ESTIMADORES VIESADOS PARA MODELOS DE REGRESSÃO EM
PRESENÇA DE MULTICOLINEARIDADE

FERNANDO FERRARI
Lic. em Matemática

Orientador: Prof. Dr. ANTONIO FRANCISCO IEMMA

Tese apresentada à Escola Superior
de Agricultura "Luiz de Queiroz", da
Universidade de São Paulo, para
obtenção do título de Doutor em
Agronomia: Área de Concentração:
Estatística e Experimentação
Agrônômica.

PIRACICABA
ESTADO DE SÃO PAULO - BRASIL
NOVEMBRO - 1989

Ficha catalográfica preparada pela Seção de Livros da
Divisão de Biblioteca e Documentação - PCAP/USP

Ferrari, Fernando
F375e Estimadores viesados para modelos de regressão
em presença de multicolinearidade. Piracicaba,
1989.
127p.

Tese - ESALQ
Bibliografia.

1. Análise de regressão 2. Estatística matemática
3. Modelo matemático - Estimador 4. Multico-
linearidade I. Escola Superior de Agricultura Luiz
de Queiroz, Piracicaba.

CDD 519.536

ESTIMADORES VIESADOS PARA MODELOS DE REGRESSÃO EM
PRESENÇA DE MULTICOLINEARIDADE

FERNANDO FERRARI

Aprovada em: 22.12.1989

Comissão Julgadora:

Prof. Dr. Antonio Francisco Iemma

ESALQ/USP

Prof. Dr. Décio Barbin

ESALQ/USP

Prof. Dr. Roberto Simionato Moraes

ESALQ/USP

Prof. Dr. Natal Antonio Vello

ESALQ/USP

Prof. Dr. Ubirajara Dorival Diniz

UEL/Londrina

Prof. Dr. ANTONIO FRANCISCO IEMMA
Orientador

A meu pai Dair (in memoriam),
A minha mãe Maria e a meus irmãos,

ofereço

A minha esposa Maria Regina,
A meu filho Ricardo José,

dedico

AGRADECIMENTOS

Ao Dr Antonio Francisco Iemma, Professor Associado do Departamento de Matemática e Estatística da ESALQ/USP, pela orientação segura e amiga na elaboração deste trabalho.

Aos Professores do Curso de Pós-Graduação em Estatística e Experimentação Agronômica da ESALQ/USP, na pessoa de seu coordenador, Prof. Dr. Décio Barbin, pela receptividade e ensinamentos.

Ao Departamento de Análise Numérica e Estatística do IBILCE - UNESP, São José do Rio Preto, pela oportunidade concedida e pelo incentivo na realização deste curso.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, pela bolsa de estudos (PICD) concedida.

Aos funcionários do Departamento de Matemática e Estatística do ESALQ/USP, e em especial à Srta. Rosa Maria Alves e ao Sr Octávio Frassetto, pela amizade e cooperação.

Aos colegas do Curso de Pós-Graduação, pela amizade e companherismo que sempre nos proporcionaram.

A minha esposa Maria Regina, pelos constantes estímulos, solidariedade e paciência.

	Página
3.1. Material -----	36
3.2. Métodos -----	45
3.2.1. Modelo matemático -----	45
3.2.2. Estimador de mínimos quadrados -----	47
3.2.3. Estimador regressão sobre cristas ----	50
3.2.4. Estimador em componentes principais --	56
3.2.5. Escolha dos valores de k -----	62
3.2.6. Um critério para a seleção de variá- veis -----	65
3.2.7. Critérios de comparação -----	67
4. RESULTADOS E DISCUSSÃO -----	69
5. CONCLUSÕES -----	77
5.1. Em Relação ao Estimador Regressão Sobre Cristas -----	77
5.2. Em Relação ao Estimador em Componentes Principais -----	78
5.3. Quanto a escolha do estimador a ser utilizado	79
6. REFERENCIAS BIBLIOGRAFICAS -----	81
7. APÊNDICE 1 -----	91
8. APÊNDICE 2 -----	124

ESTIMADORES VIESADOS PARA MODELOS DE REGRESSÃO EM PRESENÇA DE MULTICOLINEARIDADE

AUTOR: FERNANDO FERRARI

ORIENTADOR: PROF. DR. ANTONIO FRANCISCO IEMMA

RESUMO

No presente trabalho, analisou-se o desempenho de dois estimadores viesados, a saber, o estimador sobre cristas e o estimador em componentes principais, como métodos alternativos ao estimador de mínimos quadrados no ajuste de modelos de regressão, quando os dados têm problemas de multicolinearidade.

Para tanto, considerou-se o modelo de regressão linear múltipla:

$$y = X\beta + \epsilon$$

onde,

y = vetor de observações da variável dependente na forma padronizada, de dimensões $n \times 1$.

X = Matriz dos valores das variáveis independentes na forma padronizada, de dimensões $n \times p$, de posto $p < n$;

β = Vetor de constantes desconhecidas, de dimensões $p \times 1$;

ε = vetor de erros aleatórios não observáveis, de dimensões $n \times 1$, tal que $EC \varepsilon = 0$ e $EC \varepsilon \varepsilon' = \sigma^2 I_n$.

Como as variáveis independentes e a variável dependente são consideradas na forma padronizada, $X'X$ é a matriz de correlações entre as variáveis independentes, e $X'y$ é o vetor de correlações entre as variáveis independentes e a variável dependente.

Para obter-se uma maior abrangência nas comparações entre os estimadores adotados, procedeu-se a um estudo de simulação, onde foram analisados dois modelos distintos, conforme o número p de variáveis independentes envolvidas. Também foram analisados dois valores para o número n de observações sobre cada variável independente. Em cada um desses modelos, foram considerados quatro configurações, segundo uma estrutura de correlação, que foi estabelecida entre as variáveis independentes.

Como critérios de comparações, entre os métodos de estimação abordados, consideraram-se:

1 - $L_1^2 = (b - \beta)'(b - \beta)$ para medir a precisão da estimativa, e

2 - $L_2^2 = (b - \beta)'X'X(b - \beta)$ para medir a precisão da predição, sendo que o valor b em L_1^2 e L_2^2 se refere aos

estimadores considerados, quais sejam, estimador de mínimos quadrados, de regressão sobre cristas e de regressão em componentes principais.

BIASED ESTIMATORS FOR REGRESSION MODELS IN THE PRESENCE OF MULTICOLLINEARITY

Author: FERNANDO FERRARI

Adviser: Prof. Dr. ANTONIO FRANCISCO IEMMA

SUMMARY

This work consists of fitting linear models on data having problems of multicollinearity, and analyses two biased estimators, namely, the ridge regression estimator, and the principal components estimator. These estimators are considered as alternatives to the least square estimator.

For this purpose the following multiple linear regression model was considered:

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

where,

y is an $n \times 1$ response variable vector in the standardized form;

X is an $n \times p$ full column rank matrix of n observations on p regressor variables in the standardized form;

β is a $p \times 1$ parameter-vector;

ε is a $n \times 1$ vector of unobservable random-errors, with $E(\varepsilon) = 0$ e $E(\varepsilon\varepsilon') = \sigma^2 I_n$.

As the regressor variables and the response variables are considered in the standardized form, $X'X$ is the correlation matrix between the regressor variables, and $X'y$ is the vector of correlation between the regressor variables and the response variables.

To get a broader scope in the comparisons of the adopted estimators, simulations were made, where two distinct models were analysed according to the number p of independent variables involved. Two values for the number n of observations about each independent variable were also analysed. In each of these models four configurations, according to a structure of correlation which was established among the independent variables, were considered.

The following two criteria of comparisons among the methods of estimators studied, were considered:

1 - $L_1^2 = (b - \beta)'(b - \beta)$ to measure estimation accuracy, and

2 - $L_2^2 = (b - \beta)'X'X(b - \beta)$ to measure prediction accuracy so that the values of b in L_1^2 and L_2^2 refer to the estimators considered, namely, least squares, ridge regression and principal components regression estimators.

1. INTRODUÇÃO

Atenção especial tem sido dada, a métodos alternativos de estimação de parâmetros de um modelo de regressão linear múltipla, quando a matriz de variáveis independentes é mal-condicionada. Este interesse é devido principalmente, à deficiência do estimador usual de mínimos quadrados, em tais casos, para estabelecer estimativas pontuais confiáveis para os parâmetros do modelo.

Apesar de possuir a propriedade tão desejável de variância mínima na classe dos estimadores lineares não-viesados, com as suposições usuais de Gauss-Markov sobre o modelo, o estimador de mínimos quadrados pode, todavia, ter variâncias extremamente grandes, quando os dados são aproximadamente multicolineares, resultando numa forma de mal-condicionamento para a matriz $X'X$. Uma vez que este problema pode ocorrer em diversas áreas de estudo, muitas pesquisas têm sido desenvolvidas com o intuito de obter estimadores, os quais a despeito de serem viesados, têm melhor desempenho do que o estimador de mínimos quadrados, nestas circunstâncias.

Tendo-se em vista o grande número de métodos de estimação viesada citados na literatura estatística, no

presente trabalho propôs-se como um dos objetivos, estudar dois desses métodos, quais sejam, o da regressão sobre cristas, e o da regressão em componentes principais. Tal escolha deve-se ao fato de tais métodos serem aparentemente os mais largamente utilizados como alternativas ao método dos mínimos quadrados.

Na regressão sobre cristas, introduzida por HOERL e KENNARD (1970 a), procura-se obter estimativas mais estáveis para os parâmetros do modelo, alterando-se ligeiramente o sistema de equações normais usual de mínimos quadrados, adicionando-se uma constante à diagonal da matriz $X'X$. Dentre os diversos métodos propostos, para a obtenção dessa constante, dois deles são aqui considerados.

A regressão em componentes principais, desenvolvida dentre outros por HAWKINS (1973) e GREENBERG (1975), permite através da eliminação de determinadas variáveis independentes, contornar os problemas de multicolinearidade apresentados pelos dados. Aqui também, muitos são os procedimentos de seleção de variáveis sugeridos. O procedimento proposto neste trabalho, o qual envolve também a variável dependente, leva em consideração, tanto a quase-singularidade nos dados, como a eficácia de predição de determinada variável independente sobre a resposta.

Em razão das dificuldades que em geral se apresentam no estudo das distribuições dos estimadores viesados, um trabalho de simulação pareceu de fundamental

importância para determinar que particular método mostra-se mais adequado do que outro, e em quais circunstâncias.

2. REVISÃO DE LITERATURA

2.1. O Problema da Multicolinearidade

2.1.1. Generalidades

A multicolinearidade ocorre em uma análise de regressão múltipla, quando as observações amostrais das variáveis independentes, ou combinações lineares delas são altamente correlacionadas.

Segundo NETER, WASSERMAN e KUTNER (1983), as variáveis independentes tendem a ser correlacionadas entre si, em muitas situações não experimentais, tais como, em ciências sociais e biológicas, economia, finanças, dentre outras.

Quanto às possíveis fontes de multicolinearidade, MONTGOMERY e PECK (1981) destacam as seguintes:

- 1 - O método utilizado na obtenção dos dados;
- 2 - Restrições sobre o modelo ou sobre a população;
- 3 - Especificação do modelo;
- 4 - Modelos super parametrizados.

De acordo com BELSLEY, KUH e WELCH (1980), intuitivamente pode-se entender o virtual dano que resulta de dados provenientes de variáveis multicolineares, constatando-se que tais variáveis não fornecem informações que sejam muito diferentes daquelas já inerentes nas outras variáveis. Em vista disso, torna-se difícil inferir sobre as influências individuais dessas variáveis independentes sobre a resposta.

Para MASON, GUNST e WEBSTER (1975), a questão da multicolinearidade não é simplesmente de existência, mas sim de grau. Assim, quanto mais forte for a multicolinearidade, mais adversos serão seus efeitos sobre o estimador de mínimos quadrados dos parâmetros do modelo. Esses autores ressaltam, que tais estimadores serão afetados mais pela multicolinearidade, do que pelas relações entre a resposta e as variáveis independentes.

Dentre os efeitos que a multicolinearidade pode acarretar, sobre as estimativas de mínimos quadrados dos coeficientes de regressão $\hat{\beta}$, HOERL e KENNARD (1970 a) observam que quando se considera o quadrado da distância de $\hat{\beta}$ ao parâmetro β , isto é

$$L_1^2 = (\hat{\beta} - \beta)' (\hat{\beta} - \beta) = ||\hat{\beta} - \beta||^2$$

tem-se que o valor esperado de L_1^2 será

$$E(L_1^2) = E(\hat{\beta} - \beta)' (\hat{\beta} - \beta)$$

$$= \sum_{j=1}^P E(\hat{\beta}_j - \beta_j)^2$$

$$= \sum_{j=1}^P V(\hat{\beta}_j)$$

$$= \sigma^2 \text{Tr}(X'X)^{-1}$$

implicando que, $E(L_1^2) = \sigma^2 \sum_{j=1}^P \lambda_j^{-1}$, onde $\lambda_j > 0$, $j = 1, 2, \dots, p$, são os auto valores da matriz $X'X$, considerada na forma de correlações. Então, se a matriz $X'X$ é mal-condicionada em razão da multicolinearidade, pelo menos um dos λ_j será muito pequeno, conforme SILVEY (1969), implicando que a distância da estimativa de mínimos quadrados $\hat{\beta}$ ao parâmetro β poderá ser muito grande. Ainda, segundo esses autores tem-se, equivalentemente

$$E(L_1^2) = E(\hat{\beta}'\hat{\beta} - 2\hat{\beta}'\beta + \beta'\beta)$$

$$\text{ou} \quad E(\hat{\beta}'\hat{\beta}) = \beta'\beta + \sigma^2 \text{Tr}(X'X)^{-1}$$

ou seja, o vetor $\hat{\beta}$ geralmente possui maior norma do que o vetor β , o que implica que o método dos mínimos quadrados produz coeficientes de regressão estimados, que podem ser grandes em valor absoluto.

Sob esse aspecto, WEBSTER, GUNST e MASON (1974) indicam, dentre outros, que a multicolinearidade tende a produzir estimativas de mínimos quadrados $\hat{\beta}_j$ com

variâncias muito grandes, visto que,

$$VC(\hat{\beta}_j) = (1 - R_j^2)^{-1} \sigma^2$$

onde R_j^2 é o coeficiente de determinação múltiplo da regressão de X_j sobre as variáveis independentes restantes. Daí uma forte multicolinearidade entre X_j e algum subconjunto das demais variáveis independentes, implicará em R_j^2 próximo de 1 e conseqüentemente, $VC(\hat{\beta}_j)$ será muito grande.

MANDEL (1982) por sua vez, nota que apesar do método dos mínimos quadrados, em geral produzir estimativas dos parâmetros individuais do modelo demasiadamente instáveis na presença de multicolinearidade, isto necessariamente não implica que o modelo seja mau preditor. Desde que as predições estejam restritas a regiões do espaço das variáveis independentes onde a multicolinearidade se verifica, o modelo ajustado freqüentemente produzirá predições satisfatórias.

Quando ocorrem problemas de multicolinearidade é comum utilizar-se de técnicas, tais como, a obtenção de dados adicionais, ou reespecificação do modelo, como por exemplo, a eliminação de variáveis, conforme observam MONTGOMERY e PECK (1981). Para esses autores, o uso desses procedimentos, no entanto, nem sempre é possível ou mesmo viável. Fatos como , novos dados não-usuais ou atípicos, ou ainda multicolinearidade ocasionada, não

devido a restrições sobre o modelo ou sobre a população, fazem da coleta de dados adicionais uma técnica pouco recomendável. Em relação à eliminação de variáveis, MONTGOMERY e PECK (1981) salientam que apesar de ser uma técnica em geral altamente efetiva, ela poderá não estabelecer uma solução satisfatória, se as variáveis retiradas do modelo tiverem um grande poder explicativo em relação à resposta, prejudicando sobremaneira o poder de predição do modelo. Além disso, também é observado que muitos dos procedimentos de seleção de variáveis são seriamente distorcidos pela multicolinearidade, podendo resultar em um modelo final com grau de multicolinearidade não muito inferior ao modelo original.

2.1.2. Diagnósticos de multicolinearidade

As características desejáveis de um procedimento de diagnóstico segundo MONTGOMERY e PECK (1981) resumem-se em:

- 1 - Refletir diretamente o grau de multicolinearidade do problema, e
- 2 - Fornecer informação útil para determinar que variáveis independentes estão envolvidas.

Dentre os procedimentos mais comumente utilizados para detectar multicolinearidade, conforme indicam vários autores, estão os seguintes:

2.1.2.1. Análise da matriz de correlação

NETER, WASSERMAN e KUTNER (1983) indicam, que uma maneira muito simples de se examinar a multicolinearidade é a que consiste, na análise dos elementos não-diagonais r_{ij} na matriz $X'X$. Se as variáveis independentes X_i e X_j são, aproximadamente, linearmente dependentes, então $|r_{ij}|$ será próximo de 1. Observa-se porém, que este procedimento é adequado apenas para se detectar a dependência linear aproximada entre duas variáveis independentes, ou seja, é possível três ou mais variáveis independentes estarem envolvidas em uma relação de multicolinearidade, sem que quaisquer pares dessas variáveis sejam altamente correlacionadas. Tal situação não é diagnosticada pela matriz de correlação.

2.1.2.2. Fatores de inflação da variância

Os elementos diagonais $r_{jj}^{-1} = (1 - R_j^2)^{-1}$ da matriz $R_{xx}^{-1} = (X'X)^{-1}$, denominados por MARQUARDT (1970) de fatores de inflação da variância ("variance inflation factor" : VIF) podem ser utilizados para detectar multicolinearidade, conforme MONTGOMERY e PECK (1981). Observam esses autores, que sendo a variância do j -ésimo coeficiente da regressão de mínimos quadrados, dada por $r_{jj}^{-1} \sigma^2$, então r_{jj}^{-1} pode ser interpretado, como o fator pelo

qual a variância de $\hat{\beta}_j$ é aumentada devido à presença de dependência linear entre as variáveis independentes.

NETER, WASSERMAN e KUTNER (1983) afirmam que experiências têm mostrado, que quando quaisquer dos VIF's excedem a 10, isto é freqüentemente tomado como uma indicação de que os coeficientes de regressão associados a estes valores, têm estimativas de mínimos quadrados demasiadamente influenciadas pela multicolinearidade.

BELSLEY, KUH e WELCH (1980) atentam para o fato, que do mesmo modo que a matriz de correlações, os VIF's também não indicam multicolinearidades resultantes da dependência linear aproximada de três ou mais variáveis independentes.

2.1.2.3. Auto-valores e auto-vetores de $X'X$

BELSLEY, KUH e WELCH (1980) observam que KENDALL (1967) e SILVEY (1969) sugerem a utilização dos auto-valores de $X'X$, denotados por $\lambda_1, \lambda_2, \dots, \lambda_p$, para medir o grau de multicolinearidade nos dados. Segundo esses autores, se existe uma ou mais dependências lineares aproximadas nos dados, então um ou mais auto-valores de $X'X$ serão pequenos.

MONTGOMERY e PECK (1981) notam que alguns autores preferem, no entanto, examinar o número de condição

de $X'X$, definido, dada a simetria de $X'X$, por:

$$K(x) = \frac{\lambda_{\max}}{\lambda_{\min}},$$

salientando que em geral, se o número de condição é menor do que 100, não existe problema sério com a multicolinearidade. Número de condição entre 100 e 1000 implica em multicolinearidade de moderada a forte, e se $K(x)$ é maior do que 1000, isto é indicativo de uma multicolinearidade severa.

A análise dos auto-valores também pode ser usada para identificar a natureza da dependência linear aproximada nos dados, conforme indicam BELSLEY, KUH e WELCH (1980). Para isso, consideram a decomposição da matriz $X'X$ em:

$$X'X = VAV', \quad \text{onde}$$

A: matriz diagonal de dimensões $p \times p$, cujos elementos da diagonal são os auto-valores λ_j ($j = 1, 2, \dots, p$) de $X'X$;
V: matriz ortogonal de dimensões $p \times p$, cujas colunas são os auto-vetores normalizados de $X'X$.

Então, denotando-se as colunas de V por $v_1, v_2, \dots, v_p$, se o auto-valor λ_j é próximo de zero, indicando uma dependência linear aproximada nos dados, os elementos do auto-vetor associado v_j descrevem a natureza desta dependência linear.

2.1.2.4. Decomposição em valores singulares

De acordo com BELSLEY, KUH e WELCH (1980), qualquer matriz X de dimensões $n \times p$ (n observações e p variáveis), pode ser decomposta em

$$X = UDV'$$

onde U tem dimensões $n \times p$, V tem dimensões $p \times p$, e $U'U = V'V = I_p$, e D é uma matriz diagonal de dimensões $p \times p$ com elementos diagonais não-negativos μ_j , $j = 1, 2, \dots, p$. Os μ_j são denominados valores singulares de X , e $X = UDV'$ é denominada a decomposição em valores singulares de X .

BELSLEY, KUH e WELCH (1980) notam que a decomposição em valores singulares é estritamente relacionada com os conceitos de auto-valor e auto-vetor, desde que

$$X'X = (UDV')'(UDV') = VD^2V' = V\Lambda V',$$

de modo que o quadrado dos valores singulares de X são os auto-valores de $X'X$. V é a matriz dos auto-vetores normalizados de $X'X$ e U é uma matriz cujas colunas são os auto-vetores associados com os p auto-valores de XX' .

Aspectos computacionais do método da decomposição em valores singulares são descritos, por exemplo, em KENNEDY E GENTLE (1980), IEMMA (1988), CARVALHO FILHO e IEMMA (1988), dentre outros. Esses autores

ressaltam que a decomposição em valores singulares constitui-se num dos métodos numericamente mais estáveis na área da computação matricial. Uma prova concisa da existência de U, V e D para qualquer matriz X é dada por LAWSON e HANSON (1974).

MONTGOMERY e PECK (1981) salientam que o mal-condicionamento em X é refletido no tamanho dos valores singulares. Haverá um valor singular pequeno, para cada dependência linear aproximada, sendo que o grau do mal-condicionamento dependerá de quão pequeno é o valor singular, em relação ao valor singular $\mu_{\text{MÁX}}$.

O índice de condição da matriz X é definido em BELSLEY, KUH e WELCH (1980) e IEMMA (1988), dentre outros, como:

$$\eta_k = \frac{\mu_{\text{MÁX}}}{\mu_k}, \quad k = 1, 2, \dots, p.$$

acarretando assim, $\eta_k \geq 1$, para todo k.

