UMA ABORDAGEM GEOMÉTRICA DO ESTIMADOR DE CUMEEIRA DE C. R. RAO

Luzia Aparecida da COSTA[1]
Lucas Monteiro CHAVES[2]
Devanil Jaques de SOUZA[2]

▪ RESUMO: A geometria do estimador de cumeeira de Rao é apresentada e algumas relações com o estimador de quadrados mínimos são explicitadas. Para combinações lineares dos parâmetros um elipsóide é apresentado em que a utilização do estimador de cumeeira de Rao gera erros quadráticos médios menores do que o equivalente para quadrados mínimos. Um exemplo estudado na literatura é apresentado.

▪ PALAVRAS-CHAVE: Estimador de cumeeira; estimador de quadrados mínimos; combinação linear de parâmetros; erro quadrático.

1 Introdução

Dado o modelo linear de regressão:

$$Y_{n\times 1} = X_{n\times p}\beta_{p\times 1} + \varepsilon_{n\times 1}, cov(Y) = \sigma^2 I$$

quando $X'X$ apresenta alto grau de colinearidade, ou seja, determinante de $X'X$ próximo de zero, o estimador de quadrados mínimos do vetor de parâmetros

$$\hat{\beta} = (X'X)^{-1}X'Y$$

possui componentes com variâncias elevadas, o que torna em algumas aplicações, inviável a sua utilização. No sentido de contornar esse problema Hoerl & Kennard (1970) propuseram estimadores para β denominados estimadores

[1]Universidade Federal de Lavras – UFLA, Programa de Pós-graduação em Estatística e Experimentação Agronômica, CEP: 37200-000, Lavras, MG, Brasil. E-mail: *luziamatematica@hotmail.com*
[2]Universidade Federal de Lavras - UFLA, Departamento de Ciências Exatas - DEX, CEP: 37200-000, Lavras, MG, Brasil, E-mail: *lucas@dex.ufla.br* / devaniljaques@dex.ufla.br.

de cumeeira. A ideia consiste em que para diminuir a colinearidade da matriz $X'X$ soma-se a ela uma matriz diagonal kI, $k > 0$. Ou de uma maneira um pouco mais geral, soma-se uma matriz diagonal D com entradas positivas.

O estimador obtido é,

$$\hat{\beta}(k) = (X'X + kI)^{-1}X'Y,$$

ou

$$\hat{\beta}(D) = (X'X + D)^{-1}X'Y.$$

Em 1975 *C.R.Rao* baseado nas ideias de Hoerl & Kennard (1970) propôs uma generalização dos estimadores de cumeeira em que a diminuição da colinearidade da matriz $X'X$ é obtida somando-se uma matriz simétrica positiva definida qualquer G. O estimador de cummeira de *Rao* denotado por $\hat{\beta}_R$, (em geral o contexto deixa claro qual é a matriz G utilizada) tem a expressão:

$$\hat{\beta}_R = (X'X + kG)^{-1}X'Y.$$

2 A geometria do estimador de cumeeira de Rao

O estimador de cumeeira de $Hoerl$ & $Kennard$ pode ser obtido a partir de uma construção geométrica.

Seja $b = b(y_1,y_2,...,y_n)$ um estimador qualquer do vetor de parâmetros β. Se o vetor β é o verdadeiro vetor de parâmetros e y o vetor de dados, então o erro de estimação é $b - \beta$.

Portanto no espaço de dados tem-se $X\beta$ o verdadeiro vetor de médias e Xb a estimativa do vetor de médias e portanto um erro residual ao quadrado de

$$||X\beta - Xb||^2 = ||X(\beta - b)||^2 = (b - \beta)'X'X(b - \beta).$$

Vamos fixar um valor c para o resíduo ao quadrado. Neste caso, os valores possíveis do estimador b, com o mesmo erro residual, satisfazem a equação:

$$(b - \beta)'X'X(b - \beta) = c.$$

Esta equação define uma esfera no subespaço imagem de X, $Im(X)$ do espaço de dados, em que a variável é $X\beta$,

$$||X\beta - Xb||^2 = c.$$

No espaço de parâmetros a mesma equação, em que a variável é b, define uma elipse

$$(b - \beta)'X'X(b - \beta) = c.$$

Os eixos desta elipse possuem comprimentos dados por $\frac{1}{\sqrt{\lambda}}$ em que λ são autovalores de $X'X$. Em razão da quasecolinearidade de $X'X$, tem-se pelo menos um autovalor próximo de zero e portanto pelo menos um dos eixos é bastante alongado.

