



Estimativas de efeitos genéticos em bezerros cruzados por diferentes modelos e métodos de estimação¹

Eduardo da Cruz Gouveia Pimentel², Sandra Aidar de Queiroz³, Roberto Carvalheiro⁴, Luiz Alberto Fries^{4,5}

¹ Parte da dissertação apresentada pelo primeiro autor à Unesp/Jaboticabal, para obtenção do título de Mestre em Zootecnia (Genética e Melhoramento Animal), financiada pela CAPES.

² Doutorando em Zootecnia – Unesp/Jaboticabal – SP. Bolsista CAPES.

³ DZ – MGA – Unesp/Jaboticabal – SP. Bolsista CNPq.

⁴ Gensys Consultores Associados S/S Ltda.

⁵ Lagoa da Serra Ltda.

RESUMO - Os objetivos neste trabalho foram comparar estimativas de parâmetros genéticos obtidas por meio de dois modelos – um contendo apenas efeitos aditivos e de dominância e outro que incluiu os efeitos aditivo-conjunto (complementaridade) e epistático – e testar alternativas de critérios objetivos para determinação do coeficiente lambda na aplicação da regressão de cumeieira. Os resultados obtidos revelaram que a escolha de um critério para determinação do coeficiente lambda em regressão de cumeieira depende não apenas do conjunto de dados e do modelo utilizado, mas, sobretudo, de um conhecimento prévio acerca do fenômeno estudado e do significado prático e da interpretação dos parâmetros encontrados. Pelo uso de modelos mais completos para avaliação de efeitos genéticos em bovinos de corte, pode-se identificar a contribuição dos efeitos aditivo-conjunto e epistático, que encontram-se embutidos no efeito de heterose estimado por modelos mais simples. A regressão de cumeieira é uma ferramenta que viabiliza a obtenção dessas estimativas mesmo na presença de forte multicolinearidade.

Palavras-chave: epistasia, estimação viesada, heterose, multicolinearidade, regressão de cumeieira

Estimates of genetic effects in crossbred calves by different models and estimation methods

ABSTRACT - The purpose of this study was to compare estimates of genetic effects obtained using the additive-dominance model and another which included parameters for joint-additive (complementarity) and epistatic effects, as well as evaluate alternative objective criteria for choosing the lambda coefficient in ridge regression implementation. The results indicated that the criterion to be employed at the choice of lambda not only depends on the data set and the model used, but also on a previous knowledge about the phenomenon under study and the practical interpretation of estimated coefficients. When performing genetic effects evaluation, if other than additive and dominance effects are contemplated, it may be possible to identify and separate joint-additive and epistatic effects, which are usually inlaid in the heterotic effect estimated by the additive-dominance model. The use of ridge regression method can make such analyses possible even under strong multicollinearity.

Key Words: biased estimation, epistasis, heterosis, multicollinearity, ridge regression

Introdução

Sistemas de cruzamentos são fundamentados na exploração de efeitos raciais e de heterose. O uso eficiente dos efeitos raciais se baseia no conhecimento das características raciais diretas e maternas. Adicionalmente, o uso efetivo da heterose requer a caracterização dos efeitos de heterose nas vacas e nos bezerros cruzados (Roberson et al., 1986).

A heterose e seus efeitos, geralmente benéficos ao desempenho, têm sido atribuídos aos efeitos de dominância de genes de ação melhoradora. Em muitas investigações, essa hipótese mostra-se satisfatória na explicação da supe-

rioridade observada em grupos de indivíduos cruzados, como reportado por Gregory et al. (1991). Contudo, segundo Sharma et al. (2000), em alguns experimentos e investigações em que indivíduos da geração F₂ ou sintéticos são estudados, a hipótese da dominância tem sido insuficiente para ajustar os resultados. Certas combinações de alelos em diferentes locos interagem para determinar fenótipos extremamente diferentes dos que seriam preditos considerando-se os efeitos de cada alelo isoladamente, constituindo a ação epistática (Phillips, 1998).

De acordo com Wright (1978), ao longo do processo de formação das raças, a seleção natural deve, de algum

modo, operar sobre sistemas interativos de genes, como blocos, produzindo combinações mais eficientes. A quebra de associações epistáticas favoráveis entre genes de origem parental parece ser a principal causa do declínio no desempenho de características de crescimento em gerações avançadas de cruzamentos *Bos taurus* x *Bos indicus*. Logo, a epistasia parece ser um efeito genético importante que não deve ser ignorado em avaliações de desempenho de cruzamentos, particularmente quando a finalidade do exercício é a predição de desempenho de genótipos não testados, para decisões em futuros programas de cruzamentos (Demeke et al., 2003).

Segundo Dickerson (1973), a perda dessa superioridade epistática observada nas raças puras advém da recombinação nos gametas produzidos por pais cruzados. Em gerações avançadas de animais cruzados, portanto, todas as formas de ação e interação gênica e genotípica se manifestam e o balanço líquido entre efeitos de dominância e de epistasia é o que, normalmente, acaba sendo estimado e interpretado como heterose (Fries et al., 2000a).

Quando se consideram características de crescimento de animais cruzados em ambiente tropical, imagina-se que dois conjuntos de genes estejam atuando: um relacionado ao potencial de crescimento e outro à adaptação. Esses dois conjuntos podem ser tomados como características que se complementam e cujo produto (interação) forneceria uma base teórica para a heterose que não depende de dominância ou epistasia. Essa relação multiplicativa entre as ações aditivas para as duas características foi definida por Kinghorn (1993) como “profit heterosis” e pode ser interpretada como um efeito de complementariedade entre as características que determinam crescimento.