Utilizando-se a decomposição em valores singulares, $X = UDV'$, a matriz de variâncias-covariâncias de $\hat{\beta}$ (o estimador de mínimos quadrados de β) pode ser escrito conforme MONTGOMERY e PECK (1981) como:

$$V(\hat{\beta}) = \sigma^2 (X'X)^{-1} = \sigma^2 V \Lambda^{-1} V'$$

ou, para o k-ésimo componente de $\hat{\beta}$,

$$V(\hat{\beta}_k) = \sigma^2 \sum_{j=1}^p \frac{v_{kj}^2}{\mu_j^2} = \sigma^2 \sum_{j=1}^p \frac{v_{kj}^2}{\lambda_j}.$$

Esses autores notam, assim como IEMMA (1987), que com exceção do valor σ^2 , $V(\hat{\beta}_k)$ corresponde ao k-ésimo fator de inflação da variância, ou seja

$$VIF_k = \sum_{j=1}^p \frac{v_{kj}^2}{\mu_j^2} = \sum_{j=1}^p \frac{v_{kj}^2}{\lambda_j}.$$

e conseqüentemente, a ocorrência de um ou mais valores singulares pequenos, ou auto-valores pequenos, poderá inflacionar severamente a variância de $\hat{\beta}_k$.

BELSLEY, KUH e WELCH (1980) sugerem a decomposição da variância em proporções, tais como:

$$\pi_{jk} = \frac{(v_{kj}^2 / \mu_j^2)}{VIF_k}, \quad j = 1, 2, \dots, p$$

como medida de multicolinearidade. Assim, se os π_{jk} são ordenados em uma matriz Π , de dimensões $p \times p$, como na Tabela 2.1, então os elementos de cada coluna de Π são exatamente as proporções da variância de cada $\hat{\beta}_k$ (ou cada fator de inflação da variância) associado ao i-ésimo valor singular (ou auto-valor).

Tabela 2.1

Proporções da Decomposição da Variância

Valor Singular Associado	$\text{Var}(\hat{\beta}_1)$	$\text{Var}(\hat{\beta}_2)$	. . .	$\text{Var}(\hat{\beta}_p)$
μ_1	π_{11}	π_{12}	. . .	π_{1p}
μ_2	π_{21}	π_{22}	. . .	π_{2p}
.	.	.	. . .	.
.	.	.	. . .	.
.	.	.	. . .	.
μ_p	π_{p1}	π_{p2}	. . .	π_{pp}

Então, se uma proporção alta da variância, para dois ou mais coeficientes da regressão, está associada com um valor singular pequeno, isto será uma indicação de multicolinearidade. Por exemplo, se π_{22} e π_{24} são grandes, o terceiro valor singular está associado com uma multicolinearidade que está inflacionando as variâncias de $\hat{\beta}_2$ e de $\hat{\beta}_4$.

BELSLEY, KUH e WELCH (1980), ressaltam que os π_{jk} utilizam da informação da decomposição em valores singulares, sobre dependências aproximadas, de uma forma que examina diretamente seus efeitos sobre as estimativas da regressão. Segundo esses autores, um meio apropriado para diagnosticar multicolinearidade significativa, consiste no seguinte:

Um índice de condição alto e que esteja associado com proporções altas da decomposição da variância, para duas ou

mais variâncias dos coeficientes de regressão estimados. Índices de condição maiores do que 30, e proporções da decomposição da variância maiores do que 0,5, são valores recomendados como indicativos de multicolinearidade.

2.1.2.5. Determinante de $X'X$

O determinante de $X'X$ pode ser usado como um índice da multicolinearidade, conforme apontam MONTGOMERY e PECK (1981). Desde que a matriz $X'X$ esteja na forma de correlação, a possível variação dos valores do determinante é $0 \leq | X'X | \leq 1$. Se $| X'X | = 1$, as variáveis independentes são ortogonais, enquanto que se $| X'X | = 0$, existe uma dependência linear exata entre as variáveis independentes. O grau de multicolinearidade torna-se mais severo à medida que $|X'X|$ aproxima-se de zero. Esses autores notam, que apesar desta medida de multicolinearidade ser de fácil aplicação, ela não fornece qualquer informação sobre a fonte da multicolinearidade.

Uma outra interpretação deste diagnóstico, sugerida por WILLAN e WATTS (1978), consiste em considerar a região de confiança conjunta para β , com coeficientes de confiança de $100(1-\alpha)\%$, baseada nos dados observados; isto é:

$$(\beta - \hat{\beta})' X'X (\beta - \hat{\beta}) \leq p \hat{\sigma}^2 F_{\alpha, p, n-p-1}$$

e, a correspondente região de confiança para β , baseada em um delineamento ortogonal, isto é:

$$(\beta - \hat{\beta})'(\beta - \hat{\beta}) \leq p \hat{\sigma}^2 F_{\alpha, p, n-p-1}$$

Tais autores salientam, que o delineamento ortogonal produz a menor região de confiança conjunta para β , considerando-se fixados, um tamanho de amostra, um valor de $\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 / n$, e um dado α . A razão dos volumes das duas regiões de confiança é $|X'X|^{1/2}$, de modo que esta quantidade mede a perda do poder de estimação devido a multicolinearidade. Escrito na forma, $100(|X'X|^{-1/2} - 1)$, este valor reflete o aumento da percentagem no volume da região de confiança conjunta para β , devido às dependências lineares aproximadas em X . Por exemplo, se $|X'X| = 0,25$, então o volume da região de confiança conjunta é $100[(0,25)^{-1/2} - 1] = 100\%$ maior do que se um delineamento ortogonal tivesse sido usado.

2.1.2.6. Outros diagnósticos

Outros diagnósticos ocasionalmente úteis e que evidenciam a presença da multicolinearidade, os quais são classificados como métodos informais em NETER, WASSERMAN e KUTNER (1983) são:

1. Grandes alterações nos coeficientes de regressão estimados, quando uma variável é incluída ou excluída, ou

quando uma observação é alterada ou retirada;

2. Resultados não-significantes em testes individuais sobre os coeficientes de regressão, para variáveis independentes importantes;

3. Coeficientes de regressão estimados, com sinal algébrico que é o oposto do esperado de considerações teóricas, ou experiência anterior;

4. Intervalos de confiança com grandes amplitudes, para coeficientes de regressão representando importantes variáveis independentes.

Esses autores notam no entanto, que os métodos informais tem limitações importantes, como por exemplo, não fornecem medidas quantitativas do impacto, nem identificam a natureza da multicolinearidade.

2.2. Alternativas ao Método de Mínimos Quadrados quando ocorre a Multicolinearidade

Uma vez constatado o problema da multicolinearidade, pode ser desejável a utilização de métodos alternativos ao de mínimos quadrados na estimação dos coeficientes de regressão. Tais métodos dão origem a estimadores viesados, mas, em geral, conforme GUNST e MASON (1977 b), têm melhor desempenho do que os estimadores de mínimos quadrados.

Dentre os métodos de estimação viesada,

destacam-se aqui a regressão sobre cristas e a regressão em componentes principais.

2.2.1. Regressão sobre cristas

HOERL e KENNARD (1970 a,b) propõem o uso da regressão sobre cristas (ridge regression), na qual as estimativas dos coeficientes de regressão são obtidas, resolvendo-se uma versão ligeiramente modificada das equações normais. Especificamente, o estimador sobre cristas β^* é definido como solução de

$$(X'X + kI_p) \beta^* = X'y$$

ou

$$\beta^* = (X'X + kI_p)^{-1} X'y$$

onde $k \geq 0$ é uma constante, em geral pertencente ao intervalo [0,1].

HOERL e KENNARD (1970 a) destacam que o estimador sobre cristas β^* , goza das seguintes propriedades:

1-) β^* é uma transformação linear de $\hat{\beta}$

$$\beta^* = (X'X + kI_p)^{-1} X'y = (X'X + kI_p)^{-1} (X'X) \hat{\beta} = Z_k \hat{\beta}$$

segundo-se imediatamente que β^* é um estimador viesado de

$$\beta, \text{ pois } E (\beta^*) = Z_k \beta$$

$$2-) V(\beta^*) = \sigma^2 (X'X + kI_p)^{-1} X'X (X'X + kI_p)^{-1}$$

3-) Denotando-se por EQM(β^*) o erro quadrático médio do estimador β^* tem-se:

$$\begin{aligned} \text{EQM}(\beta^*) &= \sigma^2 \text{Tr}[(X'X + kI_p)^{-1} X'X (X'X + kI_p)^{-1}] + \\ &\quad + k^2 \beta' (X'X + kI_p)^{-2} \beta \\ &= \sigma^2 \sum_{j=1}^p \frac{\lambda_j}{(\lambda_j + k)^2} + k^2 \beta' (X'X + kI_p)^{-2} \beta \\ &= V(\beta^*) + (\text{Viés } \beta^*)^2 \end{aligned}$$

Esta propriedade tem dois corolários importantes:

a-) A variância de β^* é uma função decrescente de k;

b-) O quadrado do viés é uma função crescente de k.

4-) Se $\beta'\beta$ é limitado, então existe $k>0$ tal que

$$\text{EQM}(\beta^*) < \text{EQM}(\hat{\beta}) = \sigma^2 \sum_{i=1}^p \lambda_i^{-1}$$

Uma representação gráfica, relacionando a variância, o quadrado do viés, e o erro quadrático médio do estimador sobre cristas apresentada por HOERL e KENNARD (1970 a), página 61 é dada na Figura 2.2.1 .

Figura 2.2.1.

Segundo MONTGOMERY e PECK (1981), as estimativas da regressão sobre cristas podem ser calculadas, utilizando-se um programa de computador de mínimos quadrados ordinários, aumentando-se os dados da seguinte forma:

$$X_A = \begin{bmatrix} X \\ \sqrt{k} I_P \end{bmatrix} \quad Y_A = \begin{bmatrix} y \\ O_P \end{bmatrix}$$

onde $\sqrt{k} I_P$ é uma matriz diagonal de dimensões $p \times p$, com elementos diagonais iguais à raiz quadrada de k e O_P é um vetor de zeros de dimensões $p \times 1$. As estimativas sobre

cristas serão então calculadas por:

$$\beta^* = (X_A' X_A)^{-1} X_A' y_A = (X' X + k I_p)^{-1} X' y$$

Uma interpretação geométrica da regressão sobre cristas apresentada por MARQUARDT e SNEE (1975), página 6, para o caso de um exemplo hipotético envolvendo apenas dois parâmetros β_1 e β_2 é dado na Figura 2.2.2.

Figura 2.2.2.

Conforme observam MARQUARDT e SNEE (1975), o ponto $\hat{\beta}$ no centro das elipses, corresponde à solução de mínimos quadrados, onde a soma de quadrados residual é mínima. A elipse pequena representa o lugar geométrico dos pontos no plano $\beta_1 \beta_2$ onde a soma de quadrados residual é constante para algum valor maior do que o mínimo. O círculo com centro na origem é tangente à menor elipse no ponto β^* . A estimativa sobre cristas β^* é o menor vetor, partindo da origem, que produz uma soma de quadrados residual, igual ao valor representado pela elipse pequena. Isto é, a

estimativa sobre cristas β^* , produz o vetor dos coeficientes da regressão com a menor norma, consistente com um acréscimo especificado na soma de quadrados residual.

Por outro lado, HOCKING (1976) ressalta que grande parte da controvérsia, acerca da regressão sobre cristas, reside na escolha do parâmetro k . HOERL e KENNARD (1970 a) apresentam um teorema que garante a existência de valores k , onde o erro quadrático médio de β^* é menor do que o erro quadrático médio de $\hat{\beta}$. Esses valores, no entanto, dependem dos parâmetros σ^2 e β , o que torna difícil sua obtenção na prática, conforme notam GUNST e MASON (1977 b).

OISHI (1983) salienta que vários métodos têm sido propostos, para a escolha de um valor para k , a maioria deles obtendo k como uma função dos estimadores $\hat{\sigma}^2$ e $\hat{\beta}$, onde $\hat{\sigma}^2$ é o estimador usual de σ^2 . Esse autor observa, no entanto, que infelizmente os k 's encontrados através desses métodos são variáveis aleatórias, o que implica que, a partir de obtido k_0 , não sabemos se de fato β^* possui a propriedade descrita no teorema da existência, isto é, se $EQM(\beta^*(k_0))$ é realmente menor, do que $EQM(\hat{\beta})$.

Em vista disso, vários trabalhos de comparações, utilizando simulação foram desenvolvidos, em busca de um método ótimo para a obtenção de k . Dentre esses, destacam-se aqui:

a) Método baseado no teorema da existência

HOERL, KENNARD e BALDWIN (1975), mostram que uma escolha apropriada para k , baseada no teorema da existência de HOERL e KENNARD (1970 a) é dada por:

$$k = \frac{p \hat{\sigma}^2}{\hat{\beta}' \hat{\beta}}$$

onde $\hat{\beta}$ e $\hat{\sigma}^2$ são obtidos da solução de mínimos quadrados, e p = posto $(X'X)$.

O principal problema desse critério, segundo OISHI (1983) está no fato de que $\hat{\beta}' \hat{\beta}$ costuma ser muito grande, em razão da multicolinearidade, o que torna k demasiadamente pequeno e em consequência β^* próximo de $\hat{\beta}$. Este fato pode ter originado o método seguinte.

b) Método baseado na estimação de $\beta' \beta$

Este método tem por base a igualdade (HOERL e KENNARD, 1970 a)

$$E(\hat{\beta}' \hat{\beta}) = \beta' \beta + \sigma^2 \sum_{i=1}^p \lambda_i^{-1}$$

na qual, $\hat{\beta}' \hat{\beta}$ superestima $\beta' \beta$, gerando problemas na estimação de k através do método anterior.

Em vista disso, McDONALD e GALARNEAU (1975),

introduziram o seguinte método:

$$\text{Seja } Q = \hat{\beta}' \hat{\beta} - \sigma^2 \sum_{i=1}^p \lambda_i^{-1}$$

isto é, Q é um estimador não-viesado de $\beta' \beta$. Então escolha k de modo que:

$$\beta^{**}, \beta^{**} = Q, \text{ se } Q > 0$$

caso contrário, faça $\beta^{**} = \hat{\beta}$, isto é, $k = 0$.

c) Método iterativo

Um procedimento iterativo de estimação, foi proposto por HOERL e KENNARD (1976). Especificamente, eles sugeriram a sequência de estimativas de β e de k ;

$$\begin{aligned} \hat{\beta} & & k_0 &= \frac{p \hat{\sigma}^2}{\hat{\beta}' \hat{\beta}} \\ \beta^{**}(k_0) & & k_1 &= \frac{p \hat{\sigma}^2}{\beta^{**}(k_0)' \beta^{**}(k_0)} \\ \beta^{**}(k_1) & & k_2 &= \frac{p \hat{\sigma}^2}{\beta^{**}(k_1)' \beta^{**}(k_1)} \\ & & & \vdots \\ & & & \vdots \end{aligned}$$

A alteração relativa em k_j , é usada para o término do

procedimento. Se

$$\frac{k_{j+1} - k_j}{k_j} > 20 \bar{\rho}^{-1.2}$$

o procedimento deve continuar; caso contrário termina, e usa-se $\beta^*(k_j)$ como o estimador sobre cristas, onde $\rho = \text{traço}(X'X)^{-1}$.

d) Método de Lawless e Wang

O estimador β^* pode também ser obtido através de um enfoque Bayesiano. Para tal, se a distribuição a "priori" de β é $N(0, \sigma_*^2 I)$, LINDLEY e SMITH (1972) demonstram então, que β^* é a média da distribuição a "posteriori",

onde

$$k = \frac{\sigma^2}{\sigma_*^2}$$

Tendo em vista este fato, LAWLESS e WANG (1976), propõem o seguinte critério para a escolha de k :

$$k = \frac{p \hat{\sigma}^2}{\sum_{i=1}^p \lambda_i \hat{\alpha}_i^2}$$

onde $\hat{\alpha} = V' \hat{\beta}$; $\lambda_1, \lambda_2, \dots, \lambda_p$ são os auto-valores de $X'X$ e V é a matriz dos auto-vetores normalizados correspondentes.

e) Método baseado na estatística C_k

MALLOWS (1973) propõe "plotar" a estatística C_k contra V_k , onde

$$C_k = \frac{SQRes(k)}{\hat{\sigma}^2} - n + 2 + 2 \text{Tr}(XL)$$

$$V_k = 1 + \text{Tr}(X'XLL')$$

$$L = (X'X + kI_p)^{-1}X'$$

e $SQRes(k)$ é a soma de quadrados residual como uma função de k . A sugestão é escolher k , de forma a minimizar C_k .

f) Método baseado no traço da crista

HOERL e KENNARD (1970 a) indicam que um valor apropriado de k , pode ser determinado pela análise do traço da crista, constituído de um gráfico dos elementos β_i^* , $i = 1, 2, \dots, p$ contra k , para valores de k em geral no intervalo $[0, 1]$.

Nesse sentido, MONTGOMERY e PECK (1981) apontam, que se a multicolinearidade é alta, a instabilidade nos coeficientes de regressão, torna-se evidente através do traço da crista. Nota-se então, que à medida em que k aumenta, algumas das estimativas sobre cristas variam acentuadamente, e que para algum valor de k , as estimativas β_i^* se estabilizarão. O objetivo será então, selecionar um valor de k razoavelmente pequeno, para o qual, as estimativas β_i^* sejam estáveis.

Segundo HOERL e KENNARD (1970 b), um valor adequado para k , pode ser obtido, observando-se o traço da crista e escolhendo-se k , onde:

___ Os coeficientes já se estabilizaram e terão o comportamento, semelhante ao de um sistema ortogonal;

___ Os coeficientes com sinais aparentemente incorretos em $k = 0$, serão mudados para o sinal apropriado;

___ Os resíduos não serão muito elevados, com relação ao resíduo obtido em $k = 0$.

Uma ilustração deste método, apresentada por DRAPER e SMITH (1981), página 310, para os dados conhecidos como dados de Hald (1952), é dada na Figura 2.2.3. Aqui β_i^* , $i = 1, 2, 3, 4$, representam os coeficientes da regressão sobre cristas em função dos diferentes valores de k .

Figura 2.2.3.

Outras ilustrações deste método para dados de alguns experimentos agrônômicos podem ser encontradas em

MATSUO (1986), onde o método da regressão sobre cristas é proposto como alternativa ao método da regressão linear múltipla, quando se evidencia um alto grau de multicolinearidade entre as variáveis independentes.

HOERL e KENNARD (1970 a) introduzem também uma extensão da regressão sobre cristas, que é conhecida como regressão sobre cristas generalizada. De acordo com esses autores, a discussão da regressão sobre cristas generalizada, torna-se um tanto mais simplificada, se os dados são transformados para o espaço de variáveis independentes ortogonais. Assim, se A é a matriz diagonal de dimensões $p \times p$, cujos elementos da diagonal principal são os auto-valores $\lambda_1, \lambda_2, \dots, \lambda_p$ de $X'X$, e se V é a correspondente matriz ortogonal de auto-vetores normalizados de $X'X$, então

$$V'X'XV = A$$

Fazendo-se

$$Z = XV$$

e

$$\alpha = V'\beta$$

o modelo linear torna-se

$$\begin{aligned} y &= X\beta + \varepsilon \\ &= (ZV')(V\alpha) + \varepsilon \\ &= Z\alpha + \varepsilon \end{aligned}$$

forma que é freqüentemente denominada de forma canônica do modelo.

O estimador de mínimos quadrados de α é a

solução para

$$(Z'Z)\hat{\alpha} = Z'y$$

que é equivalente a

$$\Lambda \hat{\alpha} = Z'y$$

ou

$$\hat{\alpha} = \Lambda^{-1} Z'y$$

resultando, que o vetor das estimativas dos parâmetros originais, é dado por

$$\hat{\beta} = V\hat{\alpha}$$

Em termos da forma canônica do modelo, o estimador sobre cristas generalizado é então, a solução para

$$(\Lambda + K) \hat{\alpha}_g = Z'y$$

onde $K = \text{diag}(k_1, k_2, \dots, k_p)$. Em relação ao modelo original, os coeficientes da regressão sobre cristas generalizada, são:

$$\hat{\beta}_g = V\hat{\alpha}_g$$

Quanto à escolha dos parâmetros, na matriz K , MONTGOMERY e PECK (1981) mostram que o erro quadrático médio, para a regressão sobre cristas generalizada é dado por:

$$\begin{aligned} \text{EQM}(\hat{\beta}_g) &= E [(\hat{\beta}_g - \beta)'(\hat{\beta}_g - \beta)] \\ &= E [(\hat{\alpha}_g - \alpha)'(\hat{\alpha}_g - \alpha)] \end{aligned}$$

$$= \sigma^2 \sum_{j=1}^P \frac{\lambda_j}{(\lambda_j + k_j)^2} + \sum_{j=1}^P \frac{\alpha_j^2 k_j^2}{(\lambda_j + k_j)^2}$$

onde o primeiro e o segundo termos, à direita desta igualdade são respectivamente, a soma das variâncias das estimativas dos parâmetros, e o quadrado do viés. Esses autores apontam então, que o $EQM(\hat{\beta}_0)$ é minimizado escolhendo-se

$$k_j = \frac{\sigma^2}{\alpha_j^2}, \quad j = 1, 2, \dots, p$$

os quais, no entanto, dependem de parâmetros desconhecidos, σ^2 e α_j .

Um procedimento iterativo, para a determinação dos k_j é sugerido por HOERL e KENNARD (1970 a). Este procedimento, utiliza-se da solução de mínimos quadrados para se obter uma estimativa inicial dos k_j , como:

$$k_j^0 = \frac{\hat{\alpha}_j^2}{\hat{\alpha}_j^2}, \quad j = 1, 2, \dots, p$$

Usa-se então essas estimativas iniciais dos k_j , para calcular a estimativa sobre cristas generalizada inicial, de

$$\hat{\alpha}_0^0 = (A + K^0)^{-1} Z'y$$

onde $K^0 = \text{diag}(k_1^0, k_2^0, \dots, k_p^0)$. A estimativa inicial, $\hat{\alpha}_0^0$

é então utilizada para corrigir a estimativa de k_j :

$$k_j^1 = \frac{\hat{\sigma}^2}{(\hat{\alpha}_{\alpha,j}^0)^2} \quad , j = 1, 2, \dots, p$$

Esses valores de k_j^1 poderão ser usados para corrigir as estimativas de α . O processo iterativo deve continuar, até estabilizar as estimativas dos parâmetros, sendo que uma medida de estabilidade, freqüentemente usada, é o valor de $||\hat{\alpha}_\alpha||^2 = \hat{\alpha}_\alpha' \hat{\alpha}_\alpha$. Especificamente, se $\hat{\alpha}_\alpha' \hat{\alpha}_\alpha$ permanece aproximadamente constante da iteração $i-1$ para a iteração i , o processo termina. Em caso contrário, continua.

MONTGOMERY e PECK (1981) ressaltam também, que na regressão sobre cristas generalizada, não existe um dispositivo gráfico dos coeficientes, equivalente ao traço da crista.

