

Utilização de modelos de regressão
em experimentos do tipo mistura

Marcelo Burani Kowalski

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
MESTRE EM CIÊNCIAS

Programa: Estatística

Orientadora: Profa. Dra. Silvia Nagib Elian

São Paulo, agosto de 2011

Utilização de modelos de regressão em experimentos do tipo mistura

Esta dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa realizada por Marcelo Burani Kowalski em 26/08/2011.

O original encontra-se disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão Julgadora:

Profa. Dra. Silvia Nagib Elian - IME-USP

Profa. Dra. Airlane Pereira Alencar -IME -USP

Prof. Dr. Juvêncio Santos Nobre -UFC

Agradecimentos

Agradeço à minha orientadora, professora Silvia Nagib Elian, por acreditar no meu potencial, pelos ensinamentos e por seus conselhos nos momentos mais importantes deste trabalho.

Agradeço à minha esposa, Mariana Kowalski, por tornar meus dias mais felizes, me fazer acreditar no futuro e ser o amor da minha vida.

Agradeço ao amigo, Tiago Maia Magalhães, pela valiosa ajuda nesta dissertação.

Agradeço aos meus pais, Júlio e Liliana, por serem as minhas referências de conduta de vida, valores e ideais.

Agradeço à minha tia, Lígia Burani, por ter me recebido tão bem em São Paulo e me tratado com muito carinho, além de ter me incentivado muito a seguir estudando.

Agradeço aos meus irmãos, Vinícius e César, e ao meu avô Júlio por torcerem por mim.

Resumo

Nos experimentos do tipo mistura, um dos objetivos principais é prever, através dos modelos de regressão, uma determinada variável resposta através de um conjunto de variáveis preditoras que são componentes de uma mistura. Nestes procedimentos de regressão, destaca-se a utilização do modelo polinomial na forma canônica e algumas técnicas, como a regressão em cristas, que são necessárias para solucionar um problema muito comum neste tipo de experimento que é a multicolinearidade. Após a definição do modelo, é muito frequente o interesse em maximizar ou minimizar a variável resposta e para este propósito sugere-se o método de análise de cristas.

O objetivo deste trabalho foi detalhar estas técnicas e métodos, ressaltar suas vantagens e solucionar os possíveis problemas. A importância destes aspectos teóricos torna-se evidente quando apresentado o problema de maximização do ganho de massa de um grupo de frangos em função da composição do suplemento ingerido.

Palavras-chave: experimentos do tipo mistura, multicolinearidade, regressão em cristas, análise de cristas.

Abstract

In the experiments with mixtures, an important goal is to predict, using the regression models, a particular response variable through a set of predictor variables that are components of a mixture. In these regression procedures, there are the use of polynomial model in canonical form and some techniques such as ridge regression to solve the problem of multicollinearity in this kind of experiment. After defining the model, it is interesting to maximize or minimize the response variable, so it is suggested the ridge analysis for this purpose.

The objective of this work was to detail these techniques and methods, to show their advantages and solve the possible problems. The importance of these theoretical aspects becomes evident when presented the problem of maximizing mass gain in a group of chickens according to the composition of the supplement intake.

Keywords: experiments with mixtures, multicollinearity, ridge regression, ridge analysis.

Sumário

1	Introdução	1
2	Experimentos do tipo mistura –Aspectos gerais	3
	2.1 Introdução	3
	2.2 Uso do planejamento simplex.....	5
	2.3 Ajuste Polinomial.....	7
	2.4 Planejamento para o ajuste do modelo polinomial.....	11
	2.5 Considerações finais.....	14
3	Multicolinearidade em experimentos do tipo mistura	16
	3.1 Introdução	16
	3.2 Exemplos.....	25
	3.3 Uso da regressão em cristas nos experimentos do tipo mistura.....	43
	3.4 Considerações finais.....	50
4	Análise de Cristas em experimentos do tipo mistura	53
	4.1 Introdução.....	53
	4.2 Especificações dos experimentos do tipo mistura.....	62
	4.3 Exemplo.....	70
5	Aplicação	77
	5.1 Introdução.....	77
	5.2 Descrição das variáveis.....	78
	5.3 Análises.....	80
	5.3.1 Modelo polinomial quadrático com variável resposta MC....	80
	5.3.2 Introdução da variável complementar TSI no modelo.....	91
	5.4 Considerações finais.....	97

6	Considerações Finais	99
	Apêndice 1.....	101
	Apêndice 2.....	102
	Apêndice 3.....	103

Capítulo 1

Introdução

A utilização dos modelos de regressão nos experimentos do tipo mistura tem como objetivo prever uma determinada variável resposta através de um conjunto de variáveis preditoras que são componentes de uma mistura.