Na definição do estimador b vamos adotar uma atitude conservadora no seguinte sentido. Como todos os pontos sobre a elipse levam a um mesmo erro residual, toma-se como valor de $b = b(y_1,y_2,...,y_n)$, o ponto sobre a elipse mais próximo da origem.

Rao generalizou esta construção tomando como valor de $b = b(y_1,...,y_n)$ o ponto sobre a elipse mais próximo da origem, só que agora considerando uma distância definida por outra métrica $< , >_G$.

Da mesma forma que o estimador de cumeeira usual, o estimador de cumeeria de *Rao* pode ser obtido de forma natural a partir de problemas variacionais, definidos geometricamente.

Seja G uma matriz simétrica positiva definida. Considere no espaço de parâmetros $\mathbb{R}^p$ o produto interno dado por:

$$< v,w >_G = v'Gv.$$

Observe que esta métrica nos dá uma soma de quadrados ponderada, isto é, uma distância com pesos.

Note que os pontos equidistantes à origem na métrica $< , >_G$ definem elipses $b'Gb = d$.

Desta forma a definição do estimador de cumeeira de *Rao* é dada como solução de um problema de mínimo condicionado.

Minimizar $b'Gb$ restrito à elipse $(b-\beta)'X'X(b-\beta) = c$, equivale a minimizar a Lagrangiana:

$$L = b'Gb + \lambda((b-\beta)'X'X(b-\beta) - c).$$

Derivando em relação a λ e a b temos o sistema de equações:

$$\frac{\partial L}{\partial b} = Gb + \lambda X'X\,(b-\beta) = 0 \Rightarrow (G + \lambda X'X)b = \lambda X'X\beta$$

$$\frac{\partial L}{\partial \lambda} = (b-\beta)'\,X'X\,(b-\beta) - c = 0.$$

A solução do sistema é:

$$b = \left(\frac{1}{\lambda}G + X'X\right)^{-1} X'X\beta.$$

O valor de λ é obtido substituindo-se o valor de b na segunda equação

$$(b-\beta)'\,X'X\,(b-\beta) - c = 0,$$

e então tem-se:

$$\left(\left(\frac{1}{\lambda}G + X'X\right)^{-1} X'X\beta - \beta\right) X'X \left(\left(\frac{1}{\lambda}G + X'X\right)^{-1} X'X\beta - \beta\right) = c.$$

O estimador de cumeeira de Rao é agora obtido substituindo-se o parâmetro β pelo seu estimador de quadrados mínimos $\hat{\beta}$.

$$\begin{aligned}
\hat{\beta}_R &= \left(\frac{1}{\lambda}G + X'X\right)^{-1} X'X\hat{\beta} \\
&= \left(\frac{1}{\lambda}G + X'X\right)^{-1} X'X(X'X)^{-1}X'Y \\
&= \left(\frac{1}{\lambda}G + X'X\right)^{-1} X'Y \\
&= (kG + X'X)^{-1} X'Y,
\end{aligned}$$

em que $k = \dfrac{1}{\lambda}$.