O melhor conhecimento da magnitude de influência de outros tipos de efeitos genotípicos, além da dominância, daria subsídios para predição mais consistente dos valores esperados para o desempenho de populações cruzadas. O método mais aplicado para derivar equações de predição é o dos quadrados mínimos. Contudo, quando existem fortes relações lineares entre as variáveis independentes (multicolinearidade), a simples estimativa por quadrados mínimos dos coeficientes de regressão individuais tende a ser instável, geralmente com grande erro-padrão, e pode provocar inferências errôneas (Bergmann & Hohenboken, 1995). Kinghorn & Vercoe (1989), Cassady et al. (2002) e Roso et al. (2005a) reportaram problemas de multicolinearidade quando efeitos epistáticos foram adicionados ao modelo aditivo-dominante.

Existem métodos alternativos de estimação que fornecem uma análise mais informativa dos dados quando existe

a multicolinearidade (Chatterjee & Price, 1991; Draper & Smith, 1998). Um desses métodos é a regressão de cumeieira (Hoerl & Kennard, 1970), que consiste na adição de um coeficiente lambda à diagonal principal da matriz de correlações, visando quebrar as dependências lineares. Quanto mais alto o valor de lambda, menores as correlações efetivas entre as variáveis independentes, porém, maior o viés das estimativas. O problema na regressão de cumeieira é que o valor de lambda é definido por meio de argumentos relativamente subjetivos (Freund & Littell, 2000).

Este trabalho foi realizado com os objetivos de comparar estimativas de parâmetros genéticos obtidas por meio de dois modelos – um contendo apenas efeitos aditivos e de dominância (Fisher, 1918) e outro que incluiu também os efeitos aditivo-conjunto (complementariedade) e epistático – e testar alternativas de critérios objetivos para determinação do coeficiente lambda na aplicação da regressão de cumeieira.

Material e Métodos

Foram utilizados dados de um programa de melhoria de gado de corte (Conexão Delta G), contendo registros de 109.614 bezerros puros e cruzados Hereford x Nelore nascidos entre 1974 e 1999, distribuídos em 4.665 grupos contemporâneos, e criados em 29 fazendas localizadas nos estados de Mato Grosso, Goiás, Mato Grosso do Sul, São Paulo, Paraná e Rio Grande do Sul.

Agrupando-se os genótipos presentes no conjunto de dados em oito classes (0 a 0,125; 0,125 a 0,25; 0,25 a 0,375; e assim por diante, em que 0 corresponde a puro Hereford e 1 a puro Nelore), obtiveram-se as seguintes frequências de observações: 51.308; 658; 9.801; 14.894; 20.427; 2.710; 3.578 e 6.238, respectivamente.

Dois modelos foram empregados: o modelo 1 relacionava o ganho de peso médio diário pré-desmama (em kg/dia) a efeitos de ações aditivas direta (AA) e materna (AM) e heteroses direta (HA) e materna (HM). O modelo 2 incluiu também os efeitos aditivo-conjunto direto (AAZ) e materno (AMZ) e epistático direto (EA) e materno (EM). Os efeitos ambientais incluídos em ambos os modelos foram: grupo contemporâneo; efeitos linear e quadrático de idade da vaca de acordo com o sexo do bezerro; efeitos linear e quadrático de idade à desmama; e polinômio segmentado quadrático-quadrático de data juliana de nascimento. Os grupos contemporâneos foram definidos como os bezerros do mesmo sexo; nascidos nos mesmos ano, fazenda e estação; criados no mesmo grupo de manejo (retiro); e desmamados na mesma data juliana.

Coefficientes para AA e AM foram definidos pela contribuição dos genes Nelore na composição genética de cada indivíduo. Coeficientes para AAZ e AMZ entre Nelore e Hereford foram calculados como $AAZ=AA*(1-AA)$ e $AMZ=AM*(1-AM)$, respectivamente (Piccoli et al., 2002; Brito et al., 2002).

Quando se considera gado de corte cruzado, envolvendo raças puras *Bos taurus* e *Bos indicus* e várias composições intermediárias, é uma grande simplificação pressupor que a relação entre o desempenho em determinada característica e composição genética será linear na amplitude inteira entre 0,0 e 1,0 (digamos que este valor indique a contribuição de zebuínos para o genótipo do animal). É possível hipotetizar a ocorrência de um efeito aditivo quadrático (como um complemento ao efeito aditivo linear) sempre que um ótimo intermediário seja vislumbrado. Em um ambiente temperado, o efeito de alterar a composição genética de 0,0 para 0,25 *Bos indicus* não será, necessariamente, equivalente ao passar de 0,75 a 1,00.

Funções quadráticas de efeitos aditivos são geralmente entendidas como efeitos epistáticos (interação não-alélica) ou de dominância (interação entre alelos). Um problema extra causado pela adição de um termo quadrático seria a alta correlação entre as colunas correspondentes a AA e $AA*AA$ quando AA varia entre 0 e 1. Kinghorn (1993) afirma que um modelo que inclua o produto de genes para potencial de crescimento e genes para adaptação ao ambiente explicaria melhor o crescimento de gado de corte em ambiente tropical. As funções definidas aqui como $AA*(1-AA)$ são usadas para estimar o efeito conjunto da fração *indicus* do genótipo (AA) e da fração *taurus* do genótipo (1-AA). Essa forma de estimar o efeito aditivo-conjunto ajuda a minimizar colinearidades.