Por outro lado, GUILKEY e MURPHY (1975) desenvolvem uma forma alternativa da regressão sobre cristas que consiste em alterar somente os elementos correspondentes aos auto-valores, relativamente pequenos de $X'X$. Esta técnica é conhecida como regressão sobre cristas dirigida (directed ridge regression), e segundo GUILKEY e MURPHY (1975), as estimativas resultantes poderão ser menos viesadas do que em outros métodos de regressão sobre cristas, os quais alteram todos os elementos diagonais.

2.2.2. Regressão em componentes principais

HAWKINS (1973), GREENBERG (1975), dentre outros, propõem o uso da regressão em componentes principais, como procedimento alternativo ao de mínimos quadrados, para situações em que as variáveis independentes são altamente correlacionadas.

MANDEL (1982) por sua vez, sugere a aplicação da regressão em componentes principais, particularmente na forma da decomposição em valores singulares, salientando aspectos geométricos e algébricos da regressão linear múltipla, na presença de multicolinearidade.

Para a obtenção do estimador em componentes principais, HOCKING, SPEED e LYNN (1976) consideram o modelo na forma canônica,

$$y = Z\alpha + \varepsilon$$

onde as variáveis independentes no modelo são consideradas em ordem decrescente de seus auto-valores, isto é, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$, sendo que os "s" últimos auto-valores, $s < p$, são supostos aproximadamente iguais a zero. As colunas da matriz Z, correspondentes a esses auto-valores próximos de zero, são então excluídas da análise, e a regressão em componentes principais é obtida aplicando-se a técnica de mínimos quadrados aos componentes restantes, isto é,

$$\tilde{\alpha} = B\hat{\alpha}$$

onde B é uma matriz diagonal de dimensões $p \times p$, com elementos diagonais $b_1, b_2, \dots, b_p$, tais que $b_1 = b_2 = \dots = b_{p-s} = 1$ e $b_{p-s+1} = b_{p-s+2} = \dots = b_p = 0$

Assim o estimador em componentes principais, é

$$\tilde{\alpha} = (\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_{p-s}, 0, 0, \dots, 0)'$$

ou em termos das variáveis originais

$$\tilde{\beta} = V\tilde{\alpha} = \sum_{j=1}^{p-s} \lambda_j^{-1} v_j' X'y v_j$$

A regressão em componentes principais, pode ser vista, segundo JOHNSON, REIMER e ROTHROCK (1973), como o uso de um estimador de mínimos quadrados restrito, uma vez que $\tilde{\beta}$ é o estimador de mínimos quadrados restrito para o modelo $y = Z\alpha + \varepsilon$, sujeito a s restrições

$$V_{(s)}\beta = 0 \quad \text{onde} \quad V_{(s)} = (v_{p-s+1}, v_{p-s+2}, \dots, v_p)'$$

Por outro lado, FOMBY, HILL e JOHNSON (1978) mostram que a variância para o estimador em componentes principais, obtida retirando-se os s componentes, associados com os menores auto-valores, é pelo menos tão pequena quanto a variância de qualquer outro estimador de mínimos quadrados restrito, com um número menor ou igual de restrições.

Dentre as vantagens da regressão em

componentes principais, em relação a outros procedimentos de estimação viesada, GUNST e MASON (1977 b) apontam que tanto a teoria exata sobre as distribuições, como procedimentos de seleção de variáveis estão disponíveis.

3. MATERIAL E MÉTODOS

3.1 - Material

Em razão das dificuldades inerentes no trato das distribuições dos estimadores viesados, bem como de técnicas de seleção de variáveis, tornou-se inevitável utilizar-se de processos de simulação de dados, com o intuito de verificar a efetividade dos métodos considerados relativamente ao de mínimos quadrados, perante diferentes graus de multicolinearidade.

Por outro lado, a decisão em se utilizar dados simulados ao invés de dados reais, é feita de modo que a quase-singularidade das matrizes $X'X$, assim como valores dos parâmetros pudessem ser conhecidos.

Como pacote computacional, utilizou-se do Minitab, juntamente com o editor do Turbo Pascal para manuseio de arquivos, em microcomputadores Itautec I - 7000 PCxt II. No Apêndice 2 deste trabalho descreve-se um roteiro de operações, as quais foram efetuadas para cada amostra gerada, a fim de se obter as medidas de precisão consideradas neste estudo para os estimadores abordados.

Para o modelo linear adotado, $y = X\beta + \varepsilon$, onde

admite-se que $\varepsilon_i \sim \text{NID}(0, \sigma^2)$, σ^2 desconhecido, duas situações foram consideradas, segundo o número p de variáveis independentes envolvidas. Também foram analisados, dois valores para o número n de observações sobre cada variável independente. Assim, neste estudo de simulação, consideraram-se os seguintes valores de p e de n ;

$p = 3$ e 8 variáveis independentes e

$n = 15$ e 50 observações amostrais

A fim de estabelecer a estrutura de correlação desejada, entre as colunas da matriz X , adotou-se para cada um dos modelos as seguintes configurações:

a-) $p = 3$

As variáveis independentes são calculadas das relações,

$$X_{i1} = (1 - \alpha_1^2)^{1/2} Z_{i1} + \alpha_1 Z_{i4}, \quad i = 1, 2, \dots, n$$

$$X_{i2} = (1 - \alpha_2^2)^{1/2} Z_{i2} + \alpha_2 Z_{i4}, \quad i = 1, 2, \dots, n$$

$$X_{i3} = (1 - \alpha_3^2)^{1/2} Z_{i3} + \alpha_3 Z_{i4}, \quad i = 1, 2, \dots, n$$

onde Z_{i1} , Z_{i2} , Z_{i3} e Z_{i4} são variáveis pseudo-aleatórias $\text{N}(0,1)$ independentes e α_i , $i = 1, 2, 3$, são valores reais pertencentes ao intervalo $[0,1]$.

Designando-se então, por $\rho(X_{i1}, X_{i2})$, o coeficiente de correlação entre as variáveis X_{i1} e X_{i2} , resulta:

$$\rho(X_{i1}, X_{i2}) = \frac{\text{Cov}(X_{i1}, X_{i2})}{\sigma_{X_{i1}} \sigma_{X_{i2}}} = \frac{E(X_{i1} X_{i2}) - E(X_{i1}) E(X_{i2})}{\sigma_{X_{i1}} \sigma_{X_{i2}}}$$

$$\text{Sendo } X_{i1} X_{i2} = (1 - \alpha_1^2)^{1/2} (1 - \alpha_2^2)^{1/2} Z_{i1} Z_{i2} +$$

$$+ (1 - \alpha_1^2)^{1/2} \alpha_2 Z_{i1} Z_{i4} + (1 - \alpha_2^2)^{1/2} \alpha_1 Z_{i2} Z_{i4} + \alpha_1 \alpha_2 Z_{i4}^2$$

e em vista da independência entre as variáveis Z_{i1}, Z_{i2} e Z_{i4} , acarreta:

$$E(X_{i1} X_{i2}) = (1 - \alpha_1^2)^{1/2} (1 - \alpha_2^2)^{1/2} E(Z_{i1}) E(Z_{i2}) +$$

$$+ (1 - \alpha_1^2)^{1/2} \alpha_2 E(Z_{i1}) E(Z_{i4}) + (1 - \alpha_2^2)^{1/2} \alpha_1 E(Z_{i2}) E(Z_{i4}) +$$

$$+ \alpha_1 \alpha_2 E(Z_{i4}^2)$$

Como $E(Z_{ij}) = 0$ e $E(Z_{ij}^2) = 1$ tem-se:

$$E(X_{i1} X_{i2}) = (1 - \alpha_1^2)^{1/2} (1 - \alpha_2^2)^{1/2} 0.0 + (1 - \alpha_1^2)^{1/2} \alpha_2 0.0 +$$

$$+ (1 - \alpha_2^2)^{1/2} \alpha_1 0.0 + \alpha_1 \alpha_2 V(Z_{i4}^2)$$

$$= \alpha_1 \alpha_2 . 1$$

$$= \alpha_1 \alpha_2, \text{ que é o coeficiente de correlação}$$

$$\rho(X_{i1}, X_{i2}), \text{ pois } E(X_{i1}) = E(X_{i2}) = 0 \text{ e } \sigma_{i1} = \sigma_{i2} = 1$$

Analogamente, tem-se que:

$\alpha_1 \alpha_3$ é o coeficiente de correlação entre as variáveis X_{i1} e X_{i3} , $i = 1, 2, \dots, n$.

$\alpha_2 \alpha_3$ é o coeficiente de correlação entre as variáveis X_{i2} e X_{i3} , $i = 1, 2, \dots, n$.

Quatro combinações diferentes entre os valores α_1 , α_2 e α_3 , representadas pelas linhas da Tabela 3.1.1. são consideradas, a saber:

Tabela 3.1.1.

α_1	α_2	α_3	$\alpha_1 \alpha_2$	$\alpha_1 \alpha_3$	$\alpha_2 \alpha_3$
0,99	0,99	0,99	0,9801	0,9801	0,9801
0,30	0,99	0,99	0,2970	0,2970	0,9801
0,90	0,90	0,90	0,8100	0,8100	0,8100
0,50	0,70	0,95	0,3500	0,4750	0,6650

Para cada matriz X assim construída, tomar-se-ão três vetores dos coeficientes β ,

1- $\beta = v_3$: auto-vetor normalizado correspondente ao menor auto-valor de $X'X$.

2- $\beta = 0,5 (v_1 + v_3)$

3- $\beta = v_1$: auto-vetor normalizado correspondente ao maior auto-valor de $X'X$.

Observações sobre a variável dependente são

determinadas por:

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i, \quad i = 1, 2, \dots, n$$

onde $\beta_0 \equiv 0$ e X_{i1} , X_{i2} , e X_{i3} são as variáveis originais (não-padronizadas)

b-) $p = 8$

As variáveis independentes são calculadas das relações,

$$X_{ij} = (1 - \alpha_1^2)^{1/2} Z_{ij} + \alpha_1 Z_{i9}; \quad j = 1, 2, 3; \quad i = 1, 2, \dots, n$$

$$X_{ij} = (1 - \alpha_2^2)^{1/2} Z_{ij} + \alpha_2 Z_{i9}; \quad j = 4, 5, 6; \quad i = 1, 2, \dots, n$$

$$X_{ij} = (1 - \alpha_3^2)^{1/2} Z_{ij} + \alpha_3 Z_{i9}; \quad j = 7, 8; \quad i = 1, 2, \dots, n$$

onde Z_{i1} , Z_{i2} , ..., Z_{i9} são variáveis pseudo-aleatórias $N(0,1)$ independentes, e do mesmo modo que para o caso $p = 3$ verifica-se que:

α_1^2 é o coeficiente de correlação entre quaisquer pares de variáveis X_1 , X_2 e X_3

α_2^2 é o coeficiente de correlação entre quaisquer pares de variáveis X_4 , X_5 e X_6

α_3^2 é o coeficiente de correlação entre as variáveis X_7 e X_8

$\alpha_1 \alpha_2$ é o coeficiente de correlação entre variáveis $X_j (j = 1, 2, 3)$ e $X_{j'} (j' = 4, 5, 6)$

$\alpha_1 \alpha_3$ é o coeficiente de correlação entre variáveis $X_j (j = 1, 2, 3)$ e $X_{j''} (j'' = 7, 8)$

$\alpha_2 \alpha_3$ é o coeficiente de correlação entre variáveis $X_{j'} (j' = 4, 5, 6)$ e $X_{j''} (j'' = 7, 8)$

Consideraram-se também, neste caso, quatro combinações diferentes entre os diversos valores de α_1 , de α_2 , e de α_3 , conforme as linhas da Tabela 3.1.2.

Tabela 3.1.2.

α_1	α_2	α_3	$\alpha_1 \alpha_2$	$\alpha_1 \alpha_3$	$\alpha_2 \alpha_3$
0,99	0,99	0,99	0,9801	0,9801	0,9801
0,30	0,99	0,99	0,2970	0,2970	0,9801
0,90	0,90	0,90	0,8100	0,8100	0,8100
0,50	0,70	0,95	0,3500	0,4750	0,6650

Para cada matriz X , construída desta forma, tomaram-se três vetores dos coeficientes β ,

1-) $\beta = v_8$: auto-vetor normalizado correspondente ao menor auto-valor de $X'X$.

2-) $\beta = 0,5 (v_1 + v_2 + v_7 + v_8)$

3-) $\beta = v_1$: auto-vetor normalizado correspondente ao maior auto-valor de $X'X$.

Observações sobre a variável dependente são

determinadas por:

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_8 X_{i8} + \varepsilon_i, \quad i = 1, 2, \dots, n$$

onde $\beta_0 \equiv 0$ e $X_{i1}, X_{i2}, \dots, X_{i8}$, variáveis originais (não-padronizadas)

Tanto para o modelo com $p = 3$ como para o modelo com $p = 8$ variáveis independentes classificou-se a multicolinearidade, de uma maneira subjetiva, como forte, moderada ou fraca segundo a Tabela 3.1.3. dada a seguir.

Tabela 3.1.3

α_1	α_2	α_3	
0,99	0,99	0,99	forte
0,30	0,99	0,99	
0,90	0,90	0,90	moderada
0,50	0,70	0,95	
			fraca

Nesta tabela estão classificadas como multicolinearidade forte, tanto o caso em que $\alpha_1 = \alpha_2 = \alpha_3 = 0,99$, como também, quando $\alpha_1 = 0,30$ e $\alpha_2 = \alpha_3 = 0,99$. Deve-se ressaltar que o que difere uma situação da outra é o fato de que na primeira a multicolinearidade ocorre de uma maneira uniforme entre todas as colunas da matriz X , enquanto na segunda a multicolinearidade ocorre apenas em parte desta matriz, o que no entanto, não impede o mal-condicionamento da matriz $X'X$ e em consequência, o comprometimento do estimador de mínimos quadrados dos parâmetros do modelo.

Em razão de que $\beta' \beta = 1,0$, tem-se que, para as três direções consideradas do vetor de parâmetros β , resultará que as orientações de β e v_1 , em cada caso estudado serão dadas por:

$$\phi = v_1' \beta = 0,0; 0,5 \text{ e } 1,0 \text{ respectivamente}$$

As direções correspondentes aos auto-vetores normalizados, associados ao menor e ao maior auto-valores, foram escolhidas, pelo fato de que oferecem a vantagem mínima e máxima, respectivamente, de β^* sobre $\hat{\beta}$, conforme observa OISHI(1983), em resultado demonstrado por NEHOUSE e OMAN (1971). Além dessas duas achou-se por bem considerar também, uma terceira direção, ampliando assim o rol de situações estudadas.

Em cada um dos modelos considerados, quatro valores de σ foram analisados, a saber:

$\sigma = 0,01; 0,1; 0,5 \text{ e } 1,0$, ou equivalentemente, quatro razões-sinal-ruído (signal-to-noise ratio), definida como

$$\rho = \frac{\beta' \beta}{\sigma^2} : 10000 ; 100 ; 4 ; 1,0$$

Desta forma, considerou-se neste estudo 192 configurações diferentes, as quais são resultantes de 16 matrizes X (2 valores para p , 2 valores para n e 4 combinações entre α_1 , α_2 e α_3), 3 orientações de β com v_1 e 4 razões-sinal-ruído.

O processo de geração de amostras aleatórias de tamanho n (15, 50) foi repetido 30 vezes para cada

combinação dos valores de α_1 , α_2 e α_3 dados pelas Tabelas 3.1.1 e 3.1.2. Desta forma foram geradas um total de $16 \times 30 = 480$ matrizes X . Como para cada matriz X gerada, adotou-se três valores diferentes para ϕ e quatro valores diferentes para ρ na obtenção dos valores da variável dependente y , isto acarretou num totalde $480 \times 3 \times 4 = 5760$ vetores y , onde o valor $\beta_0 \equiv 0$. Assim produziu-se 30 conjuntos de estimativas dos parâmetros para cada combinação X , ϕ e ρ . Em cada amostra gerada, estimadores de β foram obtidos usando-se os métodos de mínimos quadrados, regressão sobre cristas e regressão em componentes principais.

3.2. - Métodos

3.2.1 - Modelo matemático

Considera-se o modelo de regressão linear múltipla,

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

onde,

y = vetor de observações da variável dependente, de dimensões $n \times 1$;

X = matriz dos valores das variáveis independentes, de dimensões $n \times p$, de posto $p < n$;

β = vetor de constantes desconhecidas, de dimensões $p \times 1$;

ε = vetor de erros aleatórios não-observáveis, de dimensões $n \times 1$, tal que $E(\varepsilon) = \phi$ e $E(\varepsilon\varepsilon') = \sigma^2 I_n$.

As variáveis independentes e a variável dependente são consideradas na forma padronizada, ou seja,

$$x_{ij} = \frac{X_{ij} - \bar{X}_j}{S_{jj}^{1/2}}, \quad i = 1, 2, \dots, n \quad \text{e} \quad j = 1, 2, \dots, p, \quad \text{são os}$$

elementos da matriz X , onde,

X_{ij} é a i -ésima observação da j -ésima variável independente; $\bar{X}_j$ é a média aritmética dos valores da j -ésima variável independente, e

$$S_{jj} = \sum_{i=1}^n (X_{ij} - \bar{X}_j)^2.$$

Do mesmo modo, $y_i = \frac{Y_i - \bar{Y}}{S_{yy}^{1/2}}$, $i = 1, 2, \dots, n$,

são os elementos do vetor y ; onde,

Y_i é a i -ésima observação da variável dependente, e

$$S_{yy} = \sum_{i=1}^n (Y_i - \bar{Y})^2.$$

Desta forma,

$$X'X = R = \begin{bmatrix} 1 & r_{12} & r_{13} \dots r_{1p} \\ r_{12} & 1 & r_{23} \dots r_{2p} \\ \vdots & \vdots & \vdots \dots \vdots \\ r_{1p} & r_{2p} & r_{3p} \dots 1 \end{bmatrix} \quad \text{e} \quad X'y = \begin{bmatrix} r_{1y} \\ r_{2y} \\ \vdots \\ r_{py} \end{bmatrix}$$

sendo,

$$r_{ij} = \frac{\sum_{k=1}^n (X_{ki} - \bar{X}_i)(X_{kj} - \bar{X}_j)}{(S_{ii}S_{jj})^{1/2}} = \frac{S_{ij}}{(S_{ii}S_{jj})^{1/2}}, \text{ o coeficiente}$$

de correlação entre as variáveis independentes X_i e X_j , e

$$r_{jy} = \frac{\sum_{k=1}^n (X_{kj} - \bar{X}_j)(Y_k - \bar{Y})}{(S_{jj}S_{yy})^{1/2}} = \frac{S_{jy}}{(S_{jj}S_{yy})^{1/2}}, \text{ o coeficiente}$$

de correlação entre a variável independente X_j e a variável dependente Y .

Para melhor interpretar as conseqüências da multicolinearidade sobre cada estimador do parâmetro β ,

estes serão expressos, como combinação linear dos auto-vetores de $X'X$.

Deste modo, se $\lambda_1, \lambda_2, \dots, \lambda_p$ são os auto-valores de $X'X$, com auto-vetores ortonormais correspondentes $v_1, v_2, \dots, v_p$ e se V é a matriz de dimensões $p \times p$, cuja i -ésima coluna é v_i , considera-se o teorema enunciado a seguir.

Teorema 3.2.1.1 [GRAYBILL, 1983]

Seja A uma matriz simétrica, de dimensões $n \times n$. Então existe uma matriz ortogonal P , tal que $P'AP = D$, onde D é uma matriz diagonal, com os auto-valores de A dispostos na diagonal.

Assim, conforme Teorema 3.2.1.1, obtém-se $V'X'XV = \Lambda$, onde $\Lambda = \text{diag} \{ \lambda_1, \lambda_2, \dots, \lambda_p \}$. Admite-se que as variáveis independentes são dispostas em ordem decrescente de seus auto-valores, ou seja, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$.

3.2.2. - Estimador de mínimos quadrados

Através do método dos mínimos quadrados, obtém-se o sistema de equações normais

$$X'X\hat{\beta} = X'y,$$

de onde resulta que,

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

é o estimador de mínimos quadrados para o parâmetro β , seguindo-se que,

$$E(\hat{\beta}) = \beta$$

Considerando-se que $V'X'XV = \Lambda$, tem-se que

$$(X'X)^{-1} = V\Lambda^{-1}V' = \sum_{i=1}^P \lambda_i^{-1} v_i v_i'$$

Conseqüentemente,

$$\hat{\beta} = \sum_{i=1}^P \lambda_i^{-1} v_i v_i' X' y = \sum_{i=1}^P \lambda_i^{-1} c_i v_i \quad (3.2.2.1.)$$

onde $c_i = v_i' X' y$.

Como admitiu-se que $\min(\lambda_i) = \lambda_p$, nota-se então, que para os casos em que X apresenta uma multicolinearidade forte apenas entre duas colunas, $\hat{\beta}$ tenderá a ser dominado por v_p . Coeficientes de variáveis não envolvidas na multicolinearidade, podem contudo, ser estimados com boa precisão, desde que os elementos correspondentes de v_p sejam pequenos, diluindo assim o efeito dominante de um λ_p^{-1} grande no último componente de $\hat{\beta}$. Para situações em que existem duas ou mais multicolinearidades os efeitos são similares. Os estimadores dos coeficientes das variáveis envolvidas nas multicolinearidades tendem a ser dominados pelos últimos componentes de (3.2.2.1.).

Variância e erro quadrático medio de $\hat{\beta}$

Sabe-se que a matriz de variância-covariância de $\hat{\beta}$ é dada por:

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

Utilizando-se a decomposição em auto-valores de $X'X$, tem-se

$$VC(\hat{\beta}) = \sigma^2 V\Lambda^{-1}V'$$

$$= \sigma^2 \sum_{i=1}^p \lambda_i^{-1} v_i v_i'$$

de onde implica que

$$VC(\hat{\beta}_j) = \sigma^2 \sum_{i=1}^p v_{ji}^2 \lambda_i^{-1}$$

Desse modo, a variância de $\hat{\beta}_j$ é uma combinação linear dos recíprocos dos auto-valores de $X'X$. Tal fato, associado ao trabalho de SILVEY (1969), segundo o qual na presença de multicolinearidades, tem-se um ou mais valores de λ_j muito pequenos, afetarão a precisão da estimativa de mínimos quadrados.

O erro quadrático médio de $\hat{\beta}$ é:

$$\begin{aligned} EQMC(\hat{\beta}) &= E \{ (\hat{\beta} - \beta)' (\hat{\beta} - \beta) \} \\ &= \sum_{j=1}^p E (\hat{\beta}_j - \beta_j)^2 \\ &= \text{Tr} \{ VC(\hat{\beta}) \} \\ &= \sigma^2 \text{Tr} \{ V\Lambda^{-1}V' \} \\ &= \sigma^2 \sum_{i=1}^p \lambda_i^{-1} \end{aligned}$$

Assim, apesar do estimador de mínimos quadrados de β , ser o estimador linear não-viesado de variância mínima, seu erro quadrático médio será grande, se multicolinearidades estiverem presentes entre as variáveis independentes.