Geometricamente o estimador é obtido como o ponto de tangência de uma elipse definida pela matriz G centrada na origem com uma elipse definida por $X'X$ centrada na estimativa de quadrados mínimos. A construção é descrita pela Figura 1:

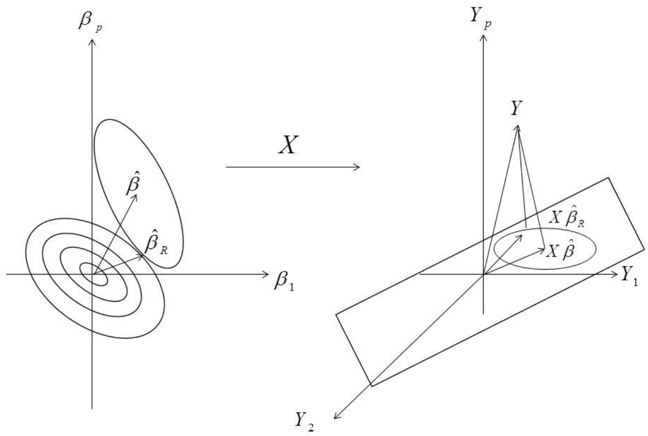

Figura 1 - Interpretação geométrica do estimador de cumeeira de Rao.

2.1 Casos particulares

Se $G = I$ então $\hat{\beta}_R = (kI + X'X)^{-1} X'Y$ é igual ao estimador de cumeeira usual de Hoerl & Kennard (1970), ou de forma mais geral se G é uma matriz diagonal D então $\hat{\beta}_R = (D + X'X)^{-1} X'Y$, denominado estimador de cummeira generalizado.

Se G é a própria matriz $X'X$ tem-se:

$$\begin{aligned}\hat{\beta}_R &= (kX'X + X'X)^{-1} X'Y \\ &= \frac{1}{1+k}(X'X)^{-1}X'X\hat{\beta} \\ &= \frac{1}{1+k}\hat{\beta}.\end{aligned}$$

Este estimador foi proposto por Mayer & Wilke (1973).

2.2 Algumas relações entre o estimador de cumeeira de Rao e o estimador de quadrados mínimos

Várias medidas de eficiência de estimadores são utilizadas quando se comparam estimadores. Vamos considerar duas delas. Dado um estimador b de β, seu erro quadrático médio EQM é definido por:

$$EQM(b) = E[(b-\beta)'(b-\beta)] = trE[(b-\beta)(b-\beta)'] = tr(b-\beta)(b-\beta)' = trM(b),$$

em que $M(b) = E\left[(b-\beta)(b-\beta)'\right]$ é a matriz de momentos do estimador b.

Dado um vetor p no espaço de parâmetros, tem-se a combinação linear das componentes do vetor β, $p'\beta$. Estimando $p'\beta$ por $p'b$, o erro quadrático médio deste estimador é:

$$\begin{aligned}EQM_p(b) &= E[(p'\beta - p'b)^2] \\ &= E[p'(b-\beta)(b-\beta)'p] \\ &= p'E[(b-\beta)(b-\beta)']p \\ &= p'M(b)p.\end{aligned}$$

Um estimador b_1 será dito melhor no sentido das combinações lineares dos parâmetros em relação a um outro estimador b_2 se para toda combinação linear $p'\beta$, $p'b_1$ tem erro quadrático médio menor ou igual que o de $p'b_2$, isto é, $EQM_p(b_1) \leq EQM_p(b_2)$ $\forall p \in \mathbb{R}^p$.

Para obter expressões para o erro quadrático médio, variância e viés dos estimadores de cumeeira de *Rao*, considere a decomposição da matriz de momentos do estimador $\hat{\beta}_R$:

$$
\begin{aligned}
M(\hat{\beta}_R) &= E[(\hat{\beta}_R - \beta)(\hat{\beta}_R - \beta)'] \\
&= E[(\hat{\beta}_R - E(\hat{\beta}_R))(\hat{\beta}_R - E(\hat{\beta}_R))'] + (\beta - E(\hat{\beta}_R)(\beta - E(\hat{\beta}_R))' \\
&= D(\hat{\beta}_R) + B(\hat{\beta}_R),
\end{aligned}
$$

em que D é a matriz de variância e covariância do estimador e B é a matriz cujo traço é o viés ao quadrado. Observe que:

$$
EQM_p(\hat{\beta}_R) = p'M(\hat{\beta}_R)p = p'(D(\hat{\beta}_R) + B(\hat{\beta}_R))p = p'D(\hat{\beta}_R)p + p'B(\hat{\beta}_R)p.
$$

$$
\begin{aligned}
D(\hat{\beta}_R) &= cov(\hat{\beta}_R) \\
&= cov\left((kG + X'X)^{-1} X'X\hat{\beta}\right) \\
&= (kG + X'X)^{-1} X'Xcov(\hat{\beta})X'X\,(kG + X'X)^{-1} \\
&= \sigma^2\,(kG + X'X)^{-1} X'X(X'X)^{-1}X'X\,(kG + X'X)^{-1} \\
&= \sigma^2\,(kG + X'X)^{-1} X'X\,(kG + X'X)^{-1}.
\end{aligned}
$$

