} \times \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1n} \\ z_{21} & z_{22} & \dots & z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \dots & z_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Nesse caso, a multiplicação matricial resulta no seguinte sistema de equações:

$$\begin{cases} m_{11}z_{11} + m_{12}z_{21} + \ldots + m_{1n}z_{n1} = 1 \\ m_{11}z_{12} + m_{12}z_{22} + \ldots + m_{1n}z_{n2} = 0 \\ \vdots \\ m_{n1}z_{1(n-1)} + m_{n2}z_{2(n-1)} + \ldots + m_{nn}z_{n(n-1)} = 0 \\ m_{n1}z_{1n} + m_{n2}z_{2n} + \ldots + m_{nn}z_{nn} = 1 \end{cases}$$

Assim, digamos que a seguinte matriz B é a inversa de A:

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix}$$

Assim, para que B seja a sua inversa, é necessário que o seguinte sistema tenha soluções únicas:

$$\begin{cases} a_1 b_{11} + 0 \times b_{21} + \ldots + 0 \times b_{n1} = 1 \\ a_1 b_{12} + 0 \times b_{22} + \ldots + 0 \times b_{n2} = 0 \\ \vdots \\ 0 \times b_{1(n-1)} + 0 \times b_{2(n-1)} + \ldots + a_n b_{n(n-1)} = 0 \\ 0 \times b_{1n} + 0 \times b_{2n} + \ldots + a_n b_{nn} = 1 \end{cases} = \begin{cases} a_1 b_{11} = 1 \\ a_1 b_{12} = 0 \\ \vdots \\ a_n b_{n(n-1)} = 0 \\ a_n b_{nn} = 1 \end{cases}$$

Assim, teremos que os elementos b_{ii} de B deverão ser $\frac{1}{a_{ii}}$, enquanto que todos os outros $n-1$ elementos de sua linha/coluna devem ser 0, já que todos os elementos da diagonal de A são diferentes de 0 (para que A sejá inversível). Assim, teremos que a matriz B, que é inversa de A, será:

$$B = \begin{bmatrix} \frac{1}{a_1} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{a_2} & 0 & \dots & 0 \\ 0 & 0 & \frac{1}{a_3} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \frac{1}{a_n} \end{bmatrix}$$

Que também é uma matriz diagonal.

2 Questão 2

2.1 Pergunta

Seja

$$\mathbf{X} = \begin{bmatrix} | & | & & | \\ x_{.1} & x_{.2} & \dots & x_{.p} \\ | & | & & | \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2k} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nk} & \dots & x_{np} \end{bmatrix}$$

Mostre que, se as colunas de $\mathbf{X}$ forem **ortogonais** entre si, isto é, $x_{.k}.x_{.l} = 0$ (produto interno entre as colunas é 0) para todo $k \neq l$, então a matriz $\mathbf{X}^\top \mathbf{X}$ é diagonal.

2.2 Resposta

O produto interno entre duas colunas da matriz X será dado por:

$$\langle x_{ia}, x_{ib} \rangle = \sum_{i=1}^{n} x_{ia} x_{ib} = 0 \tag{1}$$

Teremos que $\mathbf{X}^\top\mathbf{X}$ será:

$$X^\top X = \begin{bmatrix} x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & \dots & x_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1p} & x_{2p} & \dots & x_{np} \end{bmatrix} \times \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^n x_{i1}^2 & \sum_{i=1}^n x_{i1}x_{i2} & \dots & \sum_{i=1}^n x_{i1}x_{ip} \\ \sum_{i=1}^n x_{i2}x_{i1} & \sum_{i=1}^n x_{i2}^2 & \dots & \sum_{i=1}^n x_{i2}x_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n x_{ip}x_{i1} & \sum_{i=1}^n x_{ip}x_{i2} & \dots & \sum_{i=1}^n x_{ip}^2 \end{bmatrix}$$

Dessa forma, todos os elementos que não estão na diagonal principal de $\mathbf{X}^\top\mathbf{X}$ serão o produto interno entre duas colunas diferentes de X (como visto em (1)), então todos eles serão zero, pelo fato de que as colunas são ortogonais entre si. Assim, a matriz $X^\top X$ será:

$$X^\top X = \begin{bmatrix} \sum_{i=1}^n x_{i1}^2 & 0 & \dots & 0 \\ 0 & \sum_{i=1}^n x_{i2}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sum_{i=1}^n x_{ip}^2 \end{bmatrix}$$

Que é uma matriz diagonal.