Por outro lado, considerando-se um ponto de predição c , tem-se

$$\begin{aligned} V(c' \hat{\beta}) &= c' V(\hat{\beta}) c \\ &= \sigma^2 c' V A^{-1} V' c \\ &= \sigma^2 \sum_{i=1}^p \tilde{c}_i^2 \lambda_i^{-1}, \end{aligned}$$

onde $c'V = \tilde{c}' = (\tilde{c}_1, \tilde{c}_2, \dots, \tilde{c}_p)$. Então se o ponto de predição c , é tal que o coeficiente $\tilde{c}_p$ correspondente ao vetor v_p é grande, sua variância será grande. Conclui-se pois, que estimações relativamente imprecisas serão obtidas na direção dos auto-vetores que correspondem a multicolinearidades e que estimação relativamente precisa é possível, na direção dos demais auto-vetores.

3.2.3 - Estimador regressão sobre cristas

O estimador regressão sobre cristas é obtido resolvendo-se

$$(X'X + kI_p) \beta^* = X'y \quad (3.2.3.1.)$$

de onde tem-se

$$\beta^* = (X'X + kI_p)^{-1} X'y \quad (3.2.3.2.)$$

onde k é uma constante positiva, em geral pertencente ao intervalo $[0,1]$.

Tem-se também que,

$$\begin{aligned}
 (X'X + kI_p)v_i &= X'Xv_i + kv_i \\
 &= \lambda_i v_i + kv_i \\
 &= (\lambda_i + k)v_i \quad (3.2.3.3.)
 \end{aligned}$$

o que mostra que os auto-valores da matriz $(X'X + kI_p)$ são $\lambda_i + k$, $i = 1, 2, \dots, p$. Resulta então, que o número de condição do sistema (3.2.3.3.) é dado por

$$K(X) = \frac{\lambda_1 + k}{\lambda_p + k} < \frac{\lambda_1}{\lambda_p}$$

Em consequência, como $k > 0$, o número de condição do sistema (3.2.3.3.) é sensivelmente menor do que o do sistema de equações normais original, especialmente no caso de λ_p muito pequeno.

Observa-se ainda que o estimador sobre cristas é uma transformação linear do estimador de mínimos quadrados, pois

$$\begin{aligned}
 \beta^* &= (X'X + kI_p)^{-1} X'y \\
 &= (X'X + kI_p)^{-1} X'X \hat{\beta} \\
 &= [k(X'X)^{-1} + I_p]^{-1} \hat{\beta} \\
 &= Z_k \hat{\beta}
 \end{aligned}$$

Assim, desde que $EC \beta^* = EC Z_k' \hat{\beta} = Z_k' \beta$, tem-se que β^* é um estimador viesado de β .

Tendo em vista que $(\lambda_i + k)^{-1}$, $i = 1, 2, \dots, p$, são os auto-valores da matriz $(X'X + kI_p)^{-1}$, tem-se sua decomposição espectral, como:

$$(X'X + kI_p)^{-1} = \sum_{i=1}^p (\lambda_i + k)^{-1} v_i v_i'$$

Dessa forma, $\beta^* = (X'X + kI_p)^{-1} X'y$ pode ser escrito como:

$$\begin{aligned} \beta^* &= \sum_{i=1}^p (\lambda_i + k)^{-1} v_i v_i' X'y \\ &= \sum_{i=1}^p (\lambda_i + k)^{-1} c_i v_i \end{aligned} \quad (3.2.3.4.)$$

onde $c_i = v_i' X'y$.

Nota-se que se $X'X$ é aproximadamente singular com λ_p "pequeno", então o menor auto-valor de $(X'X + kI_p)$ será $\lambda_p + k$ e esta última matriz será não tão próxima da singularidade. Obviamente a regressão sobre cristas reduz essencialmente a contribuição sobre o estimador, dos auto-vetores mais associados com multicolinearidades.

Variância e erro quadrático médio de β^*

Considerando-se que $\beta^* = (X'X + kI_p)^{-1} X'y$,

tem-se

$$\begin{aligned}
VC(\beta^*) &= V[C'X'X + kI_p]^{-1}X'y \\
&= \sigma^2 C'X'X + kI_p]^{-1}X'XC'X'X + kI_p]^{-1} \\
&= \sigma^2 C'VAV' + kVV']^{-1}VAV'C'VAV' + kVV']^{-1} \\
&= \sigma^2 VC' \Lambda + kI_p]^{-1}V'VAV'V' C' \Lambda + kI_p]^{-1}V' \\
&= \sigma^2 VC' \Lambda + kI_p]^{-1} \Lambda C' \Lambda + kI_p]^{-1}V' \\
&= \sigma^2 V\Lambda_* V', \quad (3.2.3.5.)
\end{aligned}$$

onde Λ_* é uma matriz diagonal com $(\Lambda_*)_{ii} = \lambda_i (\lambda_i + k)^{-2}$ e portanto, $\text{Tr} \{ VC(\beta^*) \} = \sigma^2 \sum_{i=1}^p \lambda_i (\lambda_i + k)^{-2}$ é a variância total (soma das variâncias dos β_i^*) de β^* .

Em vista de que

$$V(\beta_i^*) = \frac{\sigma^2 \lambda_i}{(\lambda_i + k)^2}, \quad i = 1, 2, \dots, p, \quad \text{nota-se que}$$

esta é uma função decrescente de k .

Em relação ao erro quadrático médio de β^* , sabe-se que

$$EQM(\beta^*) = \text{Tr} \{ VC(\beta^*) \} + \{ E(\beta^*) - \beta \}' \{ E(\beta^*) - \beta \}$$

onde viu-se que

$$\text{Tr } \langle V(\beta^*) \rangle = \sigma^2 \sum_{i=1}^p \lambda_i (\lambda_i + k)^{-2}$$

Como $\beta^* = \hat{\beta}_k = [k(X'X)^{-1} + I_p]^{-1} \hat{\beta}$, tem-se

$$E(\beta^*) = [k(X'X)^{-1} + I_p]^{-1} \beta,$$

resultando que

$$\langle E(\beta^*) - \beta \rangle' \langle E(\beta^*) - \beta \rangle$$

$$= \langle [k(X'X)^{-1} + I_p]^{-1} \beta - \beta \rangle' \langle [k(X'X)^{-1} + I_p]^{-1} \beta - \beta \rangle$$

$$= \langle [k(VAV')^{-1} + I_p]^{-1} \beta - \beta \rangle' \langle [k(VAV')^{-1} + I_p]^{-1} \beta - \beta \rangle$$

$$= \beta' V [k\Lambda^{-1} + I_p]^{-1} [I_p - I_p] [k\Lambda^{-1} + I_p]^{-1} [I_p - I_p] V' \beta$$

$$= \beta' V [k\Lambda^{-1} + I_p]^{-1} [I_p] V' \beta$$

$$= \gamma' [k\Lambda^{-1} + I_p]^{-1} \gamma, \quad \text{onde } \gamma' = \beta' V$$

$$= \gamma' \Lambda_+ \gamma, \quad \text{onde } \Lambda_+ \text{ é uma matriz diagonal com}$$

$$(\Lambda_+)_{ii} = k^2 (\lambda_i + k)^{-2}$$

$$\text{Assim, } \langle E(\beta^*) - \beta \rangle' \langle E(\beta^*) - \beta \rangle = \sum_{i=1}^p \gamma_i^2 k^2 (\lambda_i + k)^{-2}$$

Consequentemente,

$$EQM(\beta^*) = \sigma^2 \sum_{i=1}^P \lambda_i (\lambda_i + k)^{-2} + \sum_{i=1}^P \gamma_i^2 k^2 (\lambda_i + k)^{-2} \quad (3.2.3.6.)$$

Considerando-se que

$$(E(\beta^*) - \beta)'(E(\beta^*) - \beta) = \sum_{i=1}^P \gamma_i^2 k^2 (\lambda_i + k)^{-2} = \sum_{i=1}^P \frac{\gamma_i^2}{\left(\frac{\lambda_i}{k} + 1\right)^2}$$

nota-se que $(E(\beta^*) - \beta)'(E(\beta^*) - \beta) = (\text{Viés } \beta^*)^2$ é uma função monótona crescente de k , e

$$\lim_{k \rightarrow \infty} (\text{viés } \beta^*)^2 = \sum_{i=1}^P \gamma_i^2 = \gamma' \gamma = \beta' V V \beta = \beta' \beta$$

Verifica-se portanto, que enquanto a variância decresce, o quadrado do viés cresce com k .

Ao usar a regressão sobre cristas, deseja-se escolher um valor de k , tal que a redução no termo variância, seja maior do que o acréscimo correspondente ao quadrado do viés. Isto feito, o erro quadrático médio do estimador sobre cristas β^* , será menor do que a variância do estimador de mínimos quadrados $\hat{\beta}$. Nos estudos de HOERL e KENNARD (1970 a) mostrou-se que existe um valor não-nulo de k para o qual o $EQM(\beta^*)$ é menor do que a variância do estimador de mínimos quadrados $\hat{\beta}$, desde que $\beta' \beta$ seja limitado.

Por outro lado, em relação à soma de quadrados residual de β^* tem-se

$$SQRes(\beta^*) = (y - X\beta^*)'(y - X\beta^*)$$

$$= y'y - \beta^{*'} X'y - y'X\beta^{*} + \beta^{*'} X'X\beta^{*} \quad (3.2.3.7.)$$

Em razão do sistema de equações normais para $\hat{\beta}$, $X'X\hat{\beta} = X'y$, tem-se:

$$\hat{\beta}'X'y = \hat{\beta}'X'X\hat{\beta} \text{ e } y'X\hat{\beta} = \hat{\beta}'X'X\hat{\beta}$$

Desta forma pode-se escrever (3.2.3.7.), como

$$\begin{aligned} \text{SQRes}(\beta^{*}) &= y'y - \hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta} - y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta} + \\ &\quad + \beta^{*'}X'X\beta^{*} - \beta^{*'}X'X\hat{\beta} - \hat{\beta}'X'X\beta^{*} \\ &= (y - X\hat{\beta})'(y - X\hat{\beta}) + (\beta^{*} - \hat{\beta})'X'X(\beta^{*} - \hat{\beta}) \\ &= \text{SQRes}(\hat{\beta}) + (\beta^{*} - \hat{\beta})'X'X(\beta^{*} - \hat{\beta}) \end{aligned}$$

Assim, vê-se que quando k aumenta, a soma de quadrados residual para β^{*} aumenta. Em consequência, pelo fato de que a soma de quadrados total é fixada, R^2 decresce quando k aumenta. Portanto, a estimativa sobre cristas necessariamente não estabelecerá o melhor "ajuste" para os dados, fato este que, segundo MONTGOMERY e PECK (1981), não é de grande relevância, desde que o interesse maior está em se obter um conjunto de estimativas estáveis para os parâmetros.

3.2.4. - Estimador em componentes principais

A obtenção do estimador em componentes principais é simplificada, quando os dados são transformados para o espaço de variáveis ortogonais. Para isso conforme teorema 3.2.1.1, considera-se a decomposição em auto-valores da matriz $X'X$, ou seja

$$V'X'XV = A$$

Fazendo-se então $Z = XV$ e $\alpha = V'\beta$, o modelo linear,

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

torna-se

$$\begin{aligned} y &= (ZV')(V\alpha) + \varepsilon \\ &= Z\alpha + \varepsilon \end{aligned} \quad (3.2.4.1.)$$

a qual denomina-se forma canônica do modelo. As colunas da matriz $Z = [Z_1, Z_2, \dots, Z_p]$, as quais definem um novo conjunto de variáveis ortogonais, visto que $Z'Z = V'X'XV = A$ são denominadas componentes principais.

Assim, o sistema de equações normais para o novo modelo (3.2.4.1.), é dado por

$$(Z'Z)\hat{\alpha} = Z'y$$

implicando que o estimador de mínimos quadrados de α , é

$$\hat{\alpha} = A^{-1}Z'y \quad (3.2.4.2.)$$

e a matriz de variâncias-covariâncias de $\hat{\alpha}$ é

$$\begin{aligned} V(\hat{\alpha}) &= \sigma^2 A^{-1}Z'ZA^{-1} \\ &= \sigma^2 A^{-1}A A^{-1} \\ &= \sigma^2 A^{-1} \end{aligned} \quad (3.2.4.3.)$$

Desta forma, um auto-valor pequeno de $X'X$, resulta que a variância do coeficiente da regressão ortogonal correspondente será grande. Além disso, como

$$Z'Z = \sum_{i=1}^p \sum_{j=1}^p Z_i Z_j' = \Lambda$$

costuma-se designar o auto-valor λ_j como a variância do j -ésimo componente principal.

Considerando-se que $\alpha = V'\beta$, tem-se que o vetor das estimativas de mínimos quadrados dos parâmetros originais β , é

$$\hat{\beta} = V \hat{\alpha} \quad (3.2.4.4.)$$

obtendo-se em consequência

$$VC(\hat{\beta}) = VC(V\hat{\alpha}) = V\Lambda^{-1}V'\sigma^2 \quad (3.2.4.5.)$$

e então,

$$VC(\hat{\beta}_j) = \sigma^2 \sum_{i=1}^p v_{ji}^2 \lambda_i^{-1}$$

conforme visto em (3.2.2.).

Admitindo-se que as variáveis independentes são dispostas em ordem decrescente de seus auto-valores, isto é, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$, o estimador em componentes principais de α é então obtido, suprimindo-se $(p-r)$ colunas da matriz Z , resultando digamos $Z_{(r)}$, e então usa-se mínimos quadrados para obter um estimador $\tilde{\alpha}$ para α , de

$$\tilde{\alpha} = (Z_{(r)}' Z_{(r)})^{-1} Z_{(r)}' y$$

$$= \Lambda_{(r)}^{-1} Z'_{(r)} y \quad (3.2.4.6.)$$

onde $\Lambda_{(r)} = \text{Diag} \{ \lambda_1, \lambda_2, \dots, \lambda_r \}$

O estimador em componentes principais de β é então

$$\begin{aligned} \tilde{\beta} &= V_{(r)} \tilde{\alpha} \\ &= V_{(r)} \Lambda_{(r)}^{-1} Z'_{(r)} y \end{aligned} \quad (3.2.4.7.)$$

onde $V_{(r)}$ é formada suprimindo-se as colunas apropriadas de V , de modo que $Z_{(r)} = X V_{(r)}$.

Média, variância e erro quadrático médio de $\tilde{\beta}$

a-) A média de $\tilde{\beta}$ é

$$E(\tilde{\beta}) = \beta - \sum_{i=r+1}^p \gamma_i v_i, \quad \text{onde } \gamma_i = v_i' \beta$$

Demonstração:

$$\text{Como } E(\hat{\beta}) = \beta = E\left(\sum_{i=1}^p \lambda_i^{-1} v_i v_i' X' y\right)$$

então,

$$\beta = E\left(\sum_{i=1}^r \lambda_i^{-1} v_i v_i' X' y\right) + E\left(\sum_{i=r+1}^p \lambda_i^{-1} v_i v_i' X' y\right)$$

Logo,

$$E\left(\sum_{i=1}^r \lambda_i^{-1} v_i v_i' X' y\right) = \beta - \sum_{i=r+1}^p \lambda_i^{-1} v_i v_i' X' E(y)$$

ou

$$\begin{aligned} E(\tilde{\beta}) &= \beta - \sum_{i=r+1}^p \lambda_i^{-1} v_i v_i' X' X \beta \\ &= \beta - \sum_{i=r+1}^p \lambda_i^{-1} v_i v_i' \lambda_i \beta \end{aligned}$$

$$\begin{aligned}
 &= \beta - \sum_{i=r+1}^p \lambda_i^{-1} \lambda_i v_i' \beta v_i \\
 &= \beta - \sum_{i=r+1}^p \gamma_i v_i, \quad \gamma_i = v_i' \beta
 \end{aligned}$$

o que mostra que $\tilde{\beta}$ é um estimador viesado para β , a menos que $v_i' \beta = 0$

b-) A variância total de $\tilde{\beta}$ (soma das variâncias dos $\tilde{\beta}_i$) é

$$V(\tilde{\beta}) = \sigma^2 \sum_{i=1}^r \lambda_i^{-1}$$

Demonstração:

Tendo-se que a matriz de variâncias-covariâncias de $\tilde{\alpha}$ é dada por

$$\begin{aligned}
 V(\tilde{\alpha}) &= \Lambda_{(r)}^{-1} Z_{(r)}' I \sigma^2 Z_{(r)} \Lambda_{(r)}^{-1} \\
 &= \sigma^2 \Lambda_{(r)}^{-1} Z_{(r)}' Z_{(r)} \Lambda_{(r)}^{-1} \\
 &= \sigma^2 \Lambda_{(r)}^{-1} \Lambda_{(r)} \Lambda_{(r)}^{-1} \\
 &= \sigma^2 \Lambda_{(r)}^{-1}
 \end{aligned}$$

e como $\tilde{\beta} = V_{(r)} \tilde{\alpha}$, tem-se que a variância total de $\tilde{\beta}$ é

$$\begin{aligned}
VC(\tilde{\beta}) &= \text{Tr} (\sigma^2 V_{(r)} \Lambda_{(r)}^{-1} V_{(r)}') \\
&= \text{Tr} (\sigma^2 V_{(r)}' V_{(r)} \Lambda_{(r)}^{-1}) \\
&= \text{Tr} (\sigma^2 \Lambda_{(r)}^{-1}) \\
&= \sigma^2 \sum_{i=1}^r \lambda_i^{-1}
\end{aligned}$$

e então obviamente, a maior redução na variância do estimador em componentes principais sobre mínimos quadrados é obtida, suprimindo-se os $(p - r)$ componentes principais correspondentes aos menores auto-valores de $X'X$.

c-) - O erro quadrático médio de $\tilde{\beta}$ é

$$EQM(\tilde{\beta}) = \sigma^2 \sum_{i=1}^r \lambda_i^{-1} + \sum_{i=r+1}^p \gamma_i^2$$

Demonstração:

Sabe-se que,

$$EQM(\tilde{\beta}) = VC(\tilde{\beta}) + (\text{viés } \tilde{\beta})^2$$

onde,

$$(\text{viés } \tilde{\beta})^2 = [E(\tilde{\beta}) - \beta]' [E(\tilde{\beta}) - \beta]$$

$$= [\beta - \sum_{i=r+1}^p \gamma_i v_i - \beta]' [\beta - \sum_{i=r+1}^p \gamma_i v_i - \beta]$$

$$= (- \sum_{i=r+1}^p \gamma_i v_i)' (- \sum_{i=r+1}^p \gamma_i v_i) = \sum_{i=r+1}^p \gamma_i^2$$

Como em b-), obteve-se que $V(\tilde{\beta}) = \sigma^2 \sum_{i=1}^r \lambda_i^{-1}$, tem-se

$$EQMC(\tilde{\beta}) = \sigma^2 \sum_{i=1}^r \lambda_i^{-1} + \sum_{i=r+1}^p \gamma_i^2$$

3.2.5. Escolha dos valores de k

O teorema seguinte confirma a existência de valores de k, para os quais, o erro quadrático médio do estimador β^* apresenta boas propriedades.

Teorema (3.2.5.1)

Se $\beta' \beta < \infty$, então existe $k > 0$ tal que

$$EQMC(\beta^*) < EQMC(\hat{\beta}) = \sigma^2 \sum_{i=1}^p \lambda_i^{-1}$$

Demonstração:

De (3.2.3.6) tem-se que o erro quadrático médio de β^* pode ser escrito como:

$$EQMC(\beta^*) = \sigma^2 \sum_{i=1}^p \frac{\lambda_i}{(\lambda_i + k)^2} + \sum_{i=1}^p \frac{\gamma_i^2}{\left(\frac{\lambda_i}{k} + 1\right)^2}$$

Derivando-se $EQMC(\beta^*)$ em relação ao parâmetro k, resulta:

$$\frac{d}{dk} EQMC(\beta^*) = -2 \sigma^2 \sum_{i=1}^p \frac{\lambda_i}{(\lambda_i + k)^3} + 2k \sum_{i=1}^p \frac{\gamma_i^2 \lambda_i}{(\lambda_i + k)^3}$$

Logo:

$$\lim_{k \rightarrow 0} \frac{d}{dk} EQM(\beta^*) = -2 \sigma^2 \sum_{i=1}^p \left(\frac{1}{\lambda_i}\right)^2 < 0$$

que indica uma inclinação negativa de $EQM(\beta^*)$ nas vizinhanças de $k = 0$. Logo, isto vai implicar na existência de valores $k > 0$ que tornam $EQM(\beta^*(k)) < EQM(\beta^*(0)) = EQM(\hat{\beta})$

Alguns desses valores de k são encontrados, resolvendo-se a inequação

$$\frac{d}{dk} EQM(\beta^*(k)) \leq 0$$

de onde obtém-se

$$k \leq \frac{p\sigma^2}{\beta' \beta} \quad \text{onde } p = \text{posto}(X'X)$$

Todavia, esses valores de k são delimitados por uma função dos parâmetros σ^2 e β , sendo portanto difícil a sua utilização na prática.

Um outro fato que também é evidenciado no teorema acima é que, quanto menor for λ_p , maior pode ser a vantagem obtida pela solução β^* em relação a $\hat{\beta}$. Essa vantagem, no entanto, somente poderá ser aproveitada a partir do momento em que se conhece o valor do parâmetro k . Contudo, esse valor de k , cujo ganho é máximo, é obtido fazendo-se

$$\frac{d}{dk} EQM(\beta^*) = 0, \text{ de onde obtém-se:}$$

$$k_{ot.} = \frac{p\sigma^2}{\beta'\beta}$$

isto é, o valor de k que minimiza $EQM(\beta^*)$ é função dos parâmetros σ^2 e β .

Este fato tem-se constituído no motivo mais forte para críticas ao método. O problema de determinar um valor para k , que satisfaça às condições ótimas acima tem movimentado pesquisadores da área e vários critérios foram propostos.

No estudo aqui desenvolvido sobre regressão sobre cristas, houve-se por bem considerar dois procedimentos distintos, para a obtenção do parâmetro k , os quais foram.

$$1-) k_1 = \frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}}$$

$$2-) k_2 = \frac{p\hat{\sigma}^2}{\sum_{i=1}^p \lambda_i \hat{\alpha}_i^2}, \quad \hat{\alpha}_i = v_i' \hat{\beta}$$

Tais valores foram adotados em razão de seu bom desempenho em trabalhos de simulação pesquisados, como em LAWLESS e WANG (1976); WICHERN e CHURCHILL (1978); GIBBONS (1981); OISHI (1983); HOERL, SCHUENEMEYER e HOERL

(1986), e também pelo fato de que estes procedimentos não requerem soluções iterativas.

3.2.6. Um critério para a seleção de variáveis

A maior dificuldade associada à aplicação da regressão em componentes principais refere-se certamente ao método a ser empregado para suprimir colunas da matriz Z .

