

## MODELOS RIDGE EM PLANEJAMENTOS DE MISTURAS: UMA APLICAÇÃO NA EXTRAÇÃO DA POLPA DE PEQUI

Tania Miranda Nepomucena, Augusto Maciel da Silva e Marcelo Angelo Cirillo\*

Departamento de Ciências Exatas, Universidade Federal de Lavras, CP 3037, 37200-000 Lavras – MG, Brasil

Recebido em 23/2/12; aceito em 6/8/12; publicado na web em 29/11/12

RIDGE MODELS IN MIXTURE PLANNING: AN APPLICATION IN THE EXTRACTION OF PEQUI PULP. Mixture Models can be used in experimental situations involving areas related to food science and chemistry. Some problems of a statistical nature can be found, such as effects of multicollinearity that result in uncertainty in the optimization of a dependent variable. This study proposes the application of the ridge model adapted for mixture planning considering the Kronecker (K-model) and Scheffe (S-Model) methods applied to response surfaces. The method determined the proportions of hexane, acetone and alcohol proportions that resulted in the maximum response of percentage of extracted pequi (*Caryocar brasiliense*) pulp oil.

Keywords: Ridge analysis; mixtures; Kronecker model.

### INTRODUÇÃO

O propósito geral em um planejamento de misturas é tornar possíveis as estimativas das propriedades de um sistema multicomponente, a partir de um número limitado de observações. Essas observações são obtidas de combinações pré-selecionadas dos componentes, na tentativa de determinar quais delas otimizam a resposta.<sup>1</sup> A teoria de experimentação e modelagem para planejamentos de misturas engloba delineamentos e regressão polinomial adequados às especificidades de ensaios envolvendo proporções dos  $q$  componentes.<sup>2-4</sup>

Em um delineamento que envolva planejamento de misturas, pode-se utilizar a técnica de análise Ridge para examinar o comportamento de superfícies de respostas quadráticas hiperdimensionais, com a finalidade de determinar a região ótima local e geral.<sup>5</sup>

Em se tratando da representação de um planejamento de mistura, define-se que uma mistura é composta por  $q$  ingredientes, na qual, geometricamente o espaço viável para a região experimental consiste de uma figura regular de dimensão  $q - 1$ , denominada *simplex*, conforme mostrado na Figura 1, onde os pontos correspondem a misturas puras, ou seja, de um único componente, misturas binárias ou de dois componentes, até consistir de misturas formadas por todos os ingredientes em quantidades proporcionais. O centroide do *simplex*

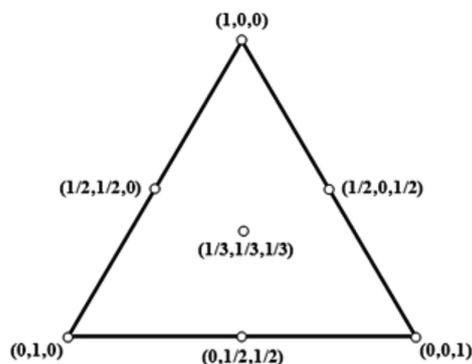


Figura 1. Distribuição dos pontos experimentais (misturas) em um simplex que garanta a restrição de que, para qualquer ponto, a soma dos três componentes deverá ser unitária

\*e-mail: macufla@dex.ufla.br

é o ponto que representa a mistura composta pelos  $q$  ingredientes distribuídos em mesma proporção.<sup>2-4</sup>

Mantendo essa representação, ressalta-se que as análises estatísticas desses experimentos devem ser feitas com cautela, uma vez que a ocorrência da multicolinearidade é sempre manifestada devido à restrição exigida, que as  $X_i$  proporções dos  $q$  componentes se distribuam de modo que  $\sum_{i=1}^q x_i = 1$ .<sup>6</sup> Usualmente, para que o efeito da multicolinearidade seja amenizado, o procedimento a ser realizado considera o fato de que a matriz de delineamento  $\mathbf{X}$  do modelo apresenta um vetor coluna que pode ser escrito como combinação linear dos demais, visto que  $x_1 + x_2 + \dots + x_q = 1$ , portanto, essa relação linear conduz a uma matriz  $\mathbf{X}'\mathbf{X}$  singular, que não fornece estimativas únicas para o vetor de parâmetros  $\beta$ .<sup>7</sup> Assim sendo, uma alternativa para remover a dependência entre as covariáveis  $x_i$  é verificada na substituição  $x_q = 1 - x_1 - x_2 - \dots - x_{q-1}$ .

A importância em se ajustar um modelo quadrático é verificada nos resultados específicos à interpretação da curvatura. Neste contexto, a observação do ponto estacionário, mormente caracteriza ponto de máximo ou mínimo. Todavia, para uma confirmação em relação a essa caracterização, todas as derivadas parciais do modelo em qualquer direção do ponto extremo deverão ser nulas. Desta forma, para que esse ponto represente um máximo, necessariamente todas as segundas derivadas deverão ser negativas; caso elas sejam todas positivas, tem-se um ponto de mínimo.