3 Questão 3

3.1 Pergunta

No modelo de regressão linear múltipla $\mathbf{Y} = \mathbf{X}\beta + \epsilon$, sob ϵ com distribuição normal multivariada, o que podemos dizer sobre a distribuição dos estimadores de mínimos quadrados $\hat{\beta}$ quando as colunas de $\mathbf{X}$ são ortogonais entre si?

3.2 Resposta

Sabemos que a distribuição de $\hat{\beta}$, sob ϵ com distribuição normal multivariada, é:

$$\hat{\beta} \sim N_{p+1}(\beta, \sigma^2(X^\top X)^{-1})$$

Assim, pelo visto na primeira questão, teremos que $(X^\top X)^{-1}$ será:

$$(X^\top X)^{-1} = \begin{bmatrix} \frac{1}{\sum_{i=1}^n x_{i1}^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\sum_{i=1}^n x_{i2}^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sum_{i=1}^n x_{ip}^2} \end{bmatrix}$$

Desse modo, a variância de um determinado $\hat{\beta}_k \in \hat{\beta}$ depende somente dos valores da coluna k da matriz X do modelo, visto que a matriz de variância-covariância de $\hat{\beta}$ será diagonal. Além disso, como $\hat{\beta} = (X^\top X)^{-1} X^\top Y$, a estimação de cada um dos β_i's será feita de maneira independente, e a sua inclusão/exclusão no modelo não afetará a estimação dos demais β's.

4 Questão 4

4.1 Pergunta

Para um modelo de regressão linear com intercepto mais p covariáveis, o R^2-ajustado, denotado R^2_α é definido como:

$$R^2_\alpha = 1 - \frac{(1 - R^2)(n-1)}{n-p-1}$$

a. Seja r o valor do R^2 para o ajuste de um determinado modelo de regressão linear com p covariáveis. Ao adicionar uma nova variável ao modelo, o novo valor de R^2 foi de $r^+ > r$. Determine, em função de n e p, quanto deve ser o aumento no R^2, $r^+ - r$, para que o R^2-ajustado também aumente.

b. Interprete o resultado do item anterior.

4.2 Resposta

4.2.1 a)

Para que haja um aumento no R^2_α, a seguinte desigualdade deve ser verdadeira:

$$1 - \frac{(1-r^+)(n-1)}{(n-p-2)} > 1 - \frac{(1-r)(n-1)}{(n-p-1)}$$

$$\frac{(1-r^+)(n-1)}{(n-p-2)} < \frac{(1-r)(n-1)}{(n-p-1)}$$

$$\frac{(1-r^+)}{(n-p-2)} < \frac{(1-r)}{(n-p-1)}$$

$$\begin{aligned}
(1-r^+)(n-p-1) &< (1-r)(n-p-2) \\
(1-r^+)(n-p-1) &< (1-r)(n-p-1) - (1-r) \\
(1-r^+)(n-p-1) - (1-r)(n-p-1) &< -(1-r) \\
(n-p-1)((1-r^+) - (1-r)) &> (1-r) \\
(r-r^+) &> \frac{(1-r)}{(n-p-1)}
\end{aligned}$$

Assim, para que haja um aumento no R^2-ajustado, é necessário que o aumento no R^2 seja maior do que $\frac{1-r}{n-p-1}$.

4.2.2 b)

Podemos ver que, para que haja aumento no R^2-ajustado, a inclusão da nova covariável deve implicar um ganho de explicação na variância maior do que o modelo antigo, ponderado pela nova quantidade de graus de liberdade dos resíduos (que será $n-p-2$).