A fórmula usada para calcular a heterozigose zebu-aurino foi a mesma usada por Bértoli (1991) e Schenkel (1993) para o caso específico de uma raça taurina e uma zebuína:

$$H = \text{raça do pai} + \text{raça da mãe} - 2*(\text{raça do pai}*\text{raça da mãe})$$

em que raça = fração zebuína do indivíduo.

O valor de H é uma medida relativa à máxima heterozigose que se pode alcançar (F_1) e esse máximo é designado 100% heterozigoto. Para o cálculo das heterozigosas direta e materna, foram necessárias as composições genéticas dos pais e avós maternos dos indivíduos. Quando não se dispunha de informação sobre a composição genética dos ancestrais maternos, foi assumido que o animal havia sido produzido por acasalamento *inter se*.

Os efeitos de EA e EM foram modelados por epistazigose, como em Fries et al. (2000a) e Roso et al.

(2005b), e calculados como a heterozigose média presente nos gametas que geraram cada indivíduo ou como a heterozigose média nos pais de um indivíduo. Segundo esses autores, a verificação de mais heterozigose em uma geração coincide com maior frequência de configurações recombinantes nos gametas e, conseqüentemente, com maior efeito epistático na geração seguinte.

A verificação da existência de multicolinearidade foi feita por meio do exame das correlações entre as variáveis independentes e dos fatores de inflação de variância (FIV), segundo recomendações de Chatterjee & Price (1991) e Freund & Littell (2000), respectivamente. Outro procedimento para diagnóstico de multicolinearidade consistiu do exame dos autovalores da matriz de correlações e do número de condição, que, para matrizes positivas definidas, pode ser calculado como a raiz quadrada da razão entre o maior e o menor autovalor.

Na implementação da regressão de cumeieira, três critérios alternativos foram testados objetivando definir o valor mais apropriado de lambda: a soma dos FIVs; a soma dos valores absolutos de t; e a média harmônica dos valores absolutos de t. O uso da soma dos FIVs baseia-se no fato de que, como indicativo da magnitude das correlações entre as variáveis, ele pode ser tão elucidativo no diagnóstico de multicolinearidade quanto na constatação de sua ausência.

A estatística t é o resultado da divisão da estimativa do parâmetro pelo respectivo erro-padrão. Segundo Fries et al. (2000b), a soma dos valores absolutos de t é uma estatística pragmática e pode ser mais informativa que o simples exame do comportamento das estimativas, uma vez que há um argumento objetivo apoiando a decisão. Os valores dos coeficientes de regressão linear tendem a diminuir com o aumento de lambda, fazendo com que o ponto onde se verifica a maior soma dos valores absolutos de t revele a maior redução nos erros-padrão das estimativas.

A média harmônica, recíproco da média aritmética dos recíprocos de uma série de números maiores que zero, tende a ser menor que a média aritmética desses números quanto maior a variância entre eles. Como se aproxima do menor valor contido na série, a média harmônica fornece melhor conhecimento sobre a dispersão observada para os valores de t. Quanto maior a média harmônica, maior a uniformidade na distribuição dos valores de t e, conseqüentemente, maior a frequência de estimativas significativas.

Para testar se a inclusão de parâmetros para os efeitos epistático e aditivo-conjunto efetivamente melhorou o modelo, em termos de ajuste, um teste F foi computado. O valor de F, conforme descrito por Weisberg (1980), foi calculado pela seguinte fórmula:

$$F = \frac{(SQR_{AD} - SQR_{ADEC}) / (gl_{AD} - gl_{ADEC})}{SQR_{ADEC} / gl_{ADEC}}$$

em que: SQR_{AD} e SQR_{ADEC} são as somas de quadrados dos resíduos dos modelos 1 e 2, respectivamente; e gl_{AD} e gl_{ADEC} são os graus de liberdade dos resíduos dos modelos 1 e 2, respectivamente.

O teste F fornece evidência contra o modelo mais simples (1) se o valor de F é maior que o valor tabelado para uma distribuição $F(gl_{AD} - gl_{ADEC}, gl_{ADEC}, \alpha)$. Outro método empregado para verificar se a exclusão de alguma variável explicativa comprometia significativamente o ajuste do modelo foi o cálculo das estatísticas $C(p)$ de Mallows (1973).

As estimativas dos parâmetros para o modelo 1 foram obtidas pelo método dos quadrados mínimos (MQM). Para o modelo 2, que continha os efeitos epistático e aditivo-conjunto, considerando a constatação da existência de multicolinearidade, optou-se por estimar os parâmetros por regressão de cumeira (RC). Todas as análises foram conduzidas utilizando-se o procedimento REG do programa SAS (SAS, 1996).

Resultados e Discussão

O valor de F (20,44) no teste conduzido para verificar a significância da diferença de ajuste entre os dois modelos foi altamente significativo ($P < 0,01$), indicando superioridade do modelo 2 em relação ao 1, em termos de ajuste.

Conforme descrito por Freund & Littell (2000), a estatística $C(p)$ de Mallows, uma medida da variância do erro mais o viés introduzido pela exclusão de uma variável do modelo, é calculada como segue:

$$C(p) = [SQR(p)/QME] - (N - 2p) + 1$$

em que QME é o quadrado médio do erro do modelo completo; $SQR(p)$, a soma de quadrados de um modelo

contendo um subconjunto de p variáveis explicativas; e N , o número de observações. Quando se observa que $C(p) > (p+1)$ para um modelo contendo p variáveis explicativas, existe evidência de viés, em razão da exclusão de uma variável importante do modelo.