Como

$$
\begin{aligned}
E\left[\widehat{\beta}_R\right] &= E\left[(kG + X'X)^{-1} X'Y\right] \\
&= (kG + X'X)^{-1} X'E\,[Y] \\
&= (kG + X'X)^{-1} X'X\beta,
\end{aligned}
$$

tem-se:

$$
\begin{aligned}
B(\hat{\beta}_R) &= \left(\beta - E\left[\widehat{\beta}_R\right]\right)\left(\beta - E\left[\widehat{\beta}_R\right]\right)' \\
&= \left(\beta - (kG + X'X)^{-1}X'X\beta\right)\left(\beta - (kG + X'X)^{-1}X'X\beta\right)' \\
&= \left[\left(I - (kG + X'X)^{-1}X'X\right)\beta\right]\left[\beta'\left(I - X'X(kG + X'X)^{-1}\right)\right] \\
&= \left[\left(I - (kG + X'X)^{-1}X'X\right)\right](\beta\beta')\left[\left(I - X'X(kG + X'X)^{-1}\right)\right].
\end{aligned}
$$

Utilizando a identidade

$$
\begin{aligned}
(I - kG + X'X)^{-1} X'X &= (kG + X'X)^{-1}(kG + X'X) - (kG + X'X)^{-1} X'X \\
&= (kG + X'X)^{-1}[(kG + X'X) - X'X] \\
&= (kG + X'X)^{-1}(kG),
\end{aligned}
$$

tem-se:

$$B(\hat{\beta}_R) = (kG + X'X)^{-1} (kG) \beta\beta' (kG) (kG + X'X)^{-1}$$

Segue então que:

$$\begin{aligned} M(\hat{\beta}_R) &= D(\hat{\beta}_R) + B(\hat{\beta}_R) \\ &= (X'X + kG)^{-1} \left[\sigma^2 (X'X)^{-1} + (kG) \beta\beta' (kG)\right] (X'X + kG)^{-1}. \end{aligned}$$

Observe que para o estimador de quadrados mínimos

$$M(\hat{\beta}) = D(\hat{\beta}) = \sigma^2 (X'X)^{-1}.$$

Vamos agora mostrar que para uma região no espaço paramétrico definida por uma elipse, se β pertence a esta elipse então, $EQM_p(\hat{\beta}_R) \leq EQM_p(\hat{\beta})$ para todo vetor p pertencente ao $\mathbb{R}^p$. Este fato implica então que para vetores β razoavelmente próximos da origem se tem uma superioridade do estimador de cumeeira de *Rao* em relação ao estimador de quadrados mínimos, segundo o critério do erro quadrático médio das funções paramétricas lineares.

Teorema 1.

$$E\left[\left(p'\widehat{\beta}_R - p'\beta\right)^2\right] \leq E\left[\left(p'\widehat{\beta} - p'\beta\right)^2\right] \forall p \in \mathbb{R}^p \Leftrightarrow \beta'\left(2G^{-1} + (X'X)^{-1}\right)^{-1}\beta \leq \sigma^2$$