Alguns autores, como por exemplo, GUNST e MASON (1977 b), MONTGOMERY e PECK (1981), baseiam-se somente na magnitude dos auto-valores da matriz $X'X$, como um critério para suprimir variáveis. Assim, os componentes principais correspondentes a auto-valores próximos de zero, os quais refletem o mal-condicionamento da matriz $X'X$ (o problema da multicolinearidade), são suprimidos da análise.

Contudo, uma crítica que se faz a este critério, refere-se ao fato de que neste caso as variáveis são suprimidas do modelo, sem se levar em consideração sua influência na predição de uma resposta estimada. Verifica-se também que diante de uma multicolinearidade forte, tal critério faz com que o modelo final apresente um número muito reduzido de variáveis, em relação ao modelo inicial.

Sob este aspecto, houve-se por bem considerar neste trabalho um método de seleção de variáveis, que

considere tanto a quase-singularidade de $X'X$, como a eficácia de predição de determinada variável independente sobre a resposta.

Assim, considerou-se a matriz aumentada descrita por WEBSTER, GUNST e MASON (1974), qual seja,

$$X^* = [Y, X]$$

onde Y e X são considerados na forma padronizada. Decorre daí que $X^{*'}X^*$ é a matriz de correlações aumentada de variáveis dependente e independentes, a qual tem auto-valores e auto-vetores definidos por,

$$|X^{*'}X^* - \delta_j I| = 0 \text{ e } (X^{*'}X^* - \delta_j I)\gamma_j = 0 ; j = 0, 1, \dots, p.$$

Os auto-vetores de $X^{*'}X^*$ definem um conjunto de eixos mutuamente ortogonais $Z_0, Z_1, \dots, Z_p$, sendo Z_j o eixo definido por γ_j , $j = 0, 1, \dots, p$. Assim, a direção do eixo Z_j relativa aos eixos originais é dada pela soma de vetores $\sum_{r=0}^p \gamma_{rj} e_r$, onde $e_0, e_1, \dots, e_p$ são vetores de comprimento unitário, da origem na direção dos eixos $Y, X_1, X_2, \dots, X_p$.

Desse modo, o primeiro elemento de γ_j representa o coseno do ângulo entre os eixos Y e Z_j , enquanto γ_{rj} ($r=1, 2, \dots, p$) representam os cosenos dos ângulos entre os eixos X_r e Z_j . Considerando-se que os auto-vetores são normalizados, γ_j é univocamente determinado a menos de um múltiplo de -1.

Por sua vez, o auto-valor correspondente a um

particular auto-vetor, mede a dispersão dos n pontos dados, na direção definida pelo auto-vetor. Em outras palavras, δ_j é a soma de quadrados das projeções dos n pontos dados, sobre o eixo Z_j . Assim, um valor pequeno de δ_j indicará que existe pouca variabilidade na direção Z_j ou seja, $z_{ij} = y_i \gamma_{0j} + \sum_{r=1}^p x_{ir} \gamma_{rj}$ é quase-nulo para $i = 1, 2, \dots, n$.

Como consequência, se γ_{0j} é próximo de zero, o eixo Z_j é aproximadamente ortogonal ao eixo Y . Portanto, se tanto δ_j como γ_{0j} são pequenos, o vetor γ_j revela o que se denomina de uma quase-singularidade não-preditiva; uma forte dependência linear, apenas entre as variáveis independentes, a qual produz pouca ou nenhuma mudança na variável dependente. Com base nesse fato, considerou-se como um critério para a seleção de variáveis, a magnitude dos auto-valores δ_j , e seus correspondentes valores γ_{0j} . Desse modo, tomou-se neste trabalho, conforme sugerido em WEBSTER, GUNST e MASON (1974), os valores $\delta_j \leq 0.05$ e $|\gamma_{0j}| < 0.10$, como referência para a seleção de variáveis, na regressão em componentes principais. Tais valores foram considerados segundo esses autores, em razão de sua eficiência em muitos conjuntos de dados analisados.

3.2.7. - Critérios de comparação

Como critérios de comparação, entre os métodos de estimação aqui abordados, considerou-se

1-) $L_1^2 = ||b - \beta||^2 = (b - \beta)'(b - \beta)$ para medir a precisão da estimativa, e

2-) $L_2^2 = ||Xb - X\beta||^2 = (b - \beta)'X'X(b - \beta)$ para medir a precisão da predição.

O valor de b em L_1^2 e L_2^2 se refere aos estimadores considerados, quais sejam, mínimos quadrados (M.Q.), regressão sobre cristas ($RC(k_1)$ e $RC(k_2)$), conforme o primeiro ou o segundo método utilizado para estimar k , regressão em componentes principais (C.P.), respectivamente.

Esses critérios de comparação foram também considerados em estudos de simulação como os de LAWLESS e WANG (1976) e HOERL, SCHUENEMEYER e HOERL (1986).

4. RESULTADOS E DISCUSSÃO

No Apêndice 1 deste trabalho, estão relacionadas 32 tabelas, resultantes do estudo de simulação realizado, a fim de se obterem os valores de L_1^2 e de L_2^2 segundo os critérios de estimação descritos em (3).

Estas tabelas (sendo 16 referentes a L_1^2 e 16 referentes a L_2^2), apresentam os valores médios de L_1^2 e de L_2^2 com seus respectivos erros padrões, obtidos com base em amostras, com 30 valores cada uma.

Os valores de L_1^2 e de L_2^2 que constam das referidas tabelas, são resultantes de $16 \times 30 = 480$ matrizes X . Tais matrizes foram obtidas, gerando-se inicialmente as amostras segundo a distribuição de probabilidade desejada, conforme descrito em (3.1), utilizando-se do gerador de números aleatórios do pacote computacional Minitab. Posteriormente, as colunas das matrizes X foram obtidas impondo-se as estruturas de correlações pre-determinadas, conforme as tabelas 3.1.1 e 3.1.2. Desse modo, como consideraram-se 3 valores para a direção ϕ , do vetor de parâmetros β com v_1 , o auto-vetor correspondente ao maior auto-valor de $X'X$, e 4 valores da razão-sinal-ruído

$\rho = \frac{\beta' \beta}{\sigma^2}$, isto nos levou a um total de $480 \times 3 \times 4 = 5760$ vetores da variável dependente y .

Assim, analisando-se as tabelas do Apêndice 1, notou-se que:

1-) O estimador de mínimos quadrados (conforme pode ser constatado nas tabelas de números 1 à 32), em geral, apresentou valores médios de L_1^2 e L_2^2 , maiores do que aqueles dos estimadores viesados considerados, o mesmo se verificando em relação a seus erros padrões. Exceções a esta afirmativa puderam, no entanto, ser notadas para algumas das situações consideradas, quais sejam:

a-) Nas configurações com valores altos da razão-sinal-ruído, principalmente $\rho = 10000$, e para as direções $\phi = 0.5$ e $\phi = 1.0$, e também na direção $\phi = 0.0$, para o valor de k_1 , o estimador regressão sobre cristas teve desempenho muito semelhante ao estimador de mínimos quadrados. Isto deveu-se ao fato de que para este valor de ρ , os valores de $\hat{\sigma}$ obtidos com base nos dados simulados, foram aproximadamente nulos, e por conseguinte, o mesmo ocorreu em geral, com os valores de k_1 e de k_2 , o que fez com que o estimador regressão sobre cristas, tendesse para o estimador de mínimos quadrados.

b-) Para configurações onde a multicolinearidade não representou um problema sério para o estimador de mínimos quadrados, visto que o número de condição da matriz $X'X$, em

geral resultou em valores inferiores a 100 (casos referidos como multicolinearidade moderada ou fraca na Tabela 3.1.3), e que estão relacionadas nas tabelas de números 5 à 8, 13 à 16, 21 à 24 e 29 à 32, notou-se que, para certos valores de ϕ e de ρ , tanto L_1^2 como L_2^2 alternaram-se em magnitudes, conforme o estimador adotado, e de acordo com os valores de p e de n considerados.

c-) Para as configurações onde os dados foram gerados, com base nos valores $\alpha_1 = 0.99$; $\alpha_2 = 0.99$; $\alpha_3 = 0.99$ (indicando uma multicolinearidade forte nos dados), nas direções $\phi = 0.5$ e $\phi = 1.0$, o estimador de mínimos quadrados, teve melhor desempenho do que o estimador em componentes principais, para valores da razão-sinal-ruído, ρ iguais a 100 e 10000 (conforme pode ser constatado nas tabelas, de números 1,2,9,10,17,18,25 e 26).

2-) O estimador regressão sobre cristas, produziu valores menores de L_1^2 e de L_2^2 do que o estimador de mínimos quadrados para a quase totalidade das situações analisadas, particularmente para os casos com multicolinearidades fortes, e razão-sinal-ruído ρ , com valores 1, 4 e 100 (conforme se observa nas tabelas, de números 1 à 4, 9 à 12, 17 à 20 e 25 à 28).

3-) Para os dados gerados segundo uma multicolinearidade designada como fraca na Tabela 3.1.3 e $p = 3$ (tabelas 7,8,

15 e 16) o estimador regressão sobre cristas teve desempenho bem parecido ao estimador de mínimos quadrados, particularmente na direção $\phi = 1.0$ e também nas direções $\phi = 0.0$ e $\phi = 0.5$ nos casos em que $\rho = 100$ e $\rho = 10000$. Para os valores $\rho = 1$ e $\rho = 4$ e direções $\phi = 0.0$ e $\phi = 0.5$, também não se observa uma supremacia de maneira uniforme de um desses estimadores sobre o outro, ocorrendo sim uma alternância entre as magnitudes, tanto de L_1^2 como de L_2^2 conforme o valor de n adotado. Para esta mesma situação de multicolinearidade, tomando agora o valor $\rho = .8$ (tabelas 23, 24, 31 e 32), observa-se uma certa supremacia do estimador regressão sobre cristas para a maioria dos valores de ρ e de ϕ .

4-) Para os dados obtidos, segundo uma multicolinearidade designada como moderada na Tabela 3.1.3, em particular para $n = 50$, verifica-se (conforme pode ser constatado nas tabelas 13, 14, 29 e 30) uma alternância entre os valores de L_1^2 , bem como de L_2^2 para os estimadores regressão sobre cristas e mínimos quadrados, conforme os valores de ρ e de ϕ considerados. Para $n = 15$ no entanto, notou-se um melhor desempenho do estimador regressão sobre cristas em quase todas as situações adotadas (como pode ser notado nas tabelas, de números 5, 6, 21 e 22).

5-) Em se tratando dos dois valores de k , considerados na regressão sobre cristas, obteve-se geralmente melhores

resultados, com o valor k_2 do que com k_1 , especialmente para os casos com multicolinearidades fortes e nas direções $\phi = 0.0$ e $\phi = 0.5$ (conforme pode ser verificado nas tabelas, de números 1 à 4, 9 à 12, 17 à 20 e 25 à 28). Por outro lado, observaram-se que estes valores de k_1 e de k_2 muito raramente não pertenceram ao intervalo $[0,1]$ e que os mesmos ora alternaram-se em magnitudes, ora forneceram valores aproximadamente iguais, conforme a configuração considerada.

6-) Comparando-se o desempenho do estimador em componentes principais com o estimador de mínimos quadrados podemos ressaltar:

a-) Nas situações onde a multicolinearidade não se caracterizou como problema sério, para as estimativas de mínimos quadrados (multicolinearidade fraca ou moderada), com exceção da configuração com $p = 8$ e $n = 15$, e $\alpha_1 = 0.90$; $\alpha_2 = 0.90$ e $\alpha_3 = 0.90$ estes estimadores tiveram desempenho coincidentes (conforme pode ser constatado nas tabelas de números 5 à 8, 13 à 16, 23 à 24 e 29 à 32), porque de acordo com o critério adotado para a eliminação de variáveis independentes, nenhuma variável foi suprimida do modelo.

b-) Para os casos com multicolinearidades fortes (tabelas, de números 1 à 4, 9 à 12, 17 à 20 e 25 à 28), o estimador em componentes principais teve desempenho bem

melhor do que o estimador de mínimos quadrados. Nestes casos as diferenças mais acentuadas, especialmente nos valores de L_1^2 , foram em geral observadas nas direções $\phi = 0$, e $\phi = 0,5$. Exceções a esta afirmativa porém ocorreram nas configurações comentadas em 1-c, assim como para a configuração em que $p = 3$, $n = 50$ e $\alpha_1 = 0.30$, $\alpha_2 = 0.99$ e $\alpha_3 = 0.99$ (tabelas 11 e 12), nas direções $\phi = 0.0$, e também $\phi = 0.5$ para $\rho = 4$, 100 e 10000, onde o estimador de mínimos quadrados foi superior ou equivalente ao estimador em componentes principais.

7-) Em relação ao desempenho do estimador em componentes principais comparativamente ao estimador regressão sobre cristas, podemos observar:

a-) Para dados com multicolinearidades designadas como fortes na Tabela 3.1.3 e que estão relacionadas nas tabelas de números 1 à 4, 9 à 12, 17 à 20 e 25 à 28, com exceção da direção $\phi = 0.0$ e da razão-sinal-ruído $\rho = 10000$, o estimador em componentes principais, raramente foi superior ao estimador sobre cristas para ambos os casos estudados, quais sejam, RCC k_1 e RCC k_2 . Observou-se contudo, que em boa parte das configurações, possivelmente a maioria delas, particularmente para a direção $\phi = 0.0$ e também para $\phi = 0.5$, com exceção dos valores $\rho = 100$ e $\rho = 10000$, já comentados em 1-c, o estimador em componentes principais teve valores de L_1^2 e de L_2^2 situados entre os dois valores

obtidos através da regressão sobre cristas.

b-) Para os casos onde os dados apresentaram uma multicolinearidade fraca ou moderada, casos estes onde o estimador em componentes principais, em geral coincidiu com o estimador de mínimos quadrados, notou-se com exceção do valor $\rho = 10000$, onde o estimador em componentes principais e o estimador regressão sobre cristas em geral coincidiram, que tanto os valores de L_1^2 como os de L_2^2 , se alternaram em magnitudes, dependendo da direção ϕ , e da razão-sinal-ruído ρ , consideradas, isto é, não houve uma supremacia predominante de qualquer deles sobre o outro (conforme pode ser verificado nas tabelas, de números de 5 à 8, 13 à 16, 21 à 24 e 29 à 32).

8-) Com relação aos dois valores de p e aos dois valores de n , considerados nos modelos estudados, notou-se em especial que:

a-) Quando os dados foram simulados, adotando-se primeiramente $p = 3$ variáveis independentes (tabelas de 1 à 16), e posteriormente $p = 8$ variáveis independentes (tabelas de 17 à 32), mantendo-se a mesma estrutura de correlações entre as colunas da matriz X , o número de condição da matriz $X'X$ teve um aumento considerável, especialmente para os casos com multicolinearidades mais fortes.

b-)em se tratando dos dois valores de n considerados, quais

sejam, $n = 15$ e $n = 50$ observações sobre cada variável independente, notou-se que, em geral tanto os valores médios de L_1^2 como os de L_2^2 , com os respectivos erros padrões, para uma mesma configuração, tiveram um decréscimo acentuado para um maior número de observações. Quanto ao número de condição das matrizes $X'X$, no entanto, não se observaram alterações significativas, isto é, não houve um agravamento da multicolinearidade, com um maior número de observações.

9-) Quanto aos graus de multicolinearidade (forte, moderada e fraca) que foram considerados para a obtenção das matrizes X , deve-se ressaltar que apenas aquele classificado como forte é o que de fato possui as características desejáveis para a aplicação dos estimadores viesados. Todavia, estudaram-se também os casos classificados como multicolinearidades moderada e fraca com o intuito de verificar o desempenho dos estimadores viesados em condições que são favoráveis ao estimador de mínimos quadrados, em especial a multicolinearidade fraca.

5. CONCLUSÕES

Considerando-se os objetivos deste estudo, julgou-se desejável que as conclusões se dessem em termos dos resultados aqui obtidos, em face da simulação realizada.

Sob esse aspecto, de modo geral conclui-se:

5.1-) Em relação ao estimador regressão sobre cristas.

Os resultados indicaram que o estimador regressão sobre cristas, pode ser recomendado como uma alternativa ao estimador de mínimos quadrados na obtenção das estimativas dos parâmetros de um modelo de regressão, nos casos onde há uma multicolinearidade forte entre as colunas da matriz X . Nestas situações, este estimador mostrou-se ser mais estável em relação às medidas de precisão adotadas, quais sejam, L_1^2 e L_2^2 .

Em se tratando dos dois métodos empregados, na obtenção do parâmetro k , apesar de não ter ocorrido uma melhora uniforme, de k_2 em relação a k_1 , para todos os valores de ϕ e de ρ considerados, notou-se um melhor desempenho de k_2 para a maioria das situações abordadas,

especialmente para os dados com uma multicolinearidade forte.

5.2-) Em relação ao estimador em componentes principais.

Também aqui, pode-se notar, que o estimador em componentes principais é uma boa alternativa ao estimador de mínimos quadrados, diante de uma multicolinearidade classificada como forte; especialmente quando a multicolinearidade não se dá de maneira uniforme para todas as colunas da matriz X ; fato este que pode conduzir a um modelo final com um número reduzido de variáveis independentes.

O critério para a seleção de variáveis utilizado para a obtenção do estimador em componentes principais, mostrou-se ser de boa eficiência, apesar de que para alguns valores de ϕ e de ρ , o número de variáveis independentes que permaneceram no modelo, conforme citado acima, foi pequeno diante de uma multicolinearidade forte entre todas as colunas de X .

Como sugestão, para situações envolvendo um número muito elevado de variáveis, e que eventualmente se deve excluir algumas delas para facilidade de análise, mesmo que estas não sejam fortemente correlacionadas, acredita-se ser ainda este método aconselhável, aumentando-se os limites aqui considerados, para além de $\delta_j \leq 0,05$ e $|\gamma_{oj}| < 0,10$.

5.3-) Quanto a escolha do estimador a ser utilizado

Em se tratando de optar por um dos estimadores estudados podemos destacar:

a-) Nos casos em que não se esteja diante de uma multicolinearidade tida como forte, o que em geral conduz à um número de condição de $X'X$ inferior a 100, o estimador de mínimos quadrados em razão de suas propriedades ótimas, deve sem dúvida, ser o escolhido.

b-) Nos casos onde a multicolinearidade caracteriza-se como um problema grave para o estimador de mínimos quadrados, um método de estimação viesada deve ser preferido.

c-) Para as situações onde o número de variáveis independentes não é grande, o método da regressão sobre cristas mostra-se mais adequado. Neste caso aconselha-se antes da aplicação do método determinar-se o valor de k através de alguns critérios diferentes, bem como pelo método gráfico do traço da crista, descrito por HOERL e KENNARD (1970 a), a fim de se comparar a proximidade desses valores.

d-) Para as situações em que deparamos com um número muito elevado de variáveis independentes o estimador em componentes principais deve ser o escolhido, especialmente quando a multicolinearidade não se dá entre todas as colunas da matriz X , fato este, que em geral, faz com que os critérios para a seleção de variáveis tornem-se menos

eficientes.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- BELSLEY, D.A; KUH, E. ; WELCH, R.E. Regression diagnostics: Identifying influential data and sources off collinearity. New York, John Wiley & Sons, 1980, 292 p.
- BERK, K. N. Validating regression procedures with new data. Technometrics, Alexandria, 26(4):331-8, 1984.
- CARVALHO FILHO, A.A. A decomposição da matriz de planejamento por valores singulares em modelos de posto incompleto. Piracicaba, 1987. 116p. (Doutorado - Escola Superior de Agricultura "Luiz de Queiroz" / USP)
- CARVALHO FILHO, A.A. & IEMMA, A.F. Aplicação da decomposição por valores singulares no modelo linear de Gauss-Markov. Atas do 8º SINAPE, pag. 86, 1988.
- CHATTERJEE, S. & PRICE, B. Regression analysis by example. New York, John Wiley & Sons, 1977. 228 p.

COPAS, J.B. Regression, prediction and shrinkage.

Journal of the Royal Statistical Society. Série B,

London, 45(3):311-35, 1983.

DEMPSTER, A.P. ; SCHATZOFF, M. ; WERMUTH, N. A simulation

study of alternatives to ordinary least squares.

Journal of the American Statistical Association,

Washington, 72(357):77-106, 1977.

DRAPER, N.R. & SMITH, H. Applied regression analysis, 2

ed., New York, John Wiley & Sons, 1981., 710 p.

DWIVEDI, T.D. ; SRIVASTAVA, V.K. ; HALL, R.L. Finite sample

properties of ridge estimators. Technometrics, Alexandria,

22(2):205-12, 1980.

FAREBROTHER, R.W. The minimum mean square error linear

estimator and ridge regression. Technometrics, Alexandria,

17(1):127-8, 1975.

FOMBY, T.B. ; HILL, R.C. ; JOHNSON, S.R. An optimal property

of principal components in the context of restricted least

squares. Journal of the American Statistical Association,

Washington, 73(361):191-3, 1978.

FOX, J. Linear statistical models and related methods with applications to social research. New York, John Wiley & Sons, 1984. 449 p.

GIBBONS, D.G. A simulation study of some ridge estimators. Journal of the American Statistical Association, Washington, 76(373):131-9, 1981.

GOLDSTEIN, M & SMITH, A.F.M. Ridge-type estimators for regression analysis. Journal of the Royal Statistical Society, Série B, London, 36(2):284-91, 1974.

GRAYBILL, F.A. Introduction to matrices with application in statistics. Belmont, Wadsworth Publishing Company, 1983. 372 p.

GREENBERG, E. Minimum variance properties of principal components regression. Journal of the American Statistical Association, Washington, 70(349):194-7, 1975.

GUILKEY, D.K. & MURPHY, J.L. Directed ridge regression techniques in cases of multicollinearity. Journal of the American Statistical Association, Washington, 70(352):769-75, 1975.

GUNST, R.F. & MASON, R.L. Advantages of examining multicollinearities in regression analysis. Biometrics, Raleigh, 33:249-60, 1977a.

GUNST, R.F. & MASON, R.L. Biased estimation in regression: An evaluation using mean squared error. Journal of the American Statistical Association, Washington, 72(359):616-28, 1977b.

HAWKINS, D.M. On the investigation of alternative regressions by principal component analysis. Applied Statistics, London 22:275-286, 1973.

HAWKINS, D.M. Relations between ridge regression and eigenanalysis of the augmented correlation matrix. Technometrics, Alexandria, 17(4):477-80, 1975.

HILL, R.C. ; FOMBY, T.B.; JOHNSON, S.R. Component selection norms for principal components regression. Communication of Statistics - Theory and Methods, Dallas, A6, (4):309-34, 1977.

HOCKING, R.R. The analysis and selection of variables in linear regression. Biometrics, Raleigh, 32(1):1-49, 1976.