Os valores de $C(p)$ observados para os modelos contendo subconjuntos com 16 variáveis (1 a menos que o modelo completo) foram todos superiores a 20, indicando que a exclusão de apenas uma variável provocaria falta de ajuste.

Calculados os coeficientes de correlação simples entre as covariáveis contempladas no modelo 2, constatou-se que 86,7% deles diferiram de zero ($P < 0,05$), sendo 12,5% superiores a 0,50. As estimativas das correlações entre as oito covariáveis genéticas (Tabela 1) foram todas estatisticamente significativas ($P < 0,01$).

Ressalta-se que muitas relações lineares observadas na matriz de coeficientes podem não advir de correlações entre pares de variáveis explicativas, mas de associações entre três ou mais delas. Dos fatores de inflação de variância encontrados para as 17 covariáveis (nove ambientais e oito genéticas), 13 foram superiores a 15, dos quais quatro foram maiores que 100. Por ser uma função do coeficiente de correlação múltipla entre as variáveis explicativas, o FIV pode representar um parâmetro mais informativo sobre a presença de multicolinearidade. Verificou-se ainda que oito autovalores foram próximos de zero e que o número de condição foi de 60,44.

Definição do valor de lambda a ser empregado na regressão de cumeira

Os coeficientes negativos para os efeitos heterótico materno e aditivo-conjunto direto e positivo para epistasia direta revelados pela análise por quadrados mínimos ordinários reforçaram ainda mais as evidências de alto grau de multicolinearidade. A obtenção dessas estimativas também pode explicar, de certa forma, porque são poucos os relatos do uso de modelos mais completos para estimação de

Tabela 1 - Estimativas dos coeficientes de correlação entre as covariáveis genéticas*

Table 1 - Correlation coefficient estimates between genetic covariables*

	AA	AM	AAZ	AMZ	HA	HM	EA	EM
AA	1,00000							
AM	0,34053	1,00000						
AAZ	-0,42190	-0,16820	1,00000					
AMZ	-0,20903	-0,17446	0,33979	1,00000				
HA	-0,39159	-0,03642	0,88894	0,02385	1,00000			
HM	-0,15008	-0,10538	0,26185	0,80746	0,01300	1,00000		
EA	-0,02444	-0,16609	0,35685	0,59128	-0,01054	0,70577	1,00000	
EM	-0,18123	-0,18183	0,27327	0,76106	0,02513	0,26947	0,22589	1,00000

* AA, AM, HA, HM, AAZ, AMZ, EA e EM correspondem aos efeitos genéticos diretos e maternos de ação aditiva, heterose, efeito aditivo-conjunto e epistasia, respectivamente.

* AA, AM, HA, HM, AAZ, AMZ, EA and EM correspond to direct and maternal additive, heterotic, joint-additive and epistatic effects, respectively.

parâmetros de epistasia e de outros efeitos genéticos. Nesses casos, a regressão de cumeeira pode representar uma ferramenta que viabilize a obtenção de estimativas mais coerentes com o fenômeno em estudo.

Pela observação do gráfico dos valores das estimativas dos parâmetros contra valores de lambda variando de 0 a 1, verificou-se que as maiores alterações nos valores das estimativas ocorreram no intervalo de 0 a 0,5, ponto a partir do qual as estimativas tenderam a se estabilizar. Fenômeno semelhante pôde ser observado no comportamento dos valores de t para as estimativas, no mesmo intervalo (Figuras 1a e 1b). Gráficos dos valores do coeficiente de determinação (R^2), do quadrado médio do erro e do traço da matriz de projeção ($X(X'X)^{-1}X'$) foram construídos no intervalo de 0 a 1, porém, todas essas estatísticas se mostraram monotônicas em lambda.

Gráficos dos valores da soma dos FIVs, da média harmônica e da soma dos valores absolutos de t foram, então, construídos contra valores de lambda variando de 0 a 1 (Figuras 1c, 1d e 1e). Avaliando como critério para determinação de lambda a soma dos valores absolutos de t,

encontrou-se para essa estatística valor máximo de 87,23 para lambda igual a 0,76. Para a média harmônica, o valor máximo de 7,49 foi determinado para lambda igual a 0,57. As somas dos FIVs decresceram com o aumento de lambda, sendo que 94,38% da redução foi observada entre lambda 0 e 0,1. A observação das mudanças nos FIVs nesse intervalo de 0 a 0,1 revelou que, para lambda igual a 0,06, todos os FIVs já eram menores que dois.