A demonstração deste teorema é baseada em algumas propriedades das matrizes positivas definidas. Primeiramente temos um resultado que generaliza a desigualde de $Cauchy - Schwartz$:

$$(v'w)^2 \leq (v'v)(w'w)$$

Lema 1. Se A é uma matriz simétrica positiva definida então:

$$(v'w)^2 \leq (v'Av)\left(w'A^{-1}w\right)$$

Demonstração: A matriz A é diagonalizável e portanto existe matriz ortogonal Q tal que $Q'AQ = \Lambda$ é uma matriz diagonal com entradas iguais aos autovalores de A que são todos positivos. Seja $A^{\frac{1}{2}} = Q'\Lambda^{\frac{1}{2}}Q$ a raiz quadrada de A, logo:

$$\begin{aligned} (v'w)^2 &= (v'A^{\frac{1}{2}}A^{-\frac{1}{2}}w)^2 \\ &= \left(\left(A^{\frac{1}{2}}v\right)'\left(A^{-\frac{1}{2}}w\right)\right)^2 \\ &\leq \left(A^{\frac{1}{2}}v\right)'\left(A^{\frac{1}{2}}v\right)\left(A^{-\frac{1}{2}}w\right)'\left(A^{-\frac{1}{2}}w\right) \\ &= \left(v'A^{\frac{1}{2}}A^{\frac{1}{2}}v\right)\left(w'A^{-\frac{1}{2}}A^{-\frac{1}{2}}w\right) \\ &= (v'Av)\left(w'A^{-1}w\right). \end{aligned}$$

Tem-se então um interessante resultado de Álgebra Linear.

Lema 2. (Farebrother): Seja A uma matriz positiva definida, α um escalar positivo e q um vetor não nulo. Então uma condição suficiente necessária e para que $\alpha A - qq'$ seja positiva definida é $q'A^{-1}q < \alpha$.

Demonstração: $\alpha A - qq'$ é positiva definida se e somente se $\forall\ v \neq 0$:

$$\begin{aligned} 0 &< v'\left(\alpha A - qq'\right)v \\ &= \alpha v'Av - v'\left(qq'\right)v \\ &= \alpha v'Av - \left(v'q\right)\left(q'v\right) \\ &= \alpha v'Av - \left(q'v\right)^2 \\ &\Rightarrow \alpha > \frac{\left(v'q\right)^2}{v'Av}. \end{aligned}$$

Portanto $\alpha A - qq'$ é positiva definida se e somente se $\alpha > \frac{\left(v'q\right)^2}{v'Av}$. Dividindo ambos os lados por $q'A^{-1}q$ tem-se:

$$\frac{\alpha}{q'A^{-1}q} > \frac{\left(q'v\right)'\left(q'v\right)}{v'Avq'A^{-1}q}. \tag{1}$$

Como pelo Lema 1

$$(v'q)^2 \leq (v'Av)(q'A^{-1}q) \Rightarrow \frac{(v'q)^2}{(v'Av)(q'A^{-1}q)} \leq 1,$$

tem-se

$$\frac{\alpha}{q'A^{-1}q} > 1 \Leftrightarrow \alpha > q'A^{-1}q.$$

Uma outra observação necessária é que uma matriz $A - B$ é positiva definida se e somente se $H'AH - H'BH$ também é positiva definida para toda matriz inversível H.

De fato:

$$v'\left(H'AH - H'BH\right)v = v'H'\left(A - B\right)Hv = (Hv)'(A - B)Hv,$$

logo, para $\forall$ matriz H inversível tem-se que para $\forall v \neq 0$

$$v'(A - B)v > 0 \Leftrightarrow v'(H'AH - H'BH)v > 0.$$

Agora é possível provar o Teorema 1.

Demonstração:

$$M(\hat{\beta}) - M(\hat{\beta}_R) = \sigma^2\left(X'X\right)^{-1} - \left(X'X + kG\right)^{-1}\left[\sigma^2\left(X'X\right)^{-1} + (kG)\,\beta\beta'\,(kG)\right]\left(X'X + kG\right)^{-1}$$

é positiva definida se e somente se $(X'X + kG)\left(M(\hat{\beta}) - M(\hat{\beta}_R)\right)(X'X + kG)$ é positiva definida.

Assim sendo,

$$(X'X + kG)\,\sigma^2\,(X'X)^{-1}\,(X'X + kG) - \left[+\sigma^2\,(X'X) + (kG)\,(\beta\beta')\,(kG)\right]$$

é positiva definida.