A situação crítica que exige cautela do pesquisador em relação à caracterização do ponto estacionário refere-se ao surgimento de um ponto de sela,<sup>8</sup> ocasionado pelo fato de que em determinadas direções ocorram derivadas positivas e em outras negativas. Em se tratando da presença de observações discrepantes, a verdadeira natureza do ponto estacionário poderá ser mascarada. Para uma melhor compreensão, exemplificamos a seguinte situação: considerando uma amostra com a presença de uma ou mais observações discrepantes, o ponto estacionário poderá ser interpretado como ponto de sela. Retirando essas observações e ajustando novamente o modelo, o ponto estacionário poderá ser identificado como máximo ou mínimo, dependendo das direções das derivadas parciais. Nesse contexto, há evidências de incertezas, relacionadas à presença ou ausência dessas observações, que contribuem para que o ponto estacionário seja localizado fora da região experimental a ser considerada na otimização.

Frente a esse problema, é desejável encontrar um valor ótimo

dentro dos limites da região contemplada no experimento. Com este propósito, utiliza-se a análise Ridge para obtenção deste valor sobre uma das circunferências proporcionadas por uma região esférica centrada na origem do experimento, aplicando-se o conceito de multiplicador de Lagrange.<sup>8</sup> Especificamente em modelagem de misturas, a análise Ridge recomenda que a restrição do somatório dos componentes seja igual ao valor unitário, sendo esta incorporada na função lagrangeana.<sup>9</sup>

Convém ressaltar que as aplicações de planejamentos de misturas são comuns na área de Química, podendo-se citar trabalhos já publicados como, por exemplo, o estudo das interações dos triglicérides de cadeia média, da fibra e do cálcio na dialisabilidade do ferro e do zinco na maximização de uma formulação de alimentação enteral utilizando planejamento de misturas da metodologia de superfície de resposta<sup>10</sup> e, também, o desenvolvimento e a validação de uma metodologia de análise de misturas biodiesel:diesel empregando cromatografia gasosa de alta resolução acoplada à espectrometria de massas no modo de análise por monitoramento seletivo de íons.<sup>11</sup>

Em virtude do que foi mencionado, o objetivo deste trabalho consistiu em ajustar modelos Ridge adaptados à metodologia de superfícies de respostas utilizando os métodos de Kronecker (*K-model*) e Scheffé (*S-model*) em um delineamento envolvendo misturas formadas pelas proporções de hexano, acetona e álcool e que possam ser recomendadas sobre as mesmas condições experimentais empregadas neste experimento para obtenção da resposta máxima do percentual de óleo extraído da polpa de Pequi (*Caryocar brasiliense*).

### Modelo Scheffé (*S-model*)

O conhecido modelo de segunda ordem de Scheffé (*S-model*), para o caso particular  $q = 3$ , é representado pela Equação 4

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \varepsilon, \quad (1)$$

que se constitui na função de regressão ajustada para dados coletados no *simplex*, cuja dedução é obtida por meio de algumas alterações na Equação 2,

$$\eta = \beta_0 + \sum_{i=1}^3 \beta_i x_i + \sum_{i < j}^3 \sum_{j} \beta_{ij} x_i x_j, \quad (2)$$

conforme os seguintes passos: <sup>6</sup> o parâmetro  $\beta_0$  é substituído por  $\beta_0(x_1 + x_2 + x_3)$ , já que  $x_1 + x_2 + x_3 = 1$ ; sendo a proporção  $x_1 = 1 - x_2 - x_3$ , pode-se obter  $x_1^2 = x_1 - x_1 x_2 - x_1 x_3$  e substituir no termo  $x_1^2$  do modelo (2). Repetir o mesmo processo para os outros dois termos quadráticos puros ( $x_2^2$  e  $x_3^2$ ) com substituições similares na Equação 2. Em seguida, aplicar as possíveis operações algébricas na equação do modelo e, finalmente, retornar à notação usual (renomeando os coeficientes  $\beta_{ij}^*$  obtidos por  $\beta_{ij}$ ,  $i, j = 1, 2, 3$ ).

### Modelo de Kronecker (*K-model*)

O modelo de Kronecker, obtido por meio de substituições realizadas diretamente na Equação 1, resume-se em outra função de regressão que pode ser utilizada para analisar dados de experimentos com misturas.<sup>12</sup> Para o caso específico, modelo quadrático e  $q = 3$ , o termo  $x_1$  é substituído por  $x_1(x_1 + x_2 + x_3) = x_1^2 + x_1 x_2 + x_1 x_3$  e usam-se operações semelhantes para  $x_2$  e  $x_3$ . Após mais uma vez renomear os coeficientes  $\beta_{ij}^*$  obtidos por  $\beta_{ij}$ , a forma resultante é apresentada na Equação 3, conhecida por modelo quadrático de Kronecker (*K-model*).

$$y = \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \varepsilon. \quad (3)$$

Convém ressaltar que em ambas as abordagens (*K-model*) e (*S-model*) os modelos quadráticos apresentam linearidade nos parâmetros, portanto, o uso do método de estimação de mínimos quadrados proporciona estimativas exatas. Em situações onde os modelos a serem ajustados apresentam não linearidade nos parâmetros como, por exemplo, curvas exponenciais, as estimativas dos parâmetros são aproximadas e obtidas pelos algoritmos Newton-Raphson ou Gauss-Newton.