A variável resposta deve ser uma propriedade da mistura constituída por k componentes e deve variar de acordo com a composição da mistura. As variáveis preditoras são as proporções dos componentes satisfazendo a condição que, para uma determinada composição, a soma das proporções de todos os componentes deve ser igual a 1.

O propósito deste trabalho é apresentar os modelos mais adequados para os experimentos do tipo mistura, apontando suas vantagens e procurando solucionar os problemas que normalmente surgem durante o processo de estimação do modelo, entre eles, destaca-se a multicolinearidade. Uma vez estabelecido o modelo, um interesse muito comum dos pesquisadores é maximizar ou minimizar a variável resposta, logo para este objetivo será apresentado o processo de análise de cristas.

No Capítulo 2, inicialmente será definida a malha simplex e a partir dela serão construídos os modelos polinomiais na forma canônica do primeiro, segundo e terceiro grau, que devem ser os modelos utilizados nos experimentos do tipo mistura, de acordo com Scheffé (1958).

Um problema muito frequente nos modelos que envolvem mistura é a multicolinearidade. No Capítulo 3, este problema será definido e exemplificado, assim como as técnicas para solucioná-lo, entre elas destaca-se a técnica de regressão em cristas.

Em muitos experimentos do tipo mistura, existe, por parte do pesquisador, o interesse de encontrar o valor máximo ou mínimo da variável resposta depois do modelo definido. A maximização ou minimização pode ser realizada através de diferentes procedimentos. Assim, a técnica selecionada para esta dissertação, amplamente adotada na literatura relacionada ao assunto, foi a análise de cristas, que encontra-se estruturada e exemplificada no Capítulo 4.

No Capítulo 5, apresenta-se uma aplicação da utilização dos modelos de regressão nos experimentos do tipo mistura. O primeiro passo consiste na construção de um modelo para o experimento e, definido o modelo, um segundo objetivo é maximizar a variável resposta. No entanto, a proposta principal é verificar a eficiência das técnicas apresentadas nesta dissertação dentro de um contexto prático.

Capítulo 2

Experimentos do tipo mistura – Aspectos gerais

2.1 Introdução

Sejam X_1, X_2, \dots e X_k as variáveis preditoras e Y a variável resposta. Consideremos o modelo de regressão linear múltipla $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + e_i$, em que y_i representa uma observação da variável resposta Y , quando os respectivos valores das variáveis preditoras são $x_{i1}, x_{i2}, \dots, x_{ik}$, correspondentes à i -ésima observação da amostra de tamanho n , e e_i é o respectivo erro.

Em algumas situações experimentais, a variável resposta depende apenas das quantidades relativas das variáveis preditoras e não das quantidades absolutas. Alguns exemplos são a dureza e a resistência à compressão de uma determinada formulação de concreto destinado à construção civil, composto por areia, água e um ou mais tipos de cimento. Neste caso, as variáveis resposta seriam as medidas de dureza e resistência e dependeriam apenas das proporções

dos componentes utilizados. A textura e as características sensoriais de alimentos industrializados, como consequência dos ingredientes, e o desempenho em quilômetros por litro de um veículo de acordo com a mistura de combustíveis inserida em seu tanque também seriam possíveis variáveis resposta, determinadas apenas pelas quantidades relativas das variáveis preditoras.

De acordo com Steinberg e Hunter (1984) e vários outros autores, os experimentos do tipo mistura são definidos como um particular grupo dessas situações experimentais, em que os valores x_j representam as respectivas proporções em massa, volume, mols ou outra grandeza, sendo satisfeitas as seguintes restrições:

$$0 \leq x_j \leq 1 \quad \forall j = 1, 2, \dots, k, \quad (2.1)$$

com

$$\sum_{j=1}^k x_j = 1. \quad (2.2)$$

Estas restrições apresentam alguns problemas para a modelagem estatística, pois se o modelo inclui todas as variáveis preditoras e um termo constante, o mesmo encontra-se super-parametrizado. A ocorrência deste fato justifica-se devido à restrição (2.2), e pode ser verificada a seguir:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + e_i, \quad i = 1, \dots, n,$$

e

$$x_{ik} = 1 - x_{i1} - x_{i2} - \dots - x_{ik-1}.$$

Substituindo na equação do modelo de regressão o valor de x_{ik} obtido a partir da restrição (2.2), tem-se:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_{k-1} x_{ik-1} + \beta_k (1 - x_{i1} - \dots - x_{ik-1}) + e_i$$

e portanto,

$$y_i = \beta_0 + \beta_k + (\beta_1 - \beta_k) x_{i1} + (\beta_2 - \beta_k) x_{i2} + \dots + (\beta_{k-1} - \beta_k) x_{ik-1} + e_i.$$

Desta forma, $\beta_0 + \beta_k$, $\beta_1 - \beta_k$, $\beta_2 - \beta_k$, ... e $\beta_{k-1} - \beta_k$ podem ser estimados, mas não β_0 , β_1 , β_2 , ... e β_k separadamente.