Pela comparação entre os valores das estimativas dos parâmetros encontrados pelo MQM (lambda = 0) e por RC com definição de lambda pelos três critérios (Tabela 2), nota-se que, para lambda igual a 0,06, a estimativa do coeficiente para efeito aditivo-conjunto direto permaneceu negativa. Constata-se também que os valores observados para lambda igual a 0,57 e 0,76 foram muito próximos, indicando que, para este conjunto de dados, as estimativas dos parâmetros não seriam muito diferentes pela adoção da soma ou da média harmônica dos valores absolutos de t como critério para determinação de lambda. Pôde-se observar neste estudo que o procedimento da regressão de cumeeira atuou em dois estágios: na quebra

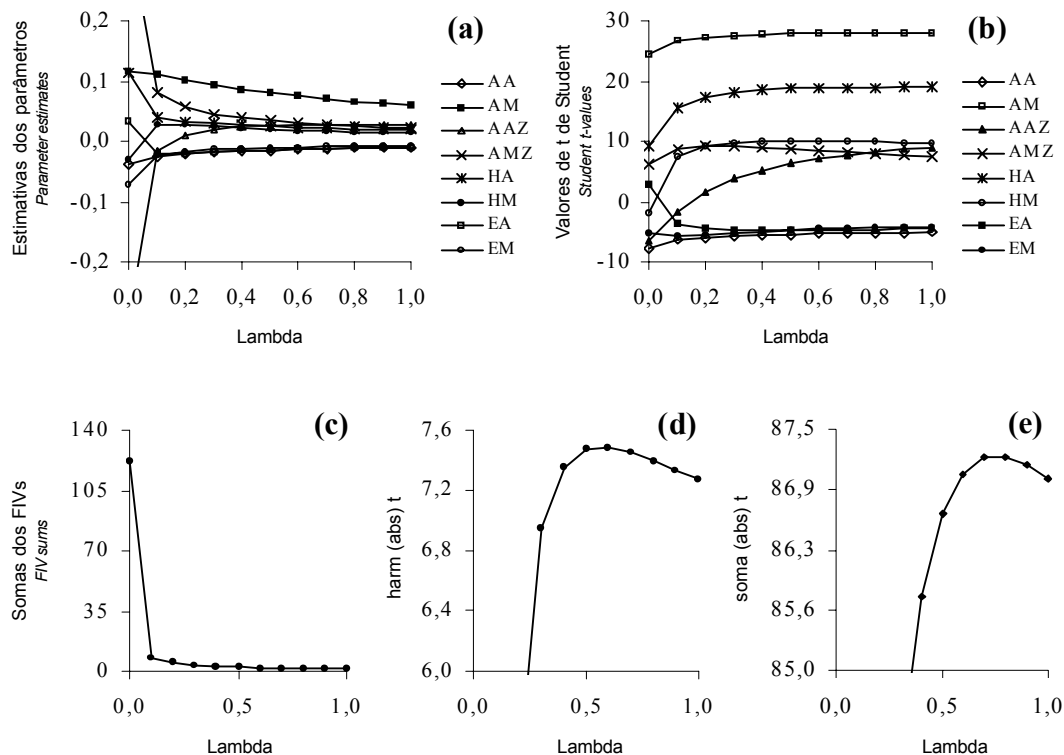


Figura 1 - Alterações nas estimativas dos parâmetros (a), nos respectivos valores de t (b), na soma dos FIVs (c), na média harmônica (d) e na soma dos valores absolutos de t (e) para valores do coeficiente lambda variando de 0 a 1. AA, AM, HA, HM, AAZ, AMZ, EA e EM correspondem aos efeitos genéticos diretos e maternos de ação aditiva, heterose, efeito aditivo-conjunto e epistasia, respectivamente.

Figure 1 - Changes on parameter estimates (a), respective Student t-values (b), sum of variance inflation factors - FIVs (c), harmonic mean (d) and sum of absolute t-values (e) for lambda values ranging from 0 to 1. AA, AM, HA, HM, AAZ, AMZ, EA and EM correspond to direct and maternal additive, heterotic, joint-additive and epistatic effects, respectively.

da multicolinearidade, com valores de lambda inferiores a 0,1; e na mudança efetiva nos valores das estimativas para o intervalo de 0,1 a 1.

Segundo Marquardt (1970), estimadores de regressão de cumeieira podem ser vistos como uma espécie de média ponderada entre a informação contida nos dados e outra informação (em termos Bayesianos, uma “informação *a priori*”) para as quais os valores atribuídos à variável resposta são arbitrariamente definidos como zero.

O procedimento conduzido para determinação do valor de lambda pode então ser interpretado, sob a ótica Bayesiana, como a especificação de 5 ou 10 distribuições *a priori*, por exemplo, com expectativa nula e cobrindo uma amplitude plausível, e a comparação de várias regressões alternativas com as 5 ou 10 médias *a posteriori*. As duas abordagens oferecem certo esclarecimento sobre o problema posto por um conjunto de dados: se a análise correta depende do modelo e da informação externa adotada, o pesquisador não deve esperar que algum procedimento específico seja automaticamente aplicável (Dempster et al., 1977).

Comparações entre as estimativas obtidas com os dois modelos

Os resultados das análises pelos dois modelos são descritos na Tabela 3. Os valores das probabilidades associadas aos testes t de Student, sob hipótese de nulidade, foram significativos ($P < 0,05$) para todas as estimativas apresentadas na tabela. O maior valor ($P = 0,0438$) foi encontrado para HM estimada pelo modelo 2 por MQM, sendo todos os outros menores que 0,01. Ainda pelo método MQM, observou-se alteração no comportamento (troca de sinal) da estimativa de heterose materna (HM) obtida com o modelo 2. Além disso, sinal negativo para efeito aditivo-conjunto direto (AAZ) e positivo para epistasia direta (EA) foram observados. Especulou-se que esses resultados estariam associados ao problema de multicolinearidade decorrente da inclusão de maior número de covariáveis no modelo. Espera-se que o efeito de heterose materna seja um dos mais atuantes na determinação do ganho de peso do nascimento à desmama, de modo que a verificação de que o maior erro-padrão relativo foi obtido para a estimativa de