## ANÁLISE RIDGE EM PLANEJAMENTOS DE MISTURAS CONSIDERANDO OS MODELOS DE KRONECKER E SCHEFFÉ

Na presença de multicolinearidade, as estimativas dos parâmetros estão associadas a variâncias inflacionadas, tornando-as não confiáveis. Um procedimento a ser adotado é encontrar um valor ótimo dentro dos limites da região experimental.

Com este enfoque, emprega-se a análise Ridge, com o propósito de examinar o comportamento de superfícies de respostas geradas por modelos quadráticos e localizar o ponto ótimo geral e local.<sup>13</sup> Além do mais, é possível determinar a região de otimização a qualquer distância de um ponto foco do delineamento<sup>5</sup> e, também, por meio de gráficos constatar os efeitos de todas as variáveis simultaneamente.<sup>14</sup>

A análise Ridge pode ser empregada diretamente em aplicações de mistura, quando acrescentadas as restrições da Equação 4,

$$\mathbf{x}'\mathbf{1} = x_1 + x_2 + \dots + x_q = 1 \quad \text{e} \quad \mathbf{x}'\mathbf{x} = x_1^2 + x_2^2 + \dots + x_q^2 = R^2 \quad (4)$$

que garantem uma solução no subespaço da mistura, localizada numa esfera de raio  $R$ .<sup>9</sup>

Para a obtenção da resposta ótima, utiliza-se a função lagrangeana apresentada na Equação 5, incorporando-se as restrições da Equação 4, sendo  $\lambda_1$  e  $\lambda_2$  conhecidos por multiplicadores de Lagrange,

$$F = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x} - \lambda_1 (\mathbf{x}'\mathbf{1} - 1) - \lambda_2 (\mathbf{x}'\mathbf{x} - R^2), \quad (5)$$

em que  $\mathbf{x}' = (x_1, x_2, \dots, x_q)$ ,  $\boldsymbol{\beta}' = (\beta_1, \beta_2, \dots, \beta_q)$  e

$$\mathbf{B} = \begin{bmatrix} \beta_{11} & \frac{1}{2}\beta_{12} & \dots & \frac{1}{2}\beta_{1q} \\ \frac{1}{2}\beta_{21} & \beta_{22} & \dots & \frac{1}{2}\beta_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}\beta_{q1} & \frac{1}{2}\beta_{q2} & \dots & \beta_{qq} \end{bmatrix} \quad (6)$$

Diferenciando-se a Equação 5 em relação a  $\mathbf{x}$ , obtém-se

$$\frac{\partial F}{\partial \mathbf{x}} = \boldsymbol{\beta} + 2\mathbf{B}\mathbf{x} - \lambda_1 \mathbf{1} - 2\lambda_2 \mathbf{x}, \quad (7)$$

que ao ser igualado a zero resulta em

$$\mathbf{x} = -\frac{1}{2}(\mathbf{B} - \lambda_2 \mathbf{I})^{-1} (\boldsymbol{\beta} - \lambda_1 \mathbf{1}) \quad (8)$$

Os elementos do vetor  $\mathbf{x}$  são as proporções da mistura que maximizam ou minimizam a resposta do sistema multicomponente.

A expressão obtida em (8) pode ser reescrita da forma,

$$\mathbf{1} = \mathbf{1}'\mathbf{x} = -\frac{1}{2}(\mathbf{B} - \lambda_2 \mathbf{I})^{-1} (\boldsymbol{\beta} - \lambda_1 \mathbf{1}), \quad (9)$$

e por meio de rearranjos chega-se a uma igualdade para  $\lambda_1$ :

$$\lambda_1 = \frac{1 + \frac{1}{2} \mathbf{1}'(\mathbf{B} - \lambda_2 \mathbf{I})^{-1} \boldsymbol{\beta}}{\frac{1}{2} \mathbf{1}'(\mathbf{B} - \lambda_2 \mathbf{I})^{-1} \mathbf{1}} \quad (10)$$

O valor para  $\lambda_2$  é definido pelo pesquisador, conforme o objetivo do estudo em obter o ponto que maximiza ou minimiza a resposta da mistura. O ponto de máximo ou mínimo fica definido por  $(\lambda_1, \lambda_2, \mathbf{x}, \mathbf{R}, \hat{y})$ .

A derivada de segunda ordem  $\frac{\partial^2 F}{\partial x_i \partial x_j} = 2(\mathbf{B} - \lambda_2 \mathbf{I})$  neste contexto não é apropriada, pois a matriz  $\mathbf{B}$  não reflete o fato que se trata de um problema restrito. Torna-se necessário considerar uma matriz com propriedades que assegurem fornecer uma solução no espaço da mistura. Uma matriz  $\mathbf{T}$  de dimensões  $(q - 1) \times q$  considerada conveniente é formada por coeficientes de polinômios ortogonais.<sup>9</sup>