Desenvolvendo tem-se:

$$\sigma^2\,(X'X) + (kG)\,\sigma^2\,(X'X)^{-1}\,(kG) + X'X\sigma^2\,(X'X)^{-1}\,kG +$$
$$+\,(kG)\,\sigma^2\,(X'X)^{-1}\,(X'X) - \sigma^2\,(X'X) - (kG)\,(\beta\beta')\,(kG) =$$

$$= \sigma^2\,(kG)\,(X'X)\,(kG) + 2\sigma^2\,(kG) - (kG)\,\beta\beta'\,(kG)\,.$$

Novamente esta matriz será positiva definida se e somente se

$$(kG)^{-1}\left[\sigma^2\,(kG)\,X'X\,(kG) + 2\sigma^2 kG - (kG)\,\beta\beta'\,(kG)\right](kG)^{-1}$$

for positiva definida, isto é, se

$$\sigma^2\,(X'X)^{-1} + 2\sigma^2\,(kG)^{-1} - \beta\beta'$$

for positiva definida. Mas, pelo lema 2,

$$\sigma^2\left[(X'X)^{-1} + \frac{2}{k}G^{-1}\right] - \beta\beta'$$

é positiva definida se e somente se

$$\sigma^2 > \beta'\left[(X'X)^{-1} + \frac{2}{k}G^{-1}\right]^{-1}\beta.$$

Tem-se então que para um dado valor da variância σ^2, fica definida uma região no espaço paramétrico limitada por um elipsóide tais que para todo vetor β nesta região, suposto ser o vetor populacional, o erro quadrático médio das funções paramétricas para o estimador de cumeeira de *Rao*, $E\left[\left(p'\widehat{\beta}_R - p'\beta\right)^2\right]$, é menor do que o de quadrados mínimos $E\left[\left(p'\widehat{\beta} - p'\beta\right)^2\right]$.

Este elipsóide define então a região em que a abordagem utilizando quadrados mínimos deve ser evitada (Figura 2).

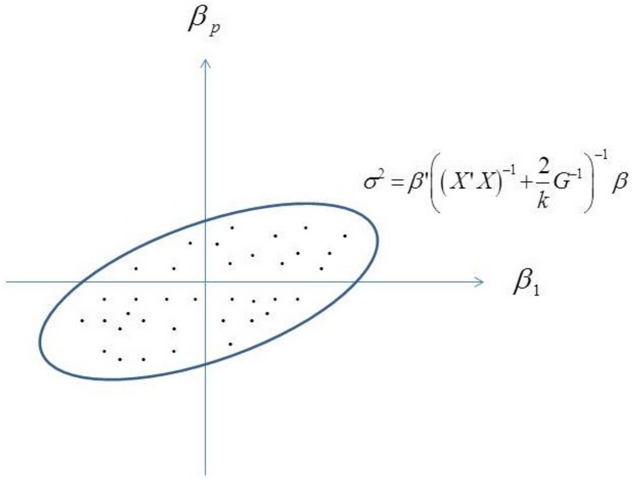

Figura 2 - Região de utilização do estimador de cumeeira de Rao.

É interessante observar que tal região não depende da direção do vetor p.

Para o estimador de cumeeira usual a região é:

$$\beta'\left(\frac{2}{k}I + (X'X)^{-1}\right)^{-1}\beta \leq \sigma^2.$$

Para o estimador de cumeeira generalizado

$$\beta'\left((X'X)^{-1} + 2(D)^{-1}\right)^{-1}\beta \leq \sigma^2,$$

em que D é uma matriz diagonal.