Tabela 2 - Estimativas dos coeficientes (erros padrão) dos efeitos genéticos* obtidos por regressão de cumeieira com valores de lambda variando de 0** a 1

Table 2 - Coefficient estimates (standard errors) of genetic effects* obtained by ridge regression with lambda values ranging from 0** to 1

Lambda	AA	AM	AAZ	AMZ	HA	HM	EA	EM
0,00	-0,03694 (0,00481)	0,11699 (0,00478)	-0,29742 (0,04619)	0,34847 (0,05706)	0,11266 (0,01215)	-0,03102 (0,01538)	0,03300 (0,01148)	-0,07061 (0,01377)
0,06	-0,02743 (0,00434)	0,11520 (0,00436)	-0,04348 (0,01249)	0,10332 (0,01296)	0,04810 (0,00348)	0,02639 (0,00467)	-0,01727 (0,00615)	-0,02617 (0,00452)
0,10	-0,02500 (0,00409)	0,11126 (0,00416)	-0,01584 (0,00908)	0,08075 (0,00927)	0,04071 (0,00262)	0,02848 (0,00384)	-0,02026 (0,00546)	-0,02173 (0,00378)
0,20	-0,02098 (0,00359)	0,10181 (0,00373)	0,00990 (0,00616)	0,05760 (0,00616)	0,03289 (0,00188)	0,02728 (0,00299)	-0,02007 (0,00448)	-0,01661 (0,00308)
0,30	-0,01821 (0,00321)	0,09365 (0,00340)	0,01945 (0,00511)	0,04678 (0,00503)	0,02912 (0,00160)	0,02484 (0,00256)	-0,01814 (0,00387)	-0,01390 (0,00273)
0,40	-0,01611 (0,00292)	0,08665 (0,00312)	0,02408 (0,00454)	0,03998 (0,00441)	0,02663 (0,00144)	0,02256 (0,00226)	-0,01627 (0,00343)	-0,01208 (0,00249)
0,50	-0,01445 (0,00268)	0,08062 (0,00289)	0,02654 (0,00415)	0,03513 (0,00401)	0,02475 (0,00132)	0,02060 (0,00205)	-0,01467 (0,00310)	-0,01073 (0,00230)
0,57	-0,01349 (0,00253)	0,07687 (0,00275)	0,02755 (0,00394)	0,03245 (0,00380)	0,02366 (0,00126)	0,01939 (0,00193)	-0,01370 (0,00291)	-0,00997 (0,00219)
0,60	-0,01311 (0,00248)	0,07537 (0,00270)	0,02786 (0,00387)	0,03143 (0,00372)	0,02324 (0,00123)	0,01891 (0,00188)	-0,01333 (0,00284)	-0,00968 (0,00214)
0,70	-0,01201 (0,00231)	0,07076 (0,00253)	0,02852 (0,00364)	0,02849 (0,00349)	0,02196 (0,00116)	0,01747 (0,00174)	-0,01219 (0,00263)	-0,00882 (0,00201)
0,76	-0,01144 (0,00222)	0,06826 (0,00244)	0,02871 (0,00352)	0,02699 (0,00337)	0,02128 (0,00112)	0,01669 (0,00167)	-0,01160 (0,00251)	-0,00838 (0,00194)
0,80	-0,01109 (0,00216)	0,06669 (0,00238)	0,02877 (0,00345)	0,02607 (0,00330)	0,02086 (0,00110)	0,01622 (0,00163)	-0,01124 (0,00245)	-0,00811 (0,00189)
0,90	-0,01030 (0,00203)	0,06306 (0,00225)	0,02876 (0,00328)	0,02405 (0,00314)	0,01989 (0,00105)	0,01513 (0,00153)	-0,01042 (0,00229)	-0,00751 (0,00179)
1,00	-0,00963 (0,00192)	0,05981 (0,00213)	0,02858 (0,00314)	0,02233 (0,00300)	0,01902 (0,00100)	0,01417 (0,00145)	-0,00971 (0,00216)	-0,00700 (0,00170)

* AA, AM, HA, HM, AAZ, AMZ, EA e EM correspondem aos efeitos genéticos diretos e maternos de ação aditiva, heterose, efeito aditivo-conjunto e epistasia, respectivamente.

** Lambda=0 equivale à implementação do método de quadrados mínimos.

*** AA, AM, HA, HM, AAZ, AMZ, EA and EM correspond to direct and maternal additive, heterotic, joint-additive and epistatic effects, respectively.

**** Lambda=0 corresponds to the implementation of the ordinary least squares method.

Tabela 3 - Estimativas dos parâmetros* (erros-padrão) pelos modelos 1 e 2, por quadrados mínimos ordinários (MQM) e por regressão de cumeieira (RC) com lambda 0,57, respectivamente

Table 3 - Parameter estimates* (standard errors) obtained with models 1 and 2, by ordinary least squares (MQM) and ridge regression (RC) with lambda 0.57, respectively

	AA	AM	AAZ	AMZ	HA	HM	EA	EM
Modelo 1	-0,03513	0,12719			0,03631	0,04738		
	(0,00466)	(0,00460)			(0,00290)	(0,00370)		
Modelo 2	-0,03694	0,11699	-0,29742	0,34847	0,11266	-0,03102	0,03300	-0,07061
	(0,00481)	(0,00478)	(0,04619)	(0,05706)	(0,01215)	(0,01538)	(0,01148)	(0,01377)
RC	-0,01349	0,07687	0,02755	0,03245	0,02366	0,01939	-0,01370	-0,00997
	(0,00253)	(0,00275)	(0,00394)	(0,00380)	(0,00126)	(0,00193)	(0,00291)	(0,00219)

* AA, AM, HA, HM, AAZ, AMZ, EA e EM correspondem aos efeitos genéticos diretos e maternos de ação aditiva, heterose, efeito aditivo-conjunto e epistasia, respectivamente.