A matriz  $\mathbf{TBT}'$  sendo positiva definida fornece um ponto mínimo na mistura; se  $\mathbf{TBT}'$  for negativa definida, tem-se um ponto de máximo. O valor atribuído ao multiplicador de Lagrange  $\lambda_2$  está relacionado aos autovalores da matriz  $\mathbf{TBT}'$ , de forma que: escolhendo-se um valor para  $\lambda_2$  superior ao maior dos autovalores de  $\mathbf{TBT}'$ , obtém-se um local de máximo no espaço da mistura; escolhendo-se um valor para  $\lambda_2$  inferior ao menor dos autovalores de  $\mathbf{TBT}'$ , obtém-se um local de mínimo no espaço da mistura; escolhendo-se um valor para  $\lambda_2$  intermediário em relação aos autovalores de  $\mathbf{TBT}'$ , obtém-se um ponto estacionário no espaço da mistura.<sup>9</sup>

## PARTE EXPERIMENTAL

Em termos de pesquisa, neste artigo ajustaram-se os modelos quadráticos, seguindo as duas abordagens - Scheffé (*S-model*) e Kronecker (*K-model*) - acrescentados da técnica de análise Ridge na modelagem das misturas dos componentes hexano, acetona e álcool etílico, cujo propósito foi maximizar o percentual de óleo extraído da polpa de Pequi.

A motivação para essa aplicação é verificada ao comparar os modelos (*S-model*) e (*K-model*) e sugerir o modelo de Kronecker (*K-model*) como alternativa, além do modelo Scheffé (*S-model*), para análise de sistemas multicomponentes, na utilização da técnica de análise Ridge diretamente no espaço das misturas sem transformações das variáveis.

Os dados utilizados para a aplicação da análise Ridge foram obtidos em um experimento realizado na Universidade Federal de Lavras, no Departamento de Ciência dos Alimentos, que objetivou estudar o processo de extração da polpa de pequi (*Caryocar brasiliense*), empregando misturas de diferentes solventes (hexano, acetona e álcool etílico). O experimento foi realizado em delineamento inteiramente casualizado, sendo os tratamentos gerados pela regra da mistura por meio do planejamento *simplex* centroide com 3 componentes e 6 repetições.

O procedimento utilizado para extrair o óleo da polpa de pequi foi realizado considerando a metodologia de Soxhlet com emprego de hexano como solvente (reagente PA da Proquímios), no período de 4 h de extração. Desta forma, o percentual (g de óleo/100 g de amostra seca) de óleo extraído da polpa de pequi foi determinado após uma sequência de passos para preparação das amostras e execução do processo.<sup>15</sup>

A Tabela 1 representa as proporções dos componentes das misturas, seguindo o delineamento mencionado.

Aplicaram-se os modelos de segunda ordem de Scheffé (1) e Kronecker (3) na modelagem estatística para obtenção da combinação dos solventes (hexano, acetona e álcool etílico) que maximiza o percentual de óleo extraído da polpa do pequi.

**Tabela 1.** Planejamento experimental das misturas e os valores médios do percentual de óleo extraído da polpa de pequi

Formulação	Proporção			Resposta média
	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	Percentual de óleo (%)
M <sub>1</sub>	1,00	0,00	0,00	60,11
M <sub>2</sub>	0,00	1,00	0,00	61,01
M <sub>3</sub>	0,00	0,00	1,00	39,72
M <sub>4</sub>	0,50	0,50	0,00	40,10
M <sub>5</sub>	0,50	0,00	0,50	62,04
M <sub>6</sub>	0,00	0,50	0,50	64,11
M <sub>7</sub>	0,33	0,33	0,34	61,01

X<sub>1</sub> proporção do hexano, X<sub>2</sub> proporção de acetona, X<sub>3</sub> proporção de álcool etílico.

A análise Ridge foi realizada em ambos os modelos (*K-model*) e (*S-model*) ajustados pelo método de mínimos quadrados, considerando as restrições dispostas na Equação 4 e a função lagrangeana definida pela Equação 5, para assegurar que a resposta máxima se encontra no espaço da mistura.

A análise Ridge deve considerar por conveniência das restrições, uma matriz  $\mathbf{T}$  ortonormal de dimensões  $(q - 1) \times q$ . Neste trabalho, como se trata de um experimento cuja mistura tem 3 componentes, foi considerada a matriz ortonormal  $\mathbf{T}_{(2 \times 3)}$  formada por coeficientes de polinômios ortogonais,<sup>9</sup> representada na Equação 11:

$$\mathbf{T} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} \quad (11)$$

Após a obtenção dos autovalores da matriz  $\mathbf{TBT}'$ , considerando os modelos Scheffé e Kronecker, foram atribuídos de forma conveniente valores a  $\lambda_2$  para obtenção de pontos do tipo  $(\lambda_1, \lambda_2, \mathbf{x}, \mathbf{R}, \hat{y})$ . Os resultados foram comparados para verificação de possíveis diferenças apresentadas pelos modelos. No desenvolvimento das análises utilizou-se software.<sup>16</sup>

## RESULTADOS E DISCUSSÃO

Os modelos de segunda ordem de Scheffé (Equação 12) e Kronecker (Equação 13) ajustados para a variável resposta representada pelo percentual de óleo extraído da polpa de pequi foram expressos por:

$$\hat{y} = 59,777x_1 + 60,674x_2 + 39,396x_3 - 75,322x_1x_2 + 55,157x_1x_3 + 61,617x_2x_3 \quad (12)$$