Para o estimador de cumeeira de $Mayer - Wilke$

$$\beta'\left((X'X)^{-1} + \frac{2}{k}(X'X)^{-1}\right)^{-1}\beta = \beta'\left(\frac{k+2}{k}(X'X)^{-1}\right)^{-1}\beta \leq \sigma^2$$

$$\Rightarrow \beta'\,(X'X)\,\beta \leq \frac{k}{k+2}\sigma^2.$$

Se utilizarmos coordenadas $\gamma = Q\beta$ tais que nestas coordenadas $X'X = \Lambda$ é a matriz diagonal com entradas os autovalores λ_i, pode-se explicitar as equações da região do elipsóide da forma:

$$\sum \frac{\gamma_i}{c_i^2} \leq 1.$$

1. Estimador de cumeeira usual

$$\gamma'\left(\Lambda^{-1}+\frac{2}{k}I\right)^{-1}\gamma\leq\sigma^2\Rightarrow c_i=\sigma\sqrt{\frac{2}{k}+\frac{1}{\lambda_i}}.$$

2. Estimador de cumeeira generalizado

$$\gamma'\left(\Lambda^{-1}+\frac{2}{k}D^{-1}\right)^{-1}\gamma\leq\sigma^2\Rightarrow c_i=\sigma\sqrt{\frac{2}{k_i}+\frac{1}{\lambda_i}}.$$

3. Estimador de cumeeira $Mayer-Wilke$

$$\gamma\Lambda\gamma\leq\frac{k\sigma^2}{k+2}\Rightarrow c_i=\sigma\sqrt{\frac{1}{\lambda_i}\left(\frac{k+2}{k}\right)}.$$

3 Um exemplo numérico

O seguinte exemplo foi apresentado detalhadamente em $Marquardt$ (1970). O objetivo é a comparação entre estimativas obtidas via quadrados mínimos, por regressão de cumeeira, utilizando inversas generalizadas e utilizando o método de seleção de covaríaveis.

Neste trabalho a comparação entre as estimativas será feita para quadrados mínimos e regressão de cumeeira como no artigo de $Marquardt$ (1970), porém agora comparando também com a regressão obtida com o estimador de Rao.

Considere o modelo linear:

$$Y=X\beta+\varepsilon$$

com:

$$X=\begin{pmatrix}\frac{3\sqrt{2}}{10} & \frac{4\sqrt{2}}{10}\\ \frac{4\sqrt{2}}{10} & \frac{3\sqrt{2}}{10}\\ \frac{5\sqrt{2}}{10} & \frac{5\sqrt{2}}{10}\end{pmatrix}\ e\ \ y=\begin{pmatrix}1\\2\\3\end{pmatrix}$$

tem-se que: $X'X=\begin{pmatrix}1 & \frac{49}{50}\\ \frac{49}{50} & 1\end{pmatrix}$ e os autovalores da matriz desta matriz são 1,98 e 0,02. O autovalor 0,02 caracteriza ocorrência de quase colinearidade.

A estimativa de quadrados mínimos é dada por

$$\hat{\beta}=(X'X)^{-1}X'y=\begin{pmatrix}5{,}3569\\-1{,}7142\end{pmatrix}$$

e a estimativa de quadrados mínimos para a variância $\hat{\sigma}^2=0{,}3636$.

Considere a elipse centrada no estimador de quadrados mínimos $\hat{\beta}$ tal que para os parâmetros β^* pertencentes a este elipsóide, a soma de quadrados de resíduos é constante e igual a $||X\beta^* - Y||^2 = 0{,}8636$. A equação desta elipse é dada por

$$\left(\beta^* - \begin{pmatrix} 5{,}3569 \\ -1{,}7142 \end{pmatrix}\right)' X'X \left(\beta^* - \begin{pmatrix} 5{,}3569 \\ -1{,}7142 \end{pmatrix}\right) = 0{,}5234.$$

No artigo de $Marquardt$(1970) é obtido que, para $k = 0{,}2$, a estimativa de cumeeira é:

$$\hat{\beta}(0{,}2) = (X'X + 0{,}2I)^{-1}X'y = \begin{pmatrix} 1{,}9757 \\ 1{,}3328 \end{pmatrix}$$

e que $\hat{\beta}(0{,}2)$ pertence à esta elipse. Note que $\hat{\beta}(0{,}2)$ também pertence ao círculo de raio $\sqrt{(1{,}9757)^2 + (1{,}3328)^2}$.