* AA, AM, HA, HM, AAZ, AMZ, EA and EM correspond to direct and maternal additive, heterotic, joint-additive and epistatic effects, respectively.

HM reforça ainda mais essa idéia. Conforme mencionado anteriormente, a estratégia adotada procurando contornar este problema foi a adoção do método de regressão de cumeieira. Os valores das estimativas dos parâmetros do modelo 2 foram determinados pela atribuição do valor 0,57 para o coeficiente lambda, ponto em que o maior valor para a média harmônica dos valores absolutos de t foi observado.

Observou-se que as estimativas de heterose obtidas pelo modelo 1 foram superiores às encontradas quando se consideraram os efeitos aditivo-conjunto e epistático, no modelo 2, pelo método RC. Os coeficientes estimados para HA (0,03631) e HM (0,04738) pelo modelo 1 se aproximam numericamente do obtido pela adição dos coeficientes estimados pelo modelo 2, por regressão de cumeieira, para HA, AAZ e EA (0,03751), e HM, AMZ e EM (0,04187), respectivamente. Isso é um indicativo de que os parâmetros para heterose são superestimados quando calculados por modelos mais simples, fato que advém da diferença em magnitude entre o coeficiente positivo (para efeito aditivo-conjunto) e o negativo (para epistasia). Fries et al. (2000a) compararam o uso de três modelos na predição do peso à desmama de bezerros Hereford x Brahman: o modelo chamado "Benchmark", contendo covariáveis para AA, AM, HA e HM; o modelo completo, incluindo covariáveis para AAZ, AMZ, EA e EM; e o modelo genotípico, incluindo efeitos para os 15 genótipos existentes na população, representando 100% de ajuste. O primeiro modelo explicou 75,7% da variância entre os genótipos e o segundo, 90,1%. Nesse caso, aproximadamente 60% da falta de ajuste do modelo mais simples pôde ser reduzida pela inclusão das covariáveis para efeito aditivo-conjunto e de epistasia.

Neste estudo, as estimativas dos coeficientes de AA (-0,03513) e AM (0,12719) para o modelo 1 foram muito próximas das obtidas por quadrados mínimos ordinários pelo modelo 2 (-0,03694 AA e 0,11699 AM), o que leva a crer que essas covariáveis não estavam tão fortemente envolvidas no problema de multicolinearidade. Quando comparadas às estimativas obtidas por regressão de cumeieira (-0,01349 AA e 0,07687 AM), a diferença aumentou consideravelmente. Esse resultado indica que a aplicação do mesmo lambda a todas as covariáveis do modelo impõe um viés desnecessário àquelas que não fazem parte da condição de multicolinearidade, evidenciando a necessidade de um método que permita a adição de coeficientes lambda condizentes com o envolvimento das respectivas covariáveis nas relações lineares entre as variáveis independentes.

A implementação da regressão de cumeieira permitiu a obtenção de estimativas biologicamente mais coerentes no modelo 2, sendo observados sinais positivos para efeito aditivo-conjunto (AAZ e AMZ) e heterose (HA e HM) e negativos para epistasia (EA e EM) sobre a característica estudada (Tabela 3).

Conclusões

A escolha de um critério para determinação do coeficiente lambda em regressão de cumeieira depende não só do conjunto de dados e do modelo utilizado, mas, sobretudo, de um conhecimento prévio acerca do fenômeno estudado e do significado prático e da interpretação dos parâmetros encontrados.

O uso de modelos mais completos para avaliação de efeitos genéticos em bovinos de corte permite identificar a contribuição dos efeitos aditivo-conjunto e epistático, que

parecem estar embutidos no efeito de heterose estimado por modelos mais simples.

A regressão de cumeira é uma ferramenta que viabiliza a obtenção de estimativas dos efeitos aditivo-conjunto e epistático mesmo na presença de forte multicolinearidade.

Os coeficientes lambda empregados na implementação da regressão de cumeira devem ser adicionados apenas às covariáveis envolvidas em problemas de multicolinearidade.

Agradecimento

À Conexão Delta G e à Gensys Consultores Associados S/S Ltda, pelo fornecimento e pela preparação dos dados utilizados neste trabalho. À CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, pela concessão da bolsa de estudos.