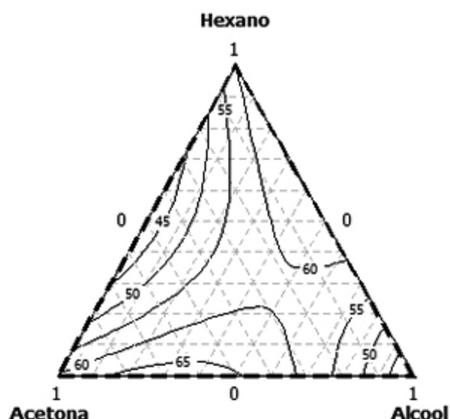
$$\hat{y} = 59,777x_1^2 + 60,674x_2^2 + 39,396x_3^2 + 45,129x_1x_2 + 154,330x_1x_3 + 161,687x_2x_3 \quad (13)$$

O modelo de Scheffé é composto pelo vetor  $\mathbf{b}$  e a matriz  $\mathbf{B}_S$ , enquanto que o modelo Kronecker apresenta apenas a matriz  $\mathbf{B}_K$ , conforme descritos nas Equações 14 e 15:

$$\mathbf{b}' = [59,777 \quad 60,674 \quad 39,396] \quad (14)$$

$$\mathbf{B}_S = \begin{bmatrix} 0,000 & -37,661 & 27,578 \\ -37,661 & 0,000 & 30,808 \\ 27,578 & 30,808 & 0,000 \end{bmatrix} \text{ e } \mathbf{B}_K = \begin{bmatrix} 59,777 & 22,564 & 77,165 \\ 22,564 & 60,674 & 80,843 \\ 77,165 & 80,843 & 39,396 \end{bmatrix} \quad (15)$$

Para a interpretação dos modelos, no sentido de se obter a combinação que determina o percentual máximo de óleo extraído da polpa de pequi, foi construído o gráfico das curvas de nível das seções da superfície gerada pelas componentes da mistura, considerando as estimativas obtidas pela modelo de Scheffé (*S-model*), conforme apresentado na Figura 2.



**Figura 2.** Curvas de nível dos fatores hexano, acetona e álcool etílico para o percentual de óleo extraído da polpa de pequi

Mantendo-se o foco no alinhamento das curvas de contorno (Figura 2), nota-se que as misturas com maiores quantidades de acetona e álcool evidenciam valores superiores, considerando-se a resposta

dada pelo percentual da extração da polpa do pequi. Entretanto, devido ao comportamento “irregular” das curvas próximas aos pontos experimentais, caracterizados pelas misturas com porcentagens dos componentes, suspeita-se da existência de que o ponto estacionário seja descrito por um ponto de sela. Frente ao exposto, justifica-se então a utilização da modelagem Ridge incorporada nos modelos quadráticos, em consonância com as abordagens aqui mencionadas, adota-se o Kronecker (*K-model*)<sup>9,12</sup> e compara-se com a análise Ridge realizada seguindo o modelo de Scheffé (*S-model*), usualmente empregado na literatura.

Os autovalores fornecidos pelas matrizes  $TB_kT'$  e  $TB_sT'$ , considerando os modelos quadráticos de Kronecker e Scheffé respectivamente, coincidiram, apresentando os seguintes valores:  $a_1 = -51,517$  e  $a_2 = 37,700$ . O caminho que segue para a solução apropriada na análise aplicada, com base nos dois modelos, foi mapeado considerando-se o maior autovalor obtido no *K-model* e *S-model*, ao estabelecer valores para  $\lambda_2$  que satisfaçam  $\lambda_2 \geq 37,700$ , o que garante o deslocamento para o ponto que fornece a resposta máxima. As Tabelas 2 e 3 mostram alguns cálculos selecionados para este caminho, partindo do centroide do espaço da mistura.

Para o modelo de segunda ordem de Scheffé (Tabela 2), o percentual máximo de óleo extraído da polpa de pequi foi definido por (39,8451; 41; 0,0145; 0,5602; 0,4253; 0,7035; 66,0196), ou seja, as proporções que otimizam a resposta correspondem a 0,0145 de hexano, 0,5602 de acetona e 0,4253 de álcool etílico, fornecendo um percentual de 66,0196 da extração do óleo da polpa de pequi. Observa-se que as restrições apresentadas na Equação 4 não foram violadas, assegurando que a solução se encontra no espaço da mistura. Para

**Tabela 2.** Caminho Ridge para obtenção da resposta máxima do percentual de óleo extraído da polpa de pequi considerando o modelo de segunda ordem de Scheffé