HOCKING, R.R.; SPEED, F.M. ; LYNN, M.J. A class of biased estimators in linear regression. Technometrics, 18(4):425-37, 1976.

HOERL, A.E. & KENNARD, R.W. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, Alexandria, 12(1):55-68, 1970a

HOERL, A.E. & KENNARD, R.W. Ridge regression: Applications to nonorthogonal problems. Technometrics, Alexandria, 12(1):69-82, 1970b

HOERL, A.E.; KENNARD, R.W. Ridge regression: Iterative estimation of the biasing parameters. Communications in Statistics, New York, A5:77-8, 1976

HOERL, A.E.; KENNARD, R.W ; BALDWIN, K.F. Ridge regression: Some simulations. Communications in Statistics, New York, A4:105-23, 1975.

HOERL, R.W. ; SCHUENEMEYER, J.H. ; HOERL, A.E. A simulation of biased estimation and subset selection regression techniques. Technometrics, Alexandria, 28(4):369-80, 1986.

IEMMA, A.F. Modelos lineares. Uma introdução para

profissionais da pesquisa agropecuária. 2^o Simpósio de Estatística Aplicada à Experimentação Agronômica e 32^a Reunião da Sociedade Brasileira de Biometria. Londrina, 1987, 263 p.

IEMMA, A.F. Matrizes para estatística. Um texto para profissionais de ciências aplicadas. Piracicaba, Escola Superior de Agricultura "Luiz de Queiróz", USP, 1988. 339 p.

IEMMA, A.F. & CARVALHO FILHO, A.A. A decomposição por valores singulares da matriz do delineamento em modelos de classificação simples. Atas do 3^o Simpósio de Estatística Aplicada à Experimentação Agronômica e 34^a Reunião da Sociedade Brasileira de Biometria, 1989. pág. 36.

JOHNSON, S.R.; REIMER, S.C.; ROTHROCK, T.P. Principal Components and the problem of multicollinearity. Metroeconomica, Bologna, 25:306-17. 1973.

KENDALL, M.G. A course in multivariate analysis. London, Griffin, 1957.

KENNEDY JR., W.J. & GENTLE, J.E. Statistical Computing, New York, Marcel Dekker, 1980. 589 p.

KUPPER, L.L. & MEYDRECH, E.F. A new approach to mean squared error estimation of response surfaces.

Biometrika, London, 60(3):373-79, 1973.

LAWLESS, J.F. & WANG, P. A simulation study of ridge and other regression estimators. Communication of Statistics -Theory and Methods, Dallas, A5(4):307-23, 1976.

LAWSON, C.L. & HANSON, R.J. Solving Least Square Problems, Englewood Cliffs, Prentice Hall, 1974. 340p.

LINDLEY, D.V. & SMITH, A.F.M. Bayes estimates for the linear model. Journal of the Royal Statistical Society, Série B, London, 34(1):1-41, 1972.

MALLOWS, CL. Some comments on Cp. Technometrics, Alexandria, 15:661-75, 1973.

MANDEL, J. Use of the singular value decomposition in regression analysis. The American Statistician, Alexandria, 36(1):15-24, 1982.

MARQUARDT, D.W. Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. Technometrics, Alexandria, 12(3):391-612, 1970.

MARQUARDT, D.W. & SNEE, R.D. Ridge regression in practice.

The American Statistician, Alexandria, 29:3-20, 1975.

MASON, R.L.; GUNST, R.F.; WEBSTER, J.T. Regression Analysis
and problems of multicollinearity.

Communications in Statistics, New York, A3:277-92, 1975.

MATSUO, T. O uso da regressão de cumeieira em experimentos
agrônômicos. Piracicaba, 1986. 89 p. (Mestrado - Escola
Superior de Agricultura "Luiz de Queiroz"/USP).

MCDONALD, G.C. & GALARNEAU, D.I. A monte carlo evaluation
of some ridge type estimators. Journal of the American
Statistical Association, Washington, 70(350):407-16, 1975.

MONTGOMERY, D.C. & PECK, E.A. Introduction to linear
regression analysis, New York, John Wiley & Sons, 1981.

504 p.

NETER, J. ; WASSERMAN, W. ; KUTNER, M.H. Applied linear
regression models. Homewood, Richard D. Irwin, 1983. 547 p.

NEWHOUSE, J.P. & OMAN, S.D. An evaluation of ridge
estimators. Technical Report, Santa Monica, R-716-PR,
1971.

OISHI, J. - Regressão sobre cristas. São Paulo, 1983. 100 p.
(Mestrado - Instituto de Matemática e Estatística / USP).

OMAN, S.D. A confidence bound approach to choosing the
biasing parameter in ridge regression. Journal of the
American Statistical Association, Washington,
76(374):452-61, 1981.

PARK, S.H. Collinearity and optimal restrictions on
regression parameters for estimating response
Technometrics, Alexandria, 23(3):289-95, 1981.

SILVEY, S.D. Multicollinearity and imprecise estimation.
Journal of the Royal Statistical Society, Série B, London,
31(3):539-52, 1969.

SWINDEL, B.F. Geometry of ridge regression illustrated:
The American Statistician, Alexandria, 35(1):12-15, 1981.

THEOBALD, C.M. Generalizations of mean square error. Applied
to ridge regression. Journal of the Royal Society,
Série B, London, 36:103-6, 1974.

WEBSTER, J.T.; GUNST, R.F. ; MASON, R.L. Latent root
regression analysis. Technometrics, Alexandria,

16(4):513-522, 1974.

WETHERILL, G.B. ; DUNCOMBE, P. ; KENWARD, M.; KÖLLERSTROM,
J.; PAULS, S.R.; VOWDEN, B.J. Regression analysis
with applications, London, Chapman and Hall, 1985. 311 p.

WICHERN, D.W. & CHURCHILL, G.A. A comparison of ridge
estimators. Technometrics, Alexandria, 20(3): 301-11,
1978.

WILLAN, A.R. & WATTS, D.G. Meaningful multicollinearity
measures. Technometrics, Alexandria, 20(4):407-12, 1978.

7. APÊNDICE 1

Tabela 1 - Valores médios e erros padrões correspondentes (entre parênteses) de L_1^2 para o modelo com $p = 3$ e $n = 15$ e $\alpha_1 = 0,99$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	8,0371 (15,3549)	10,1775 (8,7082)	27,5886 (15,8011)	49,7817 (28,3208)
RCCK ₁	3,8927 (2,9697)	3,7738 (2,8304)	19,2932 (8,5149)	49,4474 (27,8841)
RCCK ₂	1,0609 (0,1888)	0,9268 (0,1792)	0,4987 (0,3680)	31,9128 (9,9857)
CP	1,0310 (0,2061)	1,0632 (0,5354)	14,4553 (8,1813)	30,3805 (10,7161)
<u>b. $\phi = 0,5$</u>				
MQ	4,4832 (6,9842)	2,1768 (2,0335)	0,0957 (0,0541)	0,0166 (0,0069)
RCCK ₁	1,8223 (1,1174)	0,8193 (0,4208)	0,0850 (0,0605)	0,0161 (0,0066)
RCCK ₂	0,3556 (0,2314)	0,3528 (0,2161)	0,0799 (0,0571)	0,0162 (0,0066)
CP	0,2934 (0,0521)	1,2871 (0,4869)	0,1167 (0,0733)	0,1383 (0,0289)
<u>c. $\phi = 1,0$</u>				
MQ	2,3822 (3,2950)	0,8453 (0,5672)	0,1995 (0,0167)	0,1755 (0,0024)
RCCK ₁	1,0290 (0,7556)	0,4618 (0,3010)	0,1959 (0,0150)	0,1755 (0,0024)
RCCK ₂	0,3495 (0,1295)	0,4166 (0,2028)	0,1980 (0,0160)	0,1755 (0,0024)
CP	1,3281 (0,3869)	0,6440 (0,3350)	0,60213 (0,0326)	0,8357 (0,0505)

Tabela 2 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 3$ e $n = 15$ e $\alpha_1 = 0,99$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,2317 (0,1848)	0,2223 (0,1432)	0,4803 (0,1205)	0,7649 (0,0596)
RCCK ₁	0,1583 (0,1073)	0,1235 (0,1015)	0,3744 (0,1156)	0,7614 (0,0583)
RCCK ₂	0,1137 (0,0832)	0,0560 (0,0378)	0,0888 (0,0487)	0,5889 (0,0473)
CP	0,2293 (0,1632)	0,0828 (0,0684)	0,3615 (0,1011)	0,6351 (0,0380)
<u>b. $\phi = 0,5$</u>				
MQ	0,1279 (0,1541)	0,0451 (0,0329)	0,0198 (0,0031)	0,0191 (0,0007)
RCCK ₁	0,1031 (0,0913)	0,0203 (0,0239)	0,0191 (0,0030)	0,0191 (0,0007)
RCCK ₂	0,1067 (0,1138)	0,0122 (0,0112)	0,0192 (0,0031)	0,0191 (0,0007)
CP	0,0837 (0,1058)	0,0328 (0,0282)	0,0921 (0,0286)	0,3799 (0,0800)
<u>c. $\phi = 1,0$</u>				
MQ	0,7669 (0,1799)	0,5910 (0,0309)	0,5195 (0,0089)	0,5163 (0,0101)
RCCK ₁	0,7833 (0,1656)	0,5947 (0,0362)	0,5204 (0,0039)	0,5163 (0,0101)
RCCK ₂	0,7826 (0,1942)	0,5920 (0,0336)	0,5198 (0,0089)	0,5163 (0,0101)
CP	0,7508 (0,1595)	0,5877 (0,0310)	1,7410 (0,9938)	2,4741 (0,0423)

Tabela 3 - Valores médios e erros padrões correspondentes (entre parênteses) de L_1^2 para o modelo com $p = 3$ e $n = 15$ e $\alpha_1 = 0,30$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	4,4090 (8,1464)	11,7340 (18,3822)	24,1398 (16,2501)	48,2722 (49,2273)
RCCK ₁	2,2242 (1,7679)	4,3473 (3,3199)	18,0115 (10,6849)	48,1183 (48,9861)
RCCK ₂	0,9771 (0,2039)	0,8939 (0,2209)	0,9170 (0,7252)	34,4273 (26,3507)
CP	0,9919 (0,4383)	1,0160 (0,4465)	8,7979 (3,9416)	11,3774 (5,0647)
<u>b. $\phi = 0,5$</u>				
MQ	3,2533 (6,3139)	3,5339 (3,6893)	0,1074 (0,0527)	0,0755 (0,0188)
RCCK ₁	1,3387 (0,8469)	1,1765 (0,4987)	0,0918 (0,0428)	0,0751 (0,0186)
RCCK ₂	0,2531 (0,1508)	0,3039 (0,2171)	0,0866 (0,0396)	0,0751 (0,0186)
CP	0,4167 (0,2557)	0,3705 (0,2885)	0,1074 (0,0527)	0,0755 (0,0188)
<u>c. $\phi = 1,0$</u>				
MQ	1,6451 (2,7364)	1,2342 (1,0291)	0,1277 (0,0303)	0,1140 (0,0191)
RCCK ₁	0,7305 (0,6318)	0,5533 (0,3868)	0,1265 (0,0285)	0,1140 (0,0191)
RCCK ₂	0,2935 (0,0894)	0,3288 (0,1273)	0,1270 (0,0293)	0,1140 (0,0191)
CP	1,3551 (1,8060)	1,2067 (1,0585)	0,1251 (0,0301)	0,1054 (0,0131)

Tabela 4 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 3$ e $n = 15$ e $\alpha_1 = 0,30$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,2227 (0,1561)	0,1677 (0,1060)	0,5175 (0,0958)	0,7040 (0,1242)
RCCK ₁	0,1606 (0,1143)	0,0958 (0,0594)	0,4384 (0,1130)	0,7025 (0,1233)
RCCK ₂	0,1061 (0,0792)	0,0509 (0,0347)	0,1261 (0,1301)	0,6011 (0,1007)
CP	0,1875 (0,1475)	0,0867 (0,0855)	0,3864 (0,2174)	0,4171 (0,3056)
<u>b. $\phi = 0,5$</u>				
MQ	0,1318 (0,1058)	0,0834 (0,0589)	0,0729 (0,0084)	0,0753 (0,0086)
RCCK ₁	0,1039 (0,0712)	0,0514 (0,0454)	0,0712 (0,0077)	0,0753 (0,0086)
RCCK ₂	0,1201 (0,0896)	0,0377 (0,0387)	0,0711 (0,0075)	0,0753 (0,0086)
CP	0,1021 (0,0870)	0,0525 (0,0470)	0,0726 (0,0084)	0,0753 (0,0086)
<u>c. $\phi = 1,0$</u>				
MQ	0,4653 (0,1157)	0,3375 (0,0659)	0,2318 (0,0380)	0,2298 (0,0377)
RCCK ₁	0,4936 (0,1190)	0,3458 (0,0630)	0,2325 (0,0380)	0,2298 (0,0377)
RCCK ₂	0,4818 (0,1171)	0,3399 (0,0627)	0,2322 (0,0380)	0,2298 (0,0377)
CP	0,4626 (0,1143)	0,3368 (0,0664)	0,2298 (0,0360)	0,2231 (0,0340)

Tabela 5 - Valores médios e erros padrões correspondentes (entre parênteses) de L_1^2 para o modelo com $p = 3$ e $n = 15$ e $\alpha_1 = 0,90$; $\alpha_2 = 0,90$ e $\alpha_3 = 0,90$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	1,1156 (1,4668)	0,4018 (0,4664)	2,0720 (1,0743)	2,3466 (1,1836)
RCCK ₁	0,7597 (0,8139)	0,3735 (0,3865)	1,9849 (1,0392)	2,3453 (1,1826)
RCCK ₂	0,6889 (0,6296)	0,6024 (0,2818)	1,5749 (0,7986)	2,3369 (1,1744)
CP	1,1156 (1,4668)	0,4018 (0,4664)	2,0720 (1,0743)	2,3466 (1,1836)
<u>b. $\phi = 0,5$</u>				
MQ	0,7841 (0,6574)	0,2707 (0,2985)	0,0373 (0,0196)	0,0256 (0,0063)
RCCK ₁	0,4806 (0,4301)	0,2207 (0,2217)	0,0350 (0,0167)	0,0256 (0,0063)
RCCK ₂	0,3839 (0,4059)	0,2344 (0,2458)	0,0352 (0,0172)	0,0256 (0,0063)
CP	0,7841 (0,6574)	0,2707 (0,2985)	0,0373 (0,0196)	0,0256 (0,0063)
<u>c. $\phi = 1,0$</u>				
MQ	0,5626 (0,2766)	0,2479 (0,0851)	0,1459 (0,0140)	0,1461 (0,0139)
RCCK ₁	0,4386 (0,1813)	0,2358 (0,0645)	0,1462 (0,0139)	0,1461 (0,0139)
RCCK ₂	0,4380 (0,1725)	0,2395 (0,0726)	0,1460 (0,0139)	0,1461 (0,0139)
CP	0,5626 (0,3766)	0,2479 (0,0851)	0,1459 (0,0140)	0,1461 (0,0139)

Tabela 6 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 3$ e $n = 15$ e $\alpha_1 = 0,90$; $\alpha_2 = 0,90$ e $\alpha_3 = 0,90$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,3134 (0,2685)	0,1534 (0,1347)	0,3852 (0,1014)	0,3997 (0,1003)
RCCK ₁)	0,2287 (0,1877)	0,1337 (0,1137)	0,3735 (0,0969)	0,3995 (0,1003)
RCCK ₂)	0,2026 (0,1663)	0,1559 (0,0832)	0,3210 (0,0718)	0,3985 (0,0997)
CP	0,3134 (0,2685)	0,1534 (0,1347)	0,3852 (0,1014)	0,3997 (0,1003)
<u>b. $\phi = 0,5$</u>				
MQ	0,1920 (0,0922)	0,0455 (0,0348)	0,0313 (0,0033)	0,0314 (0,0053)
RCCK ₁)	0,1436 (0,0590)	0,0437 (0,0239)	0,0304 (0,0027)	0,0314 (0,0053)
RCCK ₂)	0,1314 (0,0448)	0,0449 (0,0262)	0,0306 (0,0028)	0,0314 (0,0053)
CP	0,1920 (0,0922)	0,0455 (0,0348)	0,0313 (0,0033)	0,0314 (0,0053)
<u>c. $\phi = 1,0$</u>				
MQ	0,6189 (0,1139)	0,4473 (0,0327)	0,3655 (0,0493)	0,3641 (0,0491)
RCCK ₁)	0,6381 (0,1069)	0,4760 (0,0286)	0,3663 (0,0492)	0,3641 (0,0491)
RCCK ₂)	0,6239 (0,1047)	0,4595 (0,0289)	0,3659 (0,0492)	0,3641 (0,0491)
CP	0,6189 (0,1139)	0,4473 (0,0327)	0,3655 (0,0493)	0,3641 (0,0491)

Tabela 7 - Valores médios e erros padrões correspondentes (entre parênteses) de L_1^2 para o modelo com $p = 3$ e $n = 15$ e $\alpha_1 = 0,50$; $\alpha_2 = 0,70$ e $\alpha_3 = 0,95$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,1852 (0,1448)	0,2272 (0,1589)	0,7174 (0,5988)	0,7279 (0,5138)
RCCK ₁	0,3507 (0,2390)	0,1184 (0,0784)	0,6958 (0,5884)	0,7277 (0,5137)
RCCK ₂	0,5363 (0,2068)	0,1127 (0,0951)	0,6378 (0,5193)	0,7310 (0,5127)
CP	0,1852 (0,1448)	0,2272 (0,1589)	0,7174 (0,5988)	0,7279 (0,5138)
<u>b. $\phi = 0,5$</u>				
MQ	0,1493 (0,1544)	0,0710 (0,0637)	0,0632 (0,0370)	0,0517 (0,0147)
RCCK ₁	0,1859 (0,1203)	0,0581 (0,0507)	0,0584 (0,0352)	0,0517 (0,0147)
RCCK ₂	0,1791 (0,1212)	0,0594 (0,0553)	0,0587 (0,0348)	0,0517 (0,0147)
CP	0,1493 (0,1544)	0,0710 (0,0637)	0,0632 (0,0370)	0,0517 (0,0147)
<u>c. $\phi = 1,0$</u>				
MQ	0,2551 (0,0995)	0,1721 (0,0491)	0,1166 (0,0294)	0,1086 (0,0226)
RCCK ₁	0,2628 (0,0789)	0,1779 (0,0444)	0,1168 (0,0291)	0,1086 (0,0226)
RCCK ₂	0,2546 (0,0847)	0,1744 (0,0460)	0,1167 (0,0293)	0,1086 (0,0226)
CP	0,2551 (0,0995)	0,1721 (0,0491)	0,1166 (0,0294)	0,1086 (0,0226)

Tabela 8 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 3$ e $n = 15$ e $\alpha_1 = 0,50$; $\alpha_2 = 0,70$ e $\alpha_3 = 0,95$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,1385 (0,0920)	0,1233 (0,1168)	0,2417 (0,0890)	0,2550 (0,0941)
RCCK ₁	0,1680 (0,0913)	0,0922 (0,1170)	0,2356 (0,0891)	0,2550 (0,0941)
RCCK ₂	0,1955 (0,0847)	0,0834 (0,1153)	0,2218 (0,0850)	0,2549 (0,0940)
CP	0,1385 (0,0920)	0,1233 (0,1168)	0,2417 (0,0890)	0,2550 (0,0941)
<u>b. $\phi = 0,5$</u>				
MQ	0,1054 (0,1228)	0,0574 (0,0656)	0,0526 (0,0114)	0,0555 (0,0134)
RCCK ₁	0,1274 (0,1105)	0,0500 (0,0586)	0,0503 (0,0108)	0,0555 (0,0134)
RCCK ₂	0,1223 (0,1092)	0,0511 (0,0613)	0,0505 (0,0108)	0,0555 (0,0134)
CP	0,1054 (0,1228)	0,0574 (0,0656)	0,0526 (0,0114)	0,0555 (0,0134)
<u>c. $\phi = 1,0$</u>				
MQ	0,4197 (0,1097)	0,2917 (0,0658)	0,2105 (0,0665)	0,2058 (0,0670)
RCCK ₁	0,4852 (0,1048)	0,3149 (0,0749)	0,2115 (0,0664)	0,2058 (0,0670)
RCCK ₂	0,4556 (0,1079)	0,3032 (0,0688)	0,2110 (0,0664)	0,2058 (0,0670)
CP	0,4197 (0,1097)	0,2917 (0,0658)	0,2105 (0,0665)	0,2058 (0,0670)

Tabela 9 - Valores médios e erros padrões correspondentes (entre parênteses) de L_1^2 para o modelo com $p = 3$ e $n = 50$ e $\alpha_1 = 0,99$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	2,0785 (1,7268)	2,5677 (2,9205)	33,6581 (9,4534)	61,5429 (19,4173)
RCCK ₁)	0,9645 (0,6933)	1,4248 (1,0869)	30,3240 (8,3314)	61,4717 (19,3911)
RCCK ₂)	0,9331 (0,0966)	0,8614 (0,1881)	0,4161 (0,3250)	55,8922 (16,2821)
CP	1,0125 (0,0112)	1,0038 (0,0055)	3,9528 (3,2155)	33,8749 (16,7826)
<u>b. $\phi = 0,5$</u>				
MQ	1,1799 (1,0007)	0,6940 (0,5961)	0,0377 (0,0419)	0,0136 (0,0022)
RCCK ₁)	0,4858 (0,3991)	0,3486 (0,1900)	0,0276 (0,0269)	0,0136 (0,0022)
RCCK ₂)	0,2317 (0,1037)	0,1946 (0,1266)	0,0276 (0,0276)	0,0136 (0,0022)
CP	0,2694 (0,0163)	0,2508 (0,0007)	0,1598 (0,1375)	0,2573 (0,0022)
<u>c. $\phi = 1,0$</u>				
MQ	0,8158 (0,5986)	0,4283 (0,2072)	0,1815 (0,0053)	0,1756 (0,0013)
RCCK ₁)	0,4526 (0,3037)	0,2963 (0,1023)	0,1812 (0,0049)	0,1756 (0,0013)
RCCK ₂)	0,3837 (0,1371)	0,3168 (0,0993)	0,1814 (0,0051)	0,1756 (0,0013)
CP	0,2480 (0,0209)	0,3666 (0,2194)	1,0043 (0,0045)	1,0000 (0,0000)