Para obtermos o estimador de cumeeira de Rao uma matriz G deve ser escolhida. A escolha adequada da matriz G deve ser tal que sua adição à matriz $X'X$ diminua de forma considerável a multicolinearidade. Será utilizada a matriz $G = \begin{pmatrix} 1 & 0{,}8 \\ 0{,}8 & 2 \end{pmatrix}$, no entanto é necessário uma investigação futura para fundamentar teoricamente uma escolha adequada. Utilizando um software de computação numérica obteve-se que para $k = 0{,}197$ a estimativa de cumeeira de Rao:

$$\hat{\beta}_G = (X'X + 0{,}197G)^{-1}X'y = \begin{pmatrix} 2{,}9203 \\ 0{,}1380 \end{pmatrix}$$

também pertence a esta elipse.

A estimativa de cumeeira de Rao pertence a uma elipse centrada na origem dada por

$$\beta' \begin{pmatrix} 1 & 0{,}8 \\ 0{,}8 & 2 \end{pmatrix} \beta = 9{,}2116.$$

Tem-se então as estimativas quadrados mínimos, cumeeira, cumeeira de Rao e a estimativa utilizando inversa generalizada, respectivamente

$$\begin{pmatrix} 5{,}3569 \\ -1{,}7142 \end{pmatrix}, \begin{pmatrix} 1{,}9757 \\ 1{,}3328 \end{pmatrix}, \begin{pmatrix} 2{,}9203 \\ 0{,}1380 \end{pmatrix} \text{e} \begin{pmatrix} 1{,}8213 \\ 1{,}8213 \end{pmatrix}.$$

Pode-se observar então que os três métodos de encolhimento obtiveram estimativas com ambas coordenadas positivas, resultado aparentemente mais de acordo com os dados observados como discutido no artigo de $Marquadart$, (1970).

A interpretação geométrica destas estimativas está descrita na Figura 3.

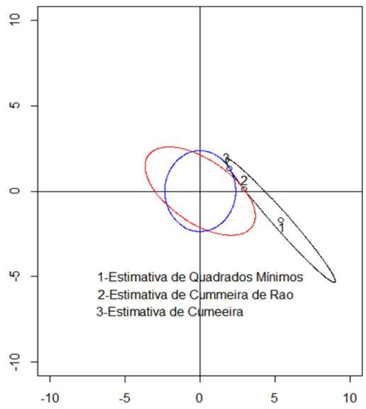

Figura 3 - Interpretação geométrica das estimativas.

L. A. COSTA; L. M. CHAVES; D. J. SOUZA. A geometric approach the ride estimator of C. R. Rao. *Rev. Bras. Biom.,* São Paulo, v.32, n.1, p.28-41, 2014.

▪ *ABSTRACT: The geometry of the ridge estimator proposed by Rao, as well as its connections with ordinary least squares estimator, are presented. For linear combinations of the parameters, it is shown the ellipsoid where Rao's ridge estimator has lower errors than the OLS estimator. It is also shown an example studied in the literature.*

▪ *KEYWORDS: Ridge estimator; minimum square estimator; linear combination of parameters; square error.*

Referências

GRUBER, M. H. J.; *Regression estimators: a comparative Study* 2.ed. Johns Hopkins: Baltimore, Maryland, 2010. 412p.

HOERL, A. E., KENNARD, R.; *Ridge Regression: Biased Estimation for Nonorthogonal Problems.* Technometrics, American Statistical Associoation, v.12, n.1, p.55-67, 1970.

MARQUARDT, D. W.; *Generalized Inverses, Ridge Regression, Biased Linear Estimation, and Nonlinear Estimation.* Technometrics, American Statistical Associoation & American Society for Quality, v.12, n.3, p.591-612, 1970.

RAO, C. R. TOUTENBURG, H.; *Linear Models, Least Squares and Alternatives* 2.ed. Springer Series in Statistics. Springer-Verlag: New York, 1999. 427p.

RAO, C. R.; *Simultaneous Estimation of Parameters in Different Linear Models and Applications to Biometric Problems*. Biometrics, International Biometrics Society, v.31, n.2, p.545-554, 1975.

Recebido em 04.02.2014.

Aprovado após revisão em 15.04.2014.