$\lambda_{s1}$	$\lambda_{s2}$	Hexano	Acetona	Álcool etílico	Raio	$\hat{y}$
-	$\infty$	0,3333	0,3324	0,3344	0,5774	58,0640
-603,8360	1000	0,3287	0,3303	0,3408	0,5774	58,0640
-1,8065	100	0,2935	0,3207	0,3857	0,5812	58,9805
34,3090	48	0,2061	0,3791	0,4148	0,5985	60,1817
35,0420	47	0,1962	0,3880	0,4158	0,6016	60,3572
35,7832	46	0,1839	0,3992	0,4169	0,6058	60,5916
36,5354	45	0,1684	0,4135	0,4181	0,6117	60,9179
37,3038	44	0,1481	0,4325	0,4194	0,6204	61,3964
38,0970	43	0,1202	0,4589	0,4210	0,6342	62,1476
38,9317	42	0,0795	0,4977	0,4228	0,6578	63,4443
<b>39,8451</b>	<b>41</b>	<b>0,0145</b>	<b>0,5602</b>	<b>0,4253</b>	<b>0,7035</b>	<b>66,0196</b>
40,9385	40	-0,1066	0,6775	0,4290	0,8090	72,47391

**Tabela 3.** Caminho Ridge para obtenção da resposta máxima do percentual de óleo extraído da polpa de pequi considerando o modelo de segunda ordem de Kronecker

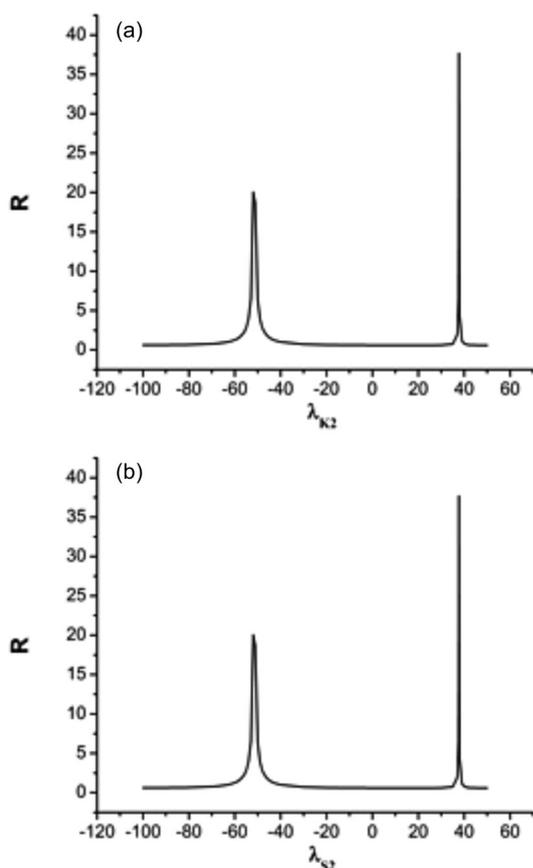
$\lambda_{k1}$	$\lambda_{k2}$	Hexano	Acetona	Álcool etílico	Raio	$\hat{y}$
-	$\infty$	0,3333	0,3324	0,3344	0,5774	58,0640
-550,7090	1000	0,3287	0,3303	0,3408	0,5774	58,0640
50,3958	100	0,2935	0,3207	0,3857	0,5812	58,9805
85,9730	48	0,2061	0,3791	0,4148	0,5985	60,1817
86,6929	47	0,1962	0,3880	0,4158	0,6016	60,3572
87,4216	46	0,1839	0,3992	0,4169	0,6058	60,5916
88,1622	45	0,1684	0,4135	0,4181	0,6117	60,9179
88,9205	44	0,1481	0,4325	0,4194	0,6204	61,3964
89,7062	43	0,1202	0,4589	0,4210	0,6342	62,1476
90,5379	42	0,0795	0,4977	0,4228	0,6578	63,4443
<b>91,4574</b>	<b>41</b>	<b>0,0145</b>	<b>0,5602</b>	<b>0,4253</b>	<b>0,7035</b>	<b>66,0196</b>
92,5785	40	-0,1067	0,6776	0,4291	0,8091	72,4739

$\lambda_2 = 40$ , a proporção de hexano é negativa, portanto o ponto (40,9385; 40; -0,1066; 0,6775; 0,4290; 0,8090; 72,47391) encontra-se fora do *simplex*, não sendo de interesse prático.

Com base nos resultados apresentados na Tabela 3, é fácil verificar o ponto que maximiza o percentual da extração do óleo da polpa de pequi empregando-se o modelo de Kronecker, definido por (91,4574; 41; 0,0145; 0,5602; 0,4253; 0,7035; 66,0196) que corresponde ao percentual 66,0196 e é obtido pelas proporções 0,0145 de hexano, 0,5602 de acetona e 0,4253 de álcool etílico. As restrições aplicadas não foram violadas para a resposta de interesse, ou seja, a soma das proporções resulta em 1, o raio da esfera é inferior a 1 e o valor de  $\lambda_2$  é superior ao maior autovalor da matriz  $\mathbf{TB}_K\mathbf{T}'$ .

Conforme apresentado na Figura 3, independente do modelo utilizado, os resultados alcançados com a aplicação da análise Ridge não apresentaram diferenças entre as respostas, nem alterações nas composições das misturas. A uniformidade de resultados obtida pelos dois modelos é uma consequência da igualdade verificada nas expressões  $\mathbf{B}_S + \frac{1}{2}\mathbf{b}\mathbf{1}' + \frac{1}{2}\mathbf{b}' = \mathbf{B}_K \Rightarrow \mathbf{x}'\mathbf{B}_S\mathbf{x} + \frac{1}{2}(\mathbf{x}'\mathbf{b} + \mathbf{b}'\mathbf{x}) = \mathbf{x}'\mathbf{B}_K\mathbf{x}$ .