Tabela 10 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 3$ e $n = 50$ e $\alpha_1 = 0,99$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,0736 (0,0529)	0,0562 (0,0438)	0,4387 (0,0279)	0,7802 (0,0335)
RCCK ₁	0,0536 (0,0422)	0,0348 (0,0245)	0,3983 (0,0293)	0,7794 (0,0335)
RCCK ₂	0,0393 (0,0253)	0,0176 (0,0118)	0,0226 (0,0210)	0,7161 (0,0183)
CP	0,0503 (0,0334)	0,0246 (0,0170)	0,0955 (0,1045)	0,5306 (0,3579)
<u>b. $\phi = 0,5$</u>				
MQ	0,0782 (0,0492)	0,0144 (0,0105)	0,0175 (0,0003)	0,0189 (0,0005)
RCCK ₁	0,0696 (0,0508)	0,0080 (0,0050)	0,0173 (0,0002)	0,0189 (0,0005)
RCCK ₂	0,0699 (0,0559)	0,0056 (0,0034)	0,0173 (0,0002)	0,0189 (0,0005)
CP	0,0607 (0,0497)	0,0057 (0,0025)	0,3781 (0,3951)	0,7404 (0,0040)
<u>c. $\phi = 1,0$</u>				
MQ	0,7443 (0,0618)	0,5893 (0,0213)	0,5220 (0,0062)	0,5196 (0,0067)
RCCK ₁	0,7529 (0,0677)	0,5920 (0,0207)	0,5223 (0,0062)	0,5196 (0,0067)
RCCK ₂	0,7470 (0,0633)	0,5902 (0,0210)	0,5221 (0,0062)	0,5196 (0,0067)
CP	0,7343 (0,0596)	0,5887 (0,0208)	2,9613 (0,0159)	2,9612 (0,0159)

Tabela 11 - Valores médios e erros padrões correspondentes (entre parênteses) de L_1^2 para o modelo com $p = 3$ e $n = 50$ e $\alpha_1 = 0,30$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,8408 (1,1007)	0,8178 (0,4824)	18,3752 (3,6691)	29,1804 (6,2913)
RCCK ₁	0,7747 (0,4583)	0,1147 (0,1083)	17,1269 (3,4003)	29,1555 (6,2794)
RCCK ₂	0,9972 (0,0556)	0,8472 (0,0908)	2,1872 (0,7038)	28,1175 (5,6533)
CP	1,0309 (0,0261)	1,0084 (0,0086)	18,3752 (3,6691)	29,1804 (6,2913)
<u>b. $\phi = 0,5$</u>				
MQ	0,7115 (0,9688)	0,0952 (0,0719)	0,0638 (0,0293)	0,0709 (0,0113)
RCCK ₁	0,3657 (0,2969)	0,0833 (0,0757)	0,0534 (0,0226)	0,0708 (0,0113)
RCCK ₂	0,3069 (0,1182)	0,0784 (0,0551)	0,0535 (0,0218)	0,0708 (0,0113)
CP	0,2939 (0,0222)	0,2124 (0,0980)	0,0638 (0,0293)	0,0709 (0,0113)
<u>c. $\phi = 1,0$</u>				
MQ	0,5164 (0,4353)	0,1531 (0,0229)	0,0999 (0,0092)	0,0969 (0,0105)
RCCK ₁	0,3176 (0,1612)	0,1379 (0,0104)	0,0997 (0,0092)	0,0969 (0,0105)
RCCK ₂	0,2790 (0,0807)	0,1414 (0,0134)	0,0997 (0,0092)	0,0969 (0,0105)
CP	0,2204 (0,0286)	0,1248 (0,0065)	0,0989 (0,0089)	0,0959 (0,0099)

Tabela 12 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 3$ e $n = 50$ e $\alpha_1 = 0,30$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,0621 (0,0559)	0,0324 (0,0067)	0,4630 (0,0232)	0,7070 (0,0234)
RCCK ₁	0,0483 (0,0347)	0,0145 (0,0105)	0,4331 (0,0247)	0,7064 (0,0232)
RCCK ₂	0,0411 (0,0187)	0,0265 (0,0053)	0,0746 (0,0220)	0,6829 (0,0145)
CP	0,0666 (0,0325)	0,0369 (0,0106)	0,4630 (0,0232)	0,7070 (0,0234)
<u>b. $\phi = 0,5$</u>				
MQ	0,0887 (0,0565)	0,0106 (0,0046)	0,0692 (0,0079)	0,0746 (0,0086)
RCCK ₁	0,0942 (0,0612)	0,0085 (0,0066)	0,0686 (0,0078)	0,0746 (0,0086)
RCCK ₂	0,0959 (0,0617)	0,0087 (0,0057)	0,0686 (0,0078)	0,0746 (0,0086)
CP	0,0782 (0,0437)	0,0149 (0,0062)	0,0692 (0,0079)	0,0746 (0,0086)
<u>c. $\phi = 1,0$</u>				
MQ	0,4576 (0,0545)	0,2615 (0,0241)	0,2048 (0,0297)	0,2027 (0,0304)
RCCK ₁	0,4766 (0,0632)	0,2685 (0,0237)	0,2051 (0,0297)	0,2027 (0,0304)
RCCK ₂	0,4700 (0,0645)	0,2650 (0,0238)	0,2049 (0,0297)	0,2027 (0,0304)
CP	0,4502 (0,0563)	0,2607 (0,0242)	0,2041 (0,0296)	0,2019 (0,0300)

Tabela 13 - Valores médios e erros padrões correspondentes (entre parênteses) de L_1^2 para o modelo com $p = 3$ e $n = 50$ e $\alpha_1 = 0,90$; $\alpha_2 = 0,90$ e $\alpha_3 = 0,90$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,0916 (0,0745)	0,3615 (0,1990)	2,3368 (0,8524)	2,6056 (0,9939)
RCCK ₁	0,2296 (0,1406)	0,2058 (0,1291)	2,3041 (0,8369)	2,6052 (0,9938)
RCCK ₂	0,5096 (0,3136)	0,0245 (0,0153)	2,1077 (0,7064)	2,6029 (0,9923)
CP	0,0916 (0,0745)	0,3615 (0,1990)	2,3368 (0,8524)	2,6056 (0,9939)
<u>b. $\phi = 0,5$</u>				
MQ	0,1049 (0,0981)	0,0164 (0,0137)	0,0263 (0,0127)	0,0224 (0,0053)
RCCK ₁	0,1199 (0,0971)	0,0191 (0,0170)	0,0251 (0,0122)	0,0224 (0,0053)
RCCK ₂	0,1094 (0,0972)	0,0166 (0,0155)	0,0254 (0,0122)	0,0224 (0,0053)
CP	0,1049 (0,0981)	0,0164 (0,0137)	0,0263 (0,0127)	0,0224 (0,0053)
<u>c. $\phi = 1,0$</u>				
MQ	0,2588 (0,0477)	0,1714 (0,0095)	0,1450 (0,0113)	0,1433 (0,0114)
RCCK ₁	0,2506 (0,0344)	0,1737 (0,0095)	0,1451 (0,0113)	0,1433 (0,0114)
RCCK ₂	0,2529 (0,0398)	0,1723 (0,0094)	0,1450 (0,0113)	0,1433 (0,0114)
CP	0,2598 (0,0477)	0,1714 (0,0095)	0,1450 (0,0113)	0,1433 (0,0114)

Tabela 14 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 3$ e $n = 50$ e $\alpha_1 = 0,90$; $\alpha_2 = 0,90$ e $\alpha_3 = 0,90$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,0410 (0,0246)	0,0674 (0,0274)	0,3408 (0,0497)	0,3754 (0,0587)
RCCK ₁	0,0574 (0,0404)	0,0452 (0,0270)	0,3363 (0,0484)	0,3754 (0,0586)
RCCK ₂	0,0916 (0,0467)	0,0191 (0,0270)	0,3101 (0,0356)	0,3750 (0,0585)
CP	0,0410 (0,0246)	0,0674 (0,0274)	0,3408 (0,0497)	0,3754 (0,0587)
<u>b. $\phi = 0,5$</u>				
MQ	0,0317 (0,0153)	0,0067 (0,0067)	0,0279 (0,0056)	0,0299 (0,0060)
RCCK ₁	0,0389 (0,0132)	0,0065 (0,0066)	0,0276 (0,0055)	0,0299 (0,0060)
RCCK ₂	0,0354 (0,0118)	0,0063 (0,0066)	0,0277 (0,0055)	0,0299 (0,0060)
CP	0,0317 (0,0153)	0,0067 (0,0067)	0,0279 (0,0056)	0,0299 (0,0060)
<u>c. $\phi = 1,0$</u>				
MQ	0,5552 (0,0432)	0,4307 (0,0441)	0,3737 (0,0464)	0,3713 (0,0465)
RCCK ₁	0,5746 (0,0449)	0,4382 (0,0445)	0,3740 (0,0464)	0,3713 (0,0465)
RCCK ₂	0,5635 (0,0435)	0,4336 (0,0441)	0,3738 (0,0464)	0,3713 (0,0465)
CP	0,5552 (0,0432)	0,4307 (0,0441)	0,3737 (0,0464)	0,3713 (0,0465)

Tabela 15 - Valores médios e erros padrões correspondentes (entre parênteses) de L_1^2 para o modelo com $p = 3$ e $n = 50$ e $\alpha_1 = 0,50$; $\alpha_2 = 0,70$ e $\alpha_3 = 0,95$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,0666 (0,0616)	0,1007 (0,0798)	0,6483 (0,2571)	0,6958 (0,2841)
RCCK ₁	0,0871 (0,0628)	0,0661 (0,0627)	0,6422 (0,2536)	0,6957 (0,2841)
RCCK ₂	0,1638 (0,1175)	0,0345 (0,0372)	0,6268 (0,2413)	0,6955 (0,2840)
CP	0,0666 (0,0616)	0,1007 (0,0798)	0,6483 (0,2571)	0,6958 (0,2841)
<u>b. $\phi = 0,5$</u>				
MQ	0,1551 (0,1132)	0,0155 (0,0190)	0,0443 (0,0128)	0,0500 (0,0120)
RCCK ₁	0,1008 (0,0486)	0,0182 (0,0221)	0,0435 (0,0126)	0,0500 (0,0120)
RCCK ₂	0,0880 (0,0413)	0,0169 (0,0210)	0,0436 (0,0125)	0,0500 (0,0120)
CP	0,1551 (0,1132)	0,0155 (0,0190)	0,0443 (0,0128)	0,0500 (0,0120)
<u>c. $\phi = 1,0$</u>				
MQ	0,2727 (0,1271)	0,1322 (0,0243)	0,0942 (0,0236)	0,0928 (0,0242)
RCCK ₁	0,2668 (0,0940)	0,1351 (0,0223)	0,0943 (0,0236)	0,0928 (0,0242)
RCCK ₂	0,2647 (0,0938)	0,1337 (0,0231)	0,0943 (0,0236)	0,0928 (0,0242)
CP	0,2727 (0,1271)	0,1322 (0,0243)	0,0942 (0,0236)	0,0928 (0,0242)

Tabela 16 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 3$ e $n = 50$ e $\alpha_1 = 0,50$; $\alpha_2 = 0,70$ e $\alpha_3 = 0,95$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,0404 (0,0305)	0,0713 (0,0827)	0,2156 (0,0960)	0,2291 (0,1021)
RCCK ₁	0,0467 (0,0304)	0,0611 (0,0800)	0,2140 (0,0955)	0,2291 (0,1021)
RCCK ₂	0,0681 (0,0372)	0,0512 (0,0736)	0,2100 (0,0931)	0,2291 (0,1021)
CP	0,0404 (0,0305)	0,0713 (0,0827)	0,2156 (0,0960)	0,2291 (0,1021)
<u>b. $\phi = 0,5$</u>				
MQ	0,0850 (0,0507)	0,0108 (0,0089)	0,0533 (0,0121)	0,0563 (0,0127)
RCCK ₁	0,0916 (0,0424)	0,0102 (0,0088)	0,0529 (0,0120)	0,0563 (0,0127)
RCCK ₂	0,0909 (0,0463)	0,0101 (0,0089)	0,0529 (0,0120)	0,0563 (0,0127)
CP	0,0850 (0,0507)	0,0108 (0,0089)	0,0533 (0,0121)	0,0563 (0,0127)
<u>c. $\phi = 1,0$</u>				
MQ	0,4370 (0,1032)	0,2523 (0,0591)	0,1911 (0,0641)	0,1888 (0,0645)
RCCK ₁	0,4615 (0,0990)	0,2599 (0,0575)	0,1913 (0,0641)	0,1888 (0,0645)
RCCK ₂	0,4503 (0,1007)	0,2562 (0,0580)	0,1912 (0,0641)	0,1888 (0,0645)
CP	0,4370 (0,1032)	0,2523 (0,0591)	0,1911 (0,0641)	0,1888 (0,0645)

Tabela 17 - Valores médios e erros padrões correspondentes (entre parênteses) de L_1^2 para o modelo com $p = 8$ e $n = 15$ e $\alpha_1 = 0,99$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	44,2458 (29,8011)	25,9392 (12,0640)	62,3810 (44,9066)	193,2401 (128,7109)
RCCK ₁	13,1886 (10,0642)	5,1722 (3,2690)	19,2204 (18,3833)	189,0730 (127,2672)
RCCK ₂	1,0178 (0,0155)	1,0356 (0,0317)	0,9973 (0,1173)	11,6397 (4,3124)
CP	2,2153 (2,6597)	2,7636 (2,2273)	2,0603 (1,7451)	29,9688 (19,7144)
<u>b. $\phi = 0,5$</u>				
MQ	17,4880 (12,9612)	5,2675 (5,6530)	0,3886 (0,1413)	0,0870 (0,0096)
RCCK ₁	5,4324 (3,8099)	1,2819 (0,6508)	0,4097 (0,1183)	0,0898 (0,0102)
RCCK ₂	0,9042 (0,2077)	0,8146 (0,1241)	0,3892 (0,1228)	0,0879 (0,0099)
CP	3,7141 (2,1373)	2,2070 (0,9528)	0,4938 (0,2401)	0,2438 (0,1051)
<u>c. $\phi = 1,0$</u>				
MQ	7,0522 (4,7864)	1,6839 (1,2579)	0,4750 (0,0374)	0,4137 (0,0021)
RCCK ₁	2,3361 (1,4777)	0,6078 (0,1219)	0,4425 (0,0171)	0,4136 (0,0020)
RCCK ₂	0,6634 (0,1637)	0,5659 (0,0553)	0,4628 (0,0281)	0,4136 (0,0020)
CP	4,3718 (3,6677)	1,4545 (1,1204)	0,7168 (0,3838)	0,6479 (0,3215)

Tabela 18 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 8$ e $n = 15$ e $\alpha_1 = 0,99$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,5436 (0,1721)	0,5824 (0,1924)	0,7129 (0,1281)	0,9296 (0,0316)
RCCK ₁	0,2730 (0,1764)	0,3442 (0,2379)	0,4965 (0,1987)	0,9178 (0,0329)
RCCK ₂	0,0644 (0,0952)	0,1030 (0,0768)	0,2282 (0,2110)	0,5125 (0,3221)
CP	0,1233 (0,1295)	0,2569 (0,1698)	0,4968 (0,6550)	0,5614 (0,4134)
<u>b. $\phi = 0,5$</u>				
MQ	0,5493 (0,1548)	0,3066 (0,0853)	0,1684 (0,0029)	0,1646 (0,0030)
RCCK ₁	0,4474 (0,1096)	0,2723 (0,0608)	0,1688 (0,0023)	0,1746 (0,0030)
RCCK ₂	0,3996 (0,0845)	0,2611 (0,0572)	0,1685 (0,0026)	0,1646 (0,0030)
CP	0,4184 (0,1104)	0,7246 (0,6604)	0,2368 (0,2194)	0,1691 (0,0112)
<u>c. $\phi = 1,0$</u>				
MQ	3,5500 (0,1110)	3,3292 (0,0317)	3,2430 (0,0419)	3,2407 (0,0424)
RCCK ₁	3,5177 (0,0884)	3,3281 (0,0261)	3,2440 (0,0415)	3,2407 (0,0424)
RCCK ₂	3,5175 (0,0879)	3,3278 (0,0291)	3,2432 (0,0419)	3,2407 (0,0424)
CP	4,4396 (1,9148)	3,4464 (1,0879)	5,7598 (2,8156)	5,5534 (2,5571)

Tabela 19 - Valores médios e erros padrões correspondentes (entre parênteses) de L_1^2 para o modelo com $p = 8$ e $n = 15$ e $\alpha_1 = 0,30$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	52,0807 (33,7302)	45,4823 (42,1309)	84,3740 (79,8898)	216,3026 (77,4373)
RCCK ₁	10,1467 (7,9182)	10,4621 (8,7127)	25,9110 (12,0855)	210,0319 (79,7334)
RCCK ₂	1,0750 (0,0931)	1,1442 (0,2036)	1,0183 (0,1103)	37,3535 (22,6965)
CP	2,4160 (1,4983)	3,0741 (3,6786)	1,3565 (0,3240)	1,1415 (0,1340)
<u>b. $\phi = 0,5$</u>				
MQ	22,0550 (21,6639)	5,6665 (6,7335)	0,2398 (0,1103)	0,0684 (0,0103)
RCCK ₁	4,1434 (3,7296)	1,8691 (1,0822)	0,2796 (0,1359)	0,0708 (0,0096)
RCCK ₂	0,6388 (0,0877)	0,6153 (0,2538)	0,2606 (0,1205)	0,0696 (0,0099)
CP	1,1985 (0,8737)	4,3393 (2,9414)	0,2325 (0,1415)	0,0672 (0,0100)
<u>c. $\phi = 1,0$</u>				
MQ	9,8797 (11,4061)	2,2893 (2,0097)	0,3844 (0,0386)	0,3316 (0,0074)
RCCK ₁	2,1162 (1,9736)	0,8763 (0,4319)	0,3513 (0,0150)	0,3316 (0,0074)
RCCK ₂	0,5315 (0,1415)	0,4912 (0,1245)	0,3667 (0,0250)	0,3316 (0,0074)
CP	2,3907 (2,6698)	2,1801 (2,0801)	0,3771 (0,0442)	0,3306 (0,0077)

Tabela 20 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 8$ e $n = 15$ e $\alpha_1 = 0,30$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,6414 (0,0978)	0,5871 (0,1845)	0,6306 (0,0827)	0,9075 (0,0216)
RCCK ₁	0,4751 (0,1385)	0,3925 (0,1844)	0,4195 (0,1371)	0,8903 (0,0239)
RCCK ₂	0,2102 (0,1459)	0,1630 (0,1122)	0,1648 (0,1480)	0,4008 (0,1968)
CP	0,3959 (0,1668)	0,3863 (0,1738)	0,2903 (0,2130)	0,5250 (0,6330)
<u>b. $\phi = 0,5$</u>				
MQ	0,4337 (0,1875)	0,1813 (0,0416)	0,0946 (0,0097)	0,0902 (0,0120)
RCCK ₁	0,3632 (0,1514)	0,1647 (0,0375)	0,0957 (0,0090)	0,0902 (0,0120)
RCCK ₂	0,3430 (0,1300)	0,1631 (0,0402)	0,0952 (0,0094)	0,0902 (0,0120)
CP	0,3356 (0,1262)	0,2084 (0,0853)	0,0939 (0,0104)	0,0896 (0,0123)
<u>c. $\phi = 1,0$</u>				
MQ	2,0795 (0,1169)	1,8948 (0,0918)	1,8205 (0,1033)	1,8176 (0,1043)
RCCK ₁	2,0572 (0,1014)	1,8960 (0,0917)	1,8226 (0,1024)	1,8176 (0,1043)
RCCK ₂	2,0625 (0,1047)	1,8957 (0,0912)	1,8210 (0,1031)	1,8176 (0,1043)
CP	0,8522 (0,9651)	1,9007 (0,6481)	1,8635 (0,1801)	1,8173 (0,1038)

Tabela 21 - Valores médios e erros padrões correspondentes (entre parênteses) de L_1^2 para o modelo com $p = 8$ e $n = 15$ e $\alpha_1 = 0,90$; $\alpha_2 = 0,90$ e $\alpha_3 = 0,90$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	8,8354 (9,2165)	6,3579 (5,9314)	13,0428 (11,6299)	20,9688 (25,4598)
RCCK ₁	3,2253 (2,2804)	2,7911 (2,1414)	8,0674 (6,6681)	20,8953 (25,3579)
RCCK ₂	1,0997 (0,2371)	1,1452 (0,4322)	0,5985 (0,2355)	17,0802 (17,7264)
CP	3,4745 (2,4283)	2,5283 (1,2749)	7,8006 (6,3249)	9,3189 (8,0490)
<u>b. $\phi = 0,5$</u>				
MQ	2,2646 (1,4103)	1,0965 (0,9676)	0,0963 (0,0423)	0,0659 (0,0225)
RCCK ₁	0,9974 (0,4180)	0,6942 (0,2778)	0,1109 (0,0582)	0,0660 (0,0225)
RCCK ₂	0,7738 (0,3018)	0,6632 (0,2721)	0,1045 (0,0516)	0,0660 (0,0225)
CP	1,8302 (1,0235)	1,0965 (0,9676)	0,0963 (0,0423)	0,0659 (0,0225)
<u>c. $\phi = 1,0$</u>				
MQ	1,1276 (0,3850)	0,6170 (0,2080)	0,3835 (0,0172)	0,3714 (0,0161)
RCCK ₁	0,6491 (0,1460)	0,4741 (0,0715)	0,3809 (0,0168)	0,3714 (0,0161)
RCCK ₂	0,6508 (0,1368)	0,5065 (0,0791)	0,3829 (0,0172)	0,3714 (0,0161)
CP	1,1276 (0,3850)	0,6170 (0,2080)	0,3933 (0,0257)	0,3694 (0,0179)

Tabela 22 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 8$ e $n = 15$ e $\alpha_1 = 0,90$; $\alpha_2 = 0,90$ e $\alpha_3 = 0,90$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,6212 (0,2413)	0,5286 (0,2130)	0,6869 (0,1302)	0,7405 (0,1362)
RCCK ₁	0,3618 (0,2018)	0,3038 (0,2273)	0,5870 (0,1319)	0,7394 (0,1359)
RCCK ₂	0,1511 (0,1432)	0,1296 (0,0981)	0,3334 (0,2320)	0,6930 (0,1352)
CP	0,5738 (0,2294)	0,4533 (0,2275)	0,5682 (0,1297)	0,4922 (0,3045)
<u>b. $\phi = 0,5$</u>				
MQ	0,4104 (0,1142)	0,2189 (0,0420)	0,1208 (0,0145)	0,1141 (0,0176)
RCCK ₁	0,3499 (0,0711)	0,2153 (0,0383)	0,1219 (0,0147)	0,1141 (0,0176)
RCCK ₂	0,3433 (0,0779)	0,2119 (0,0369)	0,1215 (0,0146)	0,1141 (0,0176)
CP	0,2905 (0,1045)	0,2189 (0,0420)	0,1208 (0,0145)	0,1141 (0,0176)
<u>c. $\phi = 1,0$</u>				
MQ	2,6316 (0,2133)	2,4835 (0,2787)	2,4173 (0,2859)	2,4137 (0,2867)
RCCK ₁	2,6390 (0,2276)	2,5048 (0,2677)	2,4198 (0,2850)	2,4137 (0,2867)
RCCK ₂	2,6293 (0,2173)	2,4879 (0,2759)	2,4177 (0,2858)	2,4137 (0,2867)
CP	2,6316 (0,2133)	2,4835 (0,2787)	2,5001 (0,3317)	2,4032 (0,2971)