Literatura Citada

- BERGMANN, J.A.G.; HOHENBOKEN, W.D. Alternatives to least squares in multiple linear regression to predict production traits. **Journal of Animal Breeding and Genetics**, v.112, p.1-16, 1995.
- BÉRTOLI, C.D. **Sistema Cruza - Controle de produção e avaliação dos valores genéticos dentro de uma população bovina sintética**. Porto Alegre: Universidade Federal do Rio Grande do Sul, 1991. 233p. Dissertação (Mestrado em Agronomia) - Universidade Federal do Rio Grande do Sul, 1991.
- BRITO, F.V.; PICCOLI, M.L.; SEVERO, J.L.P. et al. Estimating environmental and genotypic effects on preweaning weight gain of Angus x Nelore calves. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier. **Proceedings...** Montpellier: Organizing Committee, 2002. p.773-776.
- CASSADY, J.P.; YOUNG, L.D.; LEYMASTER, K.A. Heterosis and recombination effects on pig growth and carcass traits. **Journal of Animal Science**, v.80, p.2286-2302, 2002.
- CHATTERJEE, S.; PRICE, B. **Regression analysis by example**. 2.ed. New York: Wiley series in probability and mathematical statistics, 1991. 278p.
- DEMEKE, S.; NESER, F.W.C.; SCHOEMAN, S.J. Early growth performance of *Bos taurus* x *Bos indicus* cattle crosses in Ethiopia: evaluation of different crossbreeding models. **Journal of Animal Breeding and Genetics**, v.120, p.39-50, 2003.
- DEMPSTER, A.P.; SCHATZOFF, M.; WERMUTH, N. A simulation study of alternatives to ordinary least squares. **Journal of the American Statistical Association**, v.72, p.77-91, 1977.
- DICKERSON, G.E. Inbreeding and heterosis in animals. In: ANIMAL BREEDING AND GENETICS SYMPOSIUM IN HONOR OF DR. JAY L. LUSH, 1973, Champaign. **Proceedings...** Champaign: American Society of Animal Science, 1973. p.54-77.
- DRAPER, N.R.; SMITH, H. **Applied regression analysis**. 3.ed. New York: Wiley series in probability and statistics, 1998. 706p.
- FISHER, R.A. The correlations between relatives on the supposition of Mendelian inheritance. **Transactions of the Royal Society of Edinburgh**, v.52, p.399-433, 1918.
- FREUND, R.J.; LITTELL, R.C. **SAS system for regression**. 3.ed. Cary: SAS Institute Inc., 2000. 235p.
- FRIES, L.A.; JOHNSTON, D.J.; HEARNshaw, H. et al. Evidence of epistatic effects on weaning weight in crossbred beef cattle. **Asian-Australasian Journal of Animal Sciences**, v.13, supl. B, p. 242, 2000a.
- FRIES, L.A.; GRASER, H.U.; JOHNSTON, D.J. et al. Using ridge regression to estimate genetic effects in crossbred beef cattle. **Asian-Australasian Journal of Animal Sciences**, v.13, supl. B, p. 241, 2000b.
- GREGORY, K.E.; CUNDIFF, L.V.; KOCH, R.M. Breed effects and heterosis in advanced generations of composite populations for growth traits in both sexes of beef cattle. **Journal of Animal Science**, v.69, p.3202-3212. 1991.
- HOERL, A.E.; KENNARD, R.W. Ridge regression: biased estimation for nonorthogonal problems. **Technometrics**, v.12, n.1, p.55-67, 1970.
- KINGHORN, B.P. Theory of breed utilisation. In: **Design of livestock breeding programs: short course in animal breeding**. Armidale: AGBU, 1993. p.187-204.
- KINGHORN, B.P.; VERCOE, P.E. The effects of using the wrong genetic model to predict the merit of crossbred genotypes. **Animal Production**, v.49, p.209-216, 1989.
- MALLOWS, C.P. Some comments on C(p). **Technometrics**, v.15, p.661-675, 1973.
- MARQUARDT, D.W. Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. **Technometrics**, v.12, p.591-612, 1970.
- PHILLIPS, P.C. The language of gene interaction. **Genetics**, v.149, p.1167-1171, 1998.
- PICCOLI, M.L.; ROSO, V.M.; BRITO, F.V. et al. Additive, complementarity (additive*additive), dominance, and epistatic effects on preweaning weight gain of Hereford x Nelore calves. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier. **Proceedings...** Montpellier: Organizing Committee, 2002. p.275-278.
- ROBERSON, R.L.; SANDERS, J.O.; CARTWRIGHT, T.C. Direct and maternal genetic effects on preweaning characters of Brahman, Hereford and Brahman-Hereford crossbred cattle. **Journal of Animal Science**, v.63, p.438-446, 1986.
- ROSO, V.M.; SCHENKEL, F.S.; MILLER, S.P. et al. Estimation of genetic effects in the presence of multicollinearity in multibreed beef cattle evaluation. **Journal of Animal Science**, v.83, p.1788-1800, 2005a.
- ROSO, V.M.; SCHENKEL, F.S.; MILLER, S.P. et al. Additive, dominance, and epistatic loss effects on preweaning weight gain of crossbred beef cattle from different *Bos taurus* breeds. **Journal of Animal Science**, v.83, p.1780-1787, 2005b.
- SCHENKEL, F.S. **Cálculo das heterozigoses**. Porto Alegre: Gensys Consultores Associados, 1993. 3p.
- SHARMA, B.S.; PRABHAKARAN, V.T.; PIRCHNER, F. Gene action and heterosis in lifetime traits of Friesian x Sahiwal crosses. **Journal of Animal Breeding and Genetics**, v.117, p.319-330, 2000.
- STATISTICAL ANALYSIS SYSTEM – SAS. **SAS/STAT. User's guide**. version 6.11, 4.ed, v.2. Cary: 1996.
- WEISBERG, S. **Applied linear regression**. 1.ed. New York: Wiley Series in probability and mathematical statistics, 1980. 283p.
- WRIGHT, S. The relation of livestock breeding to theories of evolution. **Journal of Animal Science**, v.46, n.5, p.1192-1200, 1978.

Recebido: 25/02/05

Aprovado: 01/12/05