Ainda na Figura 3 pode-se verificar que o valor do raio R aumenta consideravelmente quando  $\lambda_2$  se aproxima dos autovalores. Para a trajetória da resposta máxima prevista, é preciso ter valores de  $\lambda_2$  para os quais  $37,700 < \lambda_2 < \infty$ .



**Figura 3.** (a) Variação do raio R em relação ao valor do  $\lambda_{K2}$  (Kronecker), (b) variação do raio R em relação ao valor do  $\lambda_{S2}$  (Scheffé)

Ao se fazer referência à contribuição de cada componente na formação do modelo quadrático, mantendo-se o nível de significância fixado em 5%, os resultados descritos na Tabela 4 confirmam que as estimativas dos parâmetros dos modelos (*S-model*) e (*K-model*) foram significativas. Assim sendo, há evidências estatísticas para afirmar que cada parâmetro de fato apresenta uma relação com a variável

dependente, no sentido de que a representatividade dos efeitos lineares e quadráticos, explícitos em cada covariável, não é nula.

**Tabela 4.** Estimativas dos parâmetros do modelo quadrático obtidas com a aplicação dos métodos de Scheffe (*S-model*) e Kroncker (*K-model*)

<i>S-model</i>			
Parâmetros	Estimativa	Erro padrão	valor-p
$\beta_1$	59,777	4,299	0,000
$\beta_2$	60,674	4,299	0,000
$\beta_3$	39,396	4,299	0,000
$\beta_{12}$	-75,322	19,801	0,000
$\beta_{13}$	55,157	19,742	0,008
$\beta_{23}$	61,617	19,742	0,003
$R^2 = 0,9707$			
<i>K-model</i>			
Parâmetros	Estimativa	Erro padrão	valor-p
$\beta_{11}$	59,777	4,299	0,000
$\beta_{22}$	60,674	4,299	0,000
$\beta_{33}$	39,396	4,299	0,000
$\beta_{12}$	45,129	17,059	0,012
$\beta_{13}$	154,330	16,993	0,000
$\beta_{23}$	161,687	16,993	0,000
$R^2 = 0,9707$			

Em se tratando da qualidade de ajuste dos modelos (*K-model*) e (*S-model*), por se tratar de parametrizações a serem utilizadas no ajuste do modelo quadrático, espera-se o mesmo valor do coeficiente de determinação  $R^2$ , bem como uma única análise de variância. Desta forma, pode-se observar, por meio da Tabela 4, que o coeficiente de determinação estimado foi de 97,07%, evidenciando que praticamente toda variabilidade amostral é explicada pelo modelo, naturalmente em ambas as parametrizações. Para uma melhor validação do modelo, os resultados da análise de variância encontram-se descritos na Tabela 5.

**Tabela 5.** Análise de variância para o modelo quadrático ajustado para o percentual de óleo extraído da polpa de pequi

FV	G.L.	SQ	QM	valor-p
Regressão	5	4023,74	804,75	0,000
Linear	2	1737,96	868,98	0,002
Quadrática	3	3738,20	1246,07	0,000
Erro	36	4022,44	111,73	
Falta de Ajuste	1	85,98	85,98	0,388
Erro Puro	35	3936,46	112,47	
Total	41	8046,18		

SQ = Soma de quadrados; QM = Quadrado médio

Em relação à interpretação dos resultados pertinentes à análise de variância (Tabela 5) convém alertar o leitor do seguinte aspecto: em síntese, quando se deseja comparar modelos de diferentes ordens (ex. lineares, quadráticos, cúbicos), conjecturando a situação de que o modelo linear foi significativo, a um nível nominal de significância previamente estabelecido, implica na afirmação que não há necessidade em ajustar um modelo de ordem superior.

Com base nessas considerações, a interpretação da análise de

variância (Tabela 5) é dada individualmente a cada modelo sem o propósito de compará-los. Nesse contexto, o resultado referente à fonte de variação da regressão quadrática foi significativo, portanto, é cabível assumir que o percentual de extração de óleo da polpa de pequi apresente uma relação funcional quadrática, com os componentes utilizados na formação das misturas.

Tendo por base essa afirmativa e o resultado não significativo do teste de falta de ajuste (valor-p = 0,388), os resultados descritos na Tabela 5 evidenciam que o modelo quadrático em ambas as parametrizações utilizadas na obtenção das estimativas é adequado aos dados experimentais, portanto, pode ser utilizado para prever a resposta máxima do percentual de óleo extraído da polpa de pequi (*Caryocar brasiliense*). Entretanto, convém ressaltar que embora os valores dos erros padrão  $\beta_{ij}$  ( $i \neq j$ ) observados nas estimativas dos parâmetros relacionadas ao efeito da interação, em ambos os modelos, observou-se ao comparar a interpretação física dos coeficientes, por meio desta aplicação, a interação entre o primeiro e segundo componentes, representada nos modelos pelo termo  $x_{12}$ , apresentaram efeitos contraditórios. Neste sentido, para esta interação o modelo (S-model) apresentou um efeito antagônico e o modelo (K-model) um efeito sinérgico.