Tabela 23 - Valores médios e erros padrões correspondentes (entre parênteses) de L_1^2 para o modelo com $p = 8$ e $n = 15$ e $\alpha_1 = 0,50$; $\alpha_2 = 0,70$ e $\alpha_3 = 0,95$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	5,0514 (4,2769)	3,6323 (1,7353)	6,5072 (2,0385)	9,5540 (4,4514)
RCCK ₁	3,0365 (2,0267)	1,5562 (1,0949)	5,0174 (1,5921)	9,5351 (4,4335)
RCCK ₂	1,5040 (0,7467)	0,7435 (0,1984)	0,5907 (0,2825)	9,1711 (3,9995)
CP	5,0514 (4,2769)	3,6323 (1,7353)	6,5072 (2,0385)	9,5540 (4,4514)
<u>b. $\phi = 0,5$</u>				
MQ	2,1469 (1,4776)	0,1684 (0,0944)	0,0724 (0,0417)	0,0317 (0,0058)
RCCK ₁	1,4591 (0,7565)	0,1931 (0,0371)	0,0807 (0,0478)	0,0317 (0,0058)
RCCK ₂	1,2896 (0,6006)	0,1937 (0,0292)	0,0776 (0,0449)	0,0317 (0,0058)
CP	2,1469 (1,4776)	0,1684 (0,0944)	0,0724 (0,0417)	0,0317 (0,0058)
<u>c. $\phi = 1,0$</u>				
MQ	1,0170 (0,4137)	0,3599 (0,0562)	0,2826 (0,0292)	0,2693 (0,0296)
RCCK ₁	0,6925 (0,3020)	0,3342 (0,0568)	0,2815 (0,0290)	0,2693 (0,0296)
RCCK ₂	0,6675 (0,2453)	0,3448 (0,0589)	0,2822 (0,0292)	0,2693 (0,0296)
CP	1,0170 (0,4137)	0,3599 (0,0562)	0,2826 (0,0292)	0,2693 (0,0296)

Tabela 24 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 8$ e $n = 15$ e $\alpha_1 = 0,50$; $\alpha_2 = 0,70$ e $\alpha_3 = 0,95$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,6919 (0,3750)	0,5068 (0,1276)	0,6609 (0,1365)	0,6623 (0,1240)
RCCK ₁	0,4611 (0,2449)	0,3381 (0,1401)	0,5827 (0,1456)	0,6608 (0,1224)
RCCK ₂	0,2301 (0,1259)	0,1621 (0,0567)	0,3244 (0,1922)	0,6095 (0,0403)
CP	0,6919 (0,3750)	0,5068 (0,1276)	0,6609 (0,1365)	0,6623 (0,1240)
<u>b. $\phi = 0,5$</u>				
MQ	0,3971 (0,1004)	0,1262 (0,0219)	0,0653 (0,0218)	0,0573 (0,0173)
RCCK ₁	0,3422 (0,0750)	0,1260 (0,0136)	0,0664 (0,0216)	0,0573 (0,0173)
RCCK ₂	0,3219 (0,0759)	0,1266 (0,0135)	0,0659 (0,0216)	0,0573 (0,0173)
CP	0,3971 (0,0759)	0,1262 (0,0219)	0,0653 (0,0218)	0,0573 (0,0173)
<u>c. $\phi = 1,0$</u>				
MQ	1,4982 (0,2961)	1,3036 (0,3005)	1,2519 (0,3072)	1,2469 (0,3069)
RCCK ₁	1,5149 (0,2831)	1,3295 (0,2960)	1,2538 (0,3067)	1,2469 (0,3069)
RCCK ₂	1,4997 (0,2809)	1,3105 (0,2986)	1,2524 (0,3070)	1,2469 (0,3069)
CP	1,4982 (0,2961)	1,3036 (0,3005)	1,2519 (0,3072)	1,2469 (0,3069)

Tabela 25 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 8$ e $n = 50$ e $\alpha_1 = 0,99$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	6,5135 (1,9605)	11,2175 (7,6207)	49,4511 (16,4663)	98,4144 (29,6702)
RCCK ₁	1,4938 (0,2820)	4,0975 (3,4328)	36,1597 (14,8001)	97,9769 (29,5402)
RCCK ₂	0,9846 (0,0223)	0,9585 (0,0462)	0,4005 (0,1687)	59,4557 (13,4453)
CP	1,0020 (0,0021)	1,0007 (0,0003)	1,0139 (0,0347)	17,6132 (6,4551)
<u>b. $\phi = 0,5$</u>				
MQ	2,4279 (0,7193)	1,1947 (0,5756)	0,1180 (0,0341)	0,0838 (0,0091)
RCCK ₁	0,7977 (0,2145)	0,7369 (0,3465)	0,1451 (0,0493)	0,0843 (0,0094)
RCCK ₂	0,6769 (0,1042)	0,6772 (0,2554)	0,1316 (0,0383)	0,0841 (0,0093)
CP	0,6859 (0,2009)	0,8149 (0,3529)	0,2930 (0,0969)	0,3102 (0,0054)
<u>c. $\phi = 1,0$</u>				
MQ	1,0878 (0,4506)	0,7120 (0,1299)	0,4250 (0,0053)	0,4141 (0,0013)
RCCK ₁	0,5784 (0,0635)	0,5323 (0,0779)	0,4233 (0,0045)	0,4140 (0,0012)
RCCK ₂	0,5713 (0,0336)	0,5924 (0,0842)	0,4248 (0,0051)	0,4140 (0,0012)
CP	0,5676 (0,0992)	0,6231 (0,1346)	0,9182 (0,2222)	0,9161 (0,2221)

Tabela 26 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 8$ e $n = 50$ e $\alpha_1 = 0,99$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,1176 (0,0221)	0,1813 (0,0726)	0,4468 (0,0914)	0,8160 (0,0211)
RCCK ₁	0,0407 (0,0166)	0,0782 (0,0482)	0,3345 (0,1053)	0,8125 (0,0208)
RCCK ₂	0,0200 (0,0124)	0,0131 (0,0049)	0,0131 (0,0095)	0,5141 (0,0785)
CP	0,0244 (0,0170)	0,0143 (0,0062)	0,0185 (0,0121)	0,2328 (0,1082)
<u>b. $\phi = 0,5$</u>				
MQ	0,4032 (0,0539)	0,2406 (0,0334)	0,1682 (0,0015)	0,1655 (0,0018)
RCCK ₁	0,3814 (0,0462)	0,2333 (0,0283)	0,1687 (0,0013)	0,1655 (0,0018)
RCCK ₂	0,3876 (0,0469)	0,2324 (0,0271)	0,1684 (0,0014)	0,1655 (0,0018)
CP	0,4468 (0,1566)	0,2305 (0,0258)	0,1894 (0,0230)	0,1722 (0,0028)
<u>c. $\phi = 1,0$</u>				
MQ	3,5061 (0,0663)	3,3177 (0,0149)	3,2546 (0,0268)	3,2522 (0,0273)
RCCK ₁	3,5063 (0,0628)	3,3204 (0,0131)	3,2553 (0,0267)	3,2522 (0,0272)
RCCK ₂	3,5061 (0,0645)	3,3180 (0,0144)	3,2547 (0,0268)	3,2522 (0,0273)
CP	3,4911 (0,0631)	3,2424 (0,0242)	7,2039 (1,7663)	7,2034 (1,7675)

Tabela 27 - Valores médios e erros padrões correspondentes (entre parênteses) de L_1^2 para o modelo com $p = 8$ e $n = 50$ e $\alpha_1 = 0,30$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	5,2187 (4,2278)	5,8514 (3,7735)	29,4976 (12,6780)	62,5203 (18,3951)
RCCK ₁	1,9563 (1,2501)	1,7180 (1,0034)	20,8928 (11,2541)	62,3128 (18,3229)
RCCK ₂	0,9912 (0,0169)	0,9336 (0,0817)	0,3250 (0,1700)	47,9441 (10,6437)
CP	1,0686 (0,0300)	1,0606 (0,0269)	1,0429 (0,0266)	47,5903 (12,9445)
<u>b. $\phi = 0,5$</u>				
MQ	2,3244 (1,7683)	0,7612 (0,1943)	0,0684 (0,0431)	0,0485 (0,0050)
RCCK ₁	0,9903 (0,5003)	0,4144 (0,1748)	0,1119 (0,0551)	0,0487 (0,0050)
RCCK ₂	0,6111 (0,1513)	0,4060 (0,1913)	0,1016 (0,0485)	0,0485 (0,0049)
CP	0,6059 (0,0371)	0,5905 (0,0621)	0,0870 (0,0418)	0,0484 (0,0049)
<u>c. $\phi = 1,0$</u>				
MQ	1,2029 (0,7140)	0,5381 (0,0379)	0,3285 (0,0060)	0,3193 (0,0054)
RCCK ₁	0,6321 (0,2467)	0,4052 (0,0272)	0,3321 (0,0130)	0,3193 (0,0054)
RCCK ₂	0,5236 (0,0989)	0,4538 (0,0352)	0,3281 (0,0059)	0,3193 (0,0054)
CP	0,7406 (0,5399)	0,5237 (0,0874)	0,3241 (0,0079)	0,3192 (0,0056)

Tabela 28 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 8$ e $n = 50$ e $\alpha_1 = 0,30$; $\alpha_2 = 0,99$ e $\alpha_3 = 0,99$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,1884 (0,0693)	0,1632 (0,0353)	0,4192 (0,0928)	0,7836 (0,0238)
RCCK ₁	0,0995 (0,0202)	0,0872 (0,0194)	0,3123 (0,1046)	0,7810 (0,0238)
RCCK ₂	0,0374 (0,0115)	0,0327 (0,0118)	0,0547 (0,0224)	0,6156 (0,0409)
CP	0,0825 (0,0258)	0,0716 (0,0269)	0,0818 (0,0438)	0,6508 (0,3032)
<u>b. $\phi = 0,5$</u>				
MQ	0,3332 (0,1060)	0,1457 (0,0186)	0,0792 (0,0072)	0,0768 (0,0073)
RCCK ₁	0,3025 (0,0913)	0,1430 (0,0191)	0,0798 (0,0071)	0,0768 (0,0073)
RCCK ₂	0,3136 (0,0939)	0,1429 (0,0185)	0,0795 (0,0072)	0,0768 (0,0073)
CP	0,2930 (0,0860)	0,2593 (0,1274)	0,0789 (0,0076)	0,0768 (0,0071)
<u>c. $\phi = 1,0$</u>				
MQ	1,9823 (0,1351)	1,7597 (0,0603)	1,6848 (0,0645)	1,6810 (0,0654)
RCCK ₁	1,9869 (0,1272)	1,7660 (0,0623)	1,6857 (0,0644)	1,6810 (0,0654)
RCCK ₂	1,9815 (0,1317)	1,7615 (0,0607)	1,6850 (0,0645)	1,6810 (0,0654)
CP	1,7344 (0,5243)	1,9576 (0,4737)	1,6853 (0,0651)	1,6817 (0,0660)

Tabela 29 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 8$ e $n = 50$ e $\alpha_1 = 0,90$; $\alpha_2 = 0,90$ e $\alpha_3 = 0,90$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,1272 (0,0309)	0,2150 (0,0387)	0,4382 (0,0319)	0,4898 (0,0405)
RCCK ₁	0,0955 (0,0193)	0,1498 (0,0425)	0,4197 (0,0346)	0,4896 (0,0405)
RCCK ₂	0,0954 (0,0131)	0,0924 (0,0610)	0,2972 (0,0505)	0,4876 (0,0406)
CP	0,1272 (0,0309)	0,2150 (0,0387)	0,4382 (0,0319)	0,4898 (0,0405)
<u>b. $\phi = 0,5$</u>				
MQ	0,3787 (0,0809)	0,2030 (0,0250)	0,1171 (0,0060)	0,1144 (0,0066)
RCCK ₁	0,3930 (0,0749)	0,2098 (0,0273)	0,1198 (0,0058)	0,1144 (0,0066)
RCCK ₂	0,3913 (0,0769)	0,2080 (0,0265)	0,1175 (0,0059)	0,1144 (0,0066)
CP	0,3787 (0,0809)	0,2030 (0,0250)	0,1171 (0,0060)	0,1144 (0,0066)
<u>c. $\phi = 1,0$</u>				
MQ	2,7255 (0,1156)	2,5201 (0,1075)	2,4465 (0,1200)	2,4439 (0,1207)
RCCK ₁	2,7654 (0,1055)	2,5349 (0,1049)	2,4473 (0,1197)	2,4439 (0,1207)
RCCK ₂	2,7343 (0,1141)	2,5229 (0,1069)	2,4467 (0,1199)	2,4439 (0,1207)
CP	2,7255 (0,1156)	2,5201 (0,1075)	2,4465 (0,1200)	2,4439 (0,1207)

Tabela 30 - Valores médios e erros padrões correspondentes (entre parênteses) de L_1^2 para o modelo com $p = 8$ e $n = 50$ e $\alpha_1 = 0,90$; $\alpha_2 = 0,90$ e $\alpha_3 = 0,90$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,6681 (0,3047)	1,1703 (0,4621)	4,0861 (0,6130)	4,4993 (0,6649)
RCCK ₁	0,6673 (0,1738)	0,6309 (0,2782)	3,8975 (0,6277)	4,4970 (0,6644)
RCCK ₂	0,8758 (0,0849)	0,3604 (0,1697)	2,6384 (0,6853)	4,4765 (0,6578)
CP	0,6681 (0,3047)	1,1703 (0,4621)	4,0861 (0,6130)	4,4993 (0,6649)
<u>b. $\phi = 0,5$</u>				
MQ	0,4373 (0,2061)	0,2138 (0,0836)	0,0703 (0,0104)	0,0642 (0,0086)
RCCK ₁	0,4626 (0,1369)	0,2351 (0,0926)	0,0736 (0,0112)	0,0642 (0,0086)
RCCK ₂	0,4548 (0,1412)	0,2279 (0,0849)	0,0721 (0,0108)	0,0642 (0,0086)
CP	0,4373 (0,2061)	0,2138 (0,0836)	0,0703 (0,0104)	0,0642 (0,0086)
<u>c. $\phi = 1,0$</u>				
MQ	0,5166 (0,0660)	0,4242 (0,0165)	0,3741 (0,0073)	0,3724 (0,0072)
RCCK ₁	0,4593 (0,0279)	0,4135 (0,0125)	0,3741 (0,0073)	0,3724 (0,0072)
RCCK ₂	0,4869 (0,0438)	0,4213 (0,0155)	0,3741 (0,0073)	0,3724 (0,0072)
CP	0,5166 (0,0660)	0,4242 (0,0165)	0,3741 (0,0073)	0,3724 (0,0072)

Tabela 31 - Valores médios e erros padrões correspondentes (entre parênteses) de L_1^2 para o modelo com $p = 8$ e $n = 50$ e $\alpha_1 = 0,50$; $\alpha_2 = 0,70$ e $\alpha_3 = 0,95$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,6436 (0,4724)	0,9352 (0,6975)	6,2844 (2,4909)	7,3692 (2,7860)
RCCK ₁)	0,4388 (0,3357)	0,3678 (0,2705)	5,9137 (2,3699)	7,3645 (2,7843)
RCCK ₂)	0,7664 (0,1751)	0,5614 (0,1992)	2,8781 (0,9664)	7,2991 (2,7415)
CP	0,6436 (0,4724)	0,9352 (0,6975)	6,2844 (2,4909)	7,3692 (2,7860)
<u>b. $\phi = 0,5$</u>				
MQ	0,3509 (0,2581)	0,1523 (0,1259)	0,0399 (0,0112)	0,0361 (0,0108)
RCCK ₁)	0,2926 (0,1651)	0,1937 (0,1156)	0,0421 (0,0119)	0,0361 (0,0108)
RCCK ₂)	0,3062 (0,1582)	0,1803 (0,1146)	0,0414 (0,0115)	0,0361 (0,0108)
CP	0,3509 (0,2581)	0,1523 (0,1259)	0,0399 (0,0112)	0,0361 (0,0108)
<u>c. $\phi = 1,0$</u>				
MQ	0,4761 (0,1055)	0,3251 (0,0299)	0,2870 (0,0186)	0,2855 (0,0185)
RCCK ₁)	0,3907 (0,0392)	0,3190 (0,0231)	0,2871 (0,0186)	0,2855 (0,0185)
RCCK ₂)	0,4081 (0,0541)	0,3225 (0,0271)	0,2870 (0,0186)	0,2855 (0,0185)
CP	0,4761 (0,1055)	0,3251 (0,0299)	0,2870 (0,0186)	0,2855 (0,0185)

Tabela 32 - Valores médios e erros padrões correspondentes (entre parênteses) de L_2^2 para o modelo com $p = 8$ e $n = 50$ e $\alpha_1 = 0,50$; $\alpha_2 = 0,70$ e $\alpha_3 = 0,95$

ESTIMADORES	$\rho = 1$	$\rho = 4$	$\rho = 100$	$\rho = 10000$
<u>a. $\phi = 0$</u>				
MQ	0,1411 (0,0370)	0,1584 (0,0447)	0,4659 (0,0480)	0,5327 (0,0526)
RCCK ₁	0,1068 (0,0291)	0,1071 (0,0383)	0,4401 (0,0515)	0,5324 (0,0526)
RCCK ₂	0,0985 (0,0258)	0,0893 (0,0316)	0,2374 (0,0781)	0,5281 (0,0509)
CP	0,1411 (0,0370)	0,1584 (0,0447)	0,4659 (0,0480)	0,5327 (0,0526)
<u>b. $\phi = 0,5$</u>				
MQ	0,2431 (0,0753)	0,0999 (0,0236)	0,0474 (0,0106)	0,0452 (0,0109)
RCCK ₁	0,2571 (0,0726)	0,1115 (0,0224)	0,0478 (0,0106)	0,0452 (0,0109)
RCCK ₂	0,2603 (0,0781)	0,1078 (0,0240)	0,0477 (0,0106)	0,0452 (0,0109)
CP	0,2431 (0,0753)	0,0999 (0,0236)	0,0474 (0,0106)	0,0452 (0,0109)
<u>c. $\phi = 1,0$</u>				
MQ	1,5687 (0,1949)	1,3740 (0,1923)	1,3126 (0,1863)	1,3098 (0,1864)
RCCK ₁	1,6073 (0,1912)	1,3901 (0,1936)	1,3134 (0,1861)	1,3098 (0,1864)
RCCK ₂	1,5805 (0,1939)	1,3778 (0,1925)	1,3128 (0,1862)	1,3098 (0,1864)
CP	1,5687 (0,1949)	1,3740 (0,1923)	1,3132 (0,1874)	1,3107 (0,1882)

8. APÊNDICE 2

Roteiro de operações efetuadas com o Minitab para a obtenção dos valores de L_1^2 e de L_2^2 para os estimadores estudados

- 1-) Geraram-se os vetores aleatórios de dimensões $n \times 1$ com distribuição $N(0,1)$;
- 2-) Obtiveram-se as colunas da matriz X segundo uma estrutura de correlações desejada, conforme descrito em (3.1);
- 3-) Determinaram-se os valores médios e as somas dos quadrados dos desvios da média para cada coluna da matriz X ;
- 4-) Obtiveram-se as matrizes X na forma padronizada subtraindo-se de cada elemento a média e dividindo-se pela soma dos quadrados dos desvios da coluna correspondente;
- 5-) Determinaram-se a seguir as matrizes X' , $X'X$, $(X'X)^{-1}$;
- 6-) Determinaram-se então o vetor dos auto-valores $\lambda_1, \dots, \lambda_p$ e a matriz $V = [v_1, \dots, v_p]$ dos auto vetores correspondentes de $X'X$;
- 7-) Multiplicaram-se a matriz X pelos três diferentes valores de β descritos em (3.1) obtendo-se o vetor $X\beta$;
- 8-) Geraram-se os vetores aleatórios ε_i , $i = 1, 2, 3, 4$ com distribuição $N(0, \sigma_i)$, $\sigma_i = 0.01, 0.1, 0.5$ e 1.0 ;
- 9-) Adicionaram-se os vetores $X\beta$ e ε_i , $i = 1, 2, 3, 4$, obtendo-se os vetores da variável dependente y para as situações estudadas (3 valores de β e 4 valores de σ);
- 10-) Determinaram-se os valores médios e as somas dos

quadrados dos desvios da média para os vetores y padronizando-os a seguir do mesmo modo que para as variáveis independentes;

11-) Obtiveram-se então as matrizes $X'y$ e a seguir o vetor $\hat{\beta} = (X'X)^{-1}X'y$;

12-) De posse dos valores de $\hat{\beta}$, β e $X'X$, obtiveram-se a seguir os valores de L_1^2 e de L_2^2 para o estimador de mínimos quadrados;

13-) Determinaram-se então os valores de $y - X\hat{\beta}$ e a seguir

$$\hat{\sigma}^2 = \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{n-p} ;$$

14-) Obtiveram-se então os valores de $k_1 = \frac{p \hat{\sigma}^2}{\hat{\beta}'\hat{\beta}}$ e

$$k_2 = \frac{p \hat{\sigma}^2}{\sum \lambda_i \hat{\alpha}_i^2} , \text{ onde } \hat{\alpha}_i = v_i' \hat{\beta} ;$$

15-) Determinaram-se as matrizes $(X'X + kI)^{-1}$ e a seguir os estimadores $\beta^* = (X'X + kI)^{-1}X'y$, obtendo-se então os valores de L_1^2 e L_2^2 para o estimador β^* ;

16-) Determinaram-se então as matrizes $Z = XV$, $\Lambda = V'X'XV$ e o vetor $\alpha = V'\beta$;

17-) Obtiveram-se a seguir as matrizes $X^* = [Y, X]$ e $X^{*'}X^*$.

Posteriormente determinaram-se os auto-valores δ_j , e os auto-vetores correspondentes γ_j da matriz $X^{*'}X^*$, $j = 0, 1, \dots, p$, através dos quais decide-se que variáveis serão suprimidas do modelo $y = Z\alpha + e$;

18-) Uma vez determinado as variáveis a serem suprimidas, digamos $p-r$, obtiveram-se o vetor $\tilde{\alpha} = \Lambda_{(r)}^{-1} Z_{(r)}' y$, onde

$$Z_{(r)} = XV_{(r)} \text{ e } \Lambda_{(r)} = (\lambda_1, \lambda_2, \dots, \lambda_r).$$

19-) Obtiveram-se então a seguir o vetor $\tilde{\beta} = V_{(r)}\tilde{\alpha}$ e os valores de L_1^2 e de L_2^2 para o estimador $\tilde{\beta}$.