Tendo por base que os modelos de primeira ordem de Kronecker (K-model) e os modelos de primeira ordem de Scheffé (S-model) são da mesma forma nos  $x_i$ s, isto é, apresentam as mesmas estimativas, há evidências estatísticas para afirmar que ao considerar o modelo (K-model) o percentual de óleo extraído da polpa de pequi (*Caryocar brasiliense*) obtido com os dois componentes  $x_1$  (hexano) e  $x_2$  (acetona) misturados é maior do que a soma dos percentuais obtidos ao se considerar os componentes individuais. No caso do modelo (S-model) o percentual obtido desta mistura será inferior aos percentuais obtidos individualmente para cada componente.

Frente ao exposto, pode-se concluir que embora os resultados inferenciais em ambos os modelos, no que tange à qualidade de ajuste e significância dos parâmetros, tenham sido equivalentes, dependendo da mistura os modelos (K-model) e (S-model) quanto à interpretação de seus efeitos de interação poderão ser contraditórios. Entretanto, por questões de natureza estatística e não física, seguindo as advertências,<sup>17</sup> o modelo (K-model) é recomendado por apresentar as seguintes características: a matriz  $X'X$  é melhor condicionada, de tal forma, que a variância dos valores preditos pelo modelo (K-model) é menor em relação ao modelo (S-model); as estimativas de mínimos quadrados dos coeficientes de qualquer termo quadrático são idênticas aos respectivos valores dos termos diretamente associados ao modelo (S-model); o fator de inflação da variância (VIF) dos coeficientes de qualquer termo quadrático serão sempre menores ou iguais ao VIF dos coeficientes dos termos correspondentes no modelo quadrático (S-model).

## CONCLUSÃO

A análise Ridge aplicada aos modelos de segunda ordem de Kronecker (K-model) e Scheffé (S-model) conduziu às mesmas proporções de hexano, acetona e álcool etílico que resultam no percentual máximo do óleo extraído da polpa do pequi.

O modelo de Kronecker (K-model) por apresentar o vetor **b** nulo facilita a implementação dos cálculos da análise Ridge em sistemas multicomponentes, podendo ser usualmente empregado em modelagem de misturas acompanhado de processos de otimização, visto que fornece os mesmos resultados em relação às estimativas proporcionadas pelo modelo Scheffé (S-model). Entretanto, devido à melhoria no condicionamento da matriz  $X'X$  traz como vantagens a redução dos erros padrão dos coeficientes estimados, reduz a correlação entre os estimadores e torna o modelo menos dependente da localização precisa dos pontos do delineamento.

Em se tratando do efeito de interação, há evidências para afirmar que os modelos (K-model) e (S-model) poderão apresentar efeitos contrastantes em relação à natureza do efeito ser sinérgico ou antagônico.

## REFERÊNCIAS

1. Reis, C.; Andrade, J. C.; *Quim. Nova* **1996**, *19*, 313.
2. Scheffé, H. J.; *Royal Stat. Soc.* **1958**, *B 20*, 344.
3. Scheffé, H.; *Royal Stat. Soc.* **1963**, *B 25*, 235.
4. de Barros Neto, B.; Scarminio, I. S.; Bruns, R. E.; *Como fazer experimentos*, Bookman: Porto Alegre, 2010.
5. Draper, N. R., Pukelsheim, F.; *Technometrics* **2002**, *44*, 250.
6. Cornell, J. A.; *Experiments with mixtures: designs, models and the analysis of mixture data*, Wiley: New York, 2002.
7. Draper, N. R.; Smith, H.; *Applied regression analysis*, Wiley: New York, 1998.
8. Box, G. E. P.; Draper, N. R.; *Response Surfaces, mixtures and ridge analyses*, Wiley: New Jersey, 2007.
9. Draper, N. R.; Pukelsheim, F.; *Statistics and Probability Letters* **2000**, *48*, 131.
10. Bueno, L.; *Quim. Nova* **2008**, *31*, 585.
11. Faria, R. C. M.; Rezende, M. J. C.; Rezende, C. M.; Pinto, A. C.; *Quim. Nova* **2007**, *30*, 1900.
12. Draper, N. R.; Pukelsheim, F.; *J. Statist. Plann. Inference* **1998**, *71*, 303.
13. Hoerl, A. E.; *Chem. Eng. Prog.* **1959**, *55*, 69.
14. Hoerl, R. W.; *Am. Statist.* **1985**, *39*, 186.
15. Aquino, L. P.; Ferrua, F. Q.; Borges, S. V.; Antoniassi, S.; Correa, J. L. G.; Cirillo, M. A.; *Cienc. Tecnol. Aliment.* **2009**, *29*, 354.
16. R development core team. *R: A language and environment for statistical computing*, R Foundation for Statistical Computing: Vienna, 2010.
17. Prescott, P.; Dean, A. M.; Draper, N. R.; Lewis, S. M.; *Technometrics* **2002**, *44*, 260.