O problema da multicolinearidade em experimentos do tipo mistura será discutido no próximo capítulo.

Apresentaremos nas seções seguintes alguns tópicos descritos em Scheffé (1958), artigo introdutório na área de experimentos do tipo mistura.

2.2 Uso do planejamento simplex

De acordo com Scheffé (1958), devido às restrições (2.1) e (2.2), o espaço de um experimento do tipo mistura contendo os k componentes pode ser entendido através de uma descrição geométrica denominada sistema de coordenadas simplex. Este sistema consiste em definir uma malha de pontos localizados no interior ou sobre os lados, vértices e faces de um simplex regular de $k-1$ dimensões. Logo, se $k=2$, o espaço simplex é um segmento de reta, correspondente a uma mistura binária, onde as extremidades representam os componentes puros. Se $k=3$, o espaço é um triângulo equilátero e para $k=4$, o espaço simplex é um tetraedro regular. Os vértices representam os componentes puros, tanto no triângulo quanto no tetraedro.

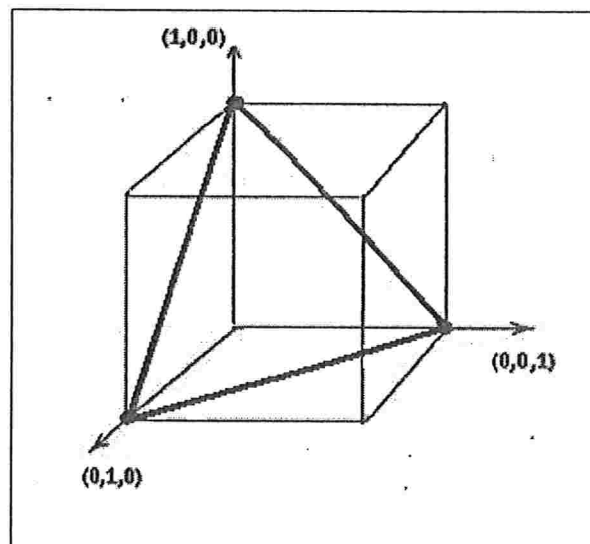


Figura 2.1: Representação de um espaço simplex com três componentes.

Na Figura 2.1, temos a representação do espaço referente a três componentes, onde o triângulo equilátero aparece orientado por um sistema de três eixos perpendiculares entre si e os três vértices representam a situação em que um componente ocorre na proporção 1 e os demais na proporção 0.

Os pontos na malha simplex devem apresentar uma distribuição espacial uniforme, deste modo definimos uma malha $\{k, m\}$, cujas propriedades são:

- as proporções utilizadas para cada componente (x_j) correspondem a $m+1$ valores igualmente espaçados limitados por 0 e 1, ou seja, $x_j \in \{0, \frac{1}{m}, \frac{2}{m}, \dots, 1\}$ e os outros componentes ($x_l, l \neq j$) devem ser fixados de modo a formar todas as misturas possíveis para cada valor de x_j utilizado.
- em uma malha $\{k, m\}$, o número de pontos no espaço simplex é dado por $\binom{m+k-1}{m}$. A demonstração desse fato encontra-se no Apêndice 1.

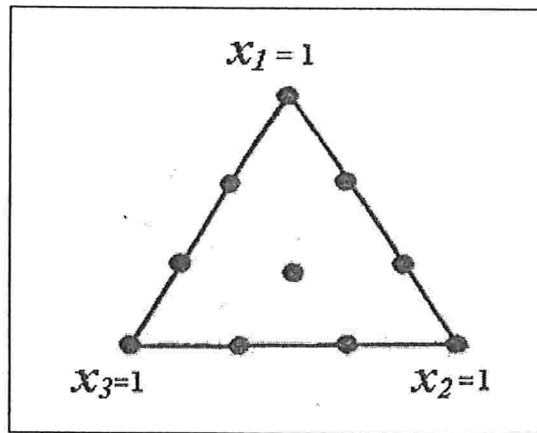


Figura 2.2: Malha{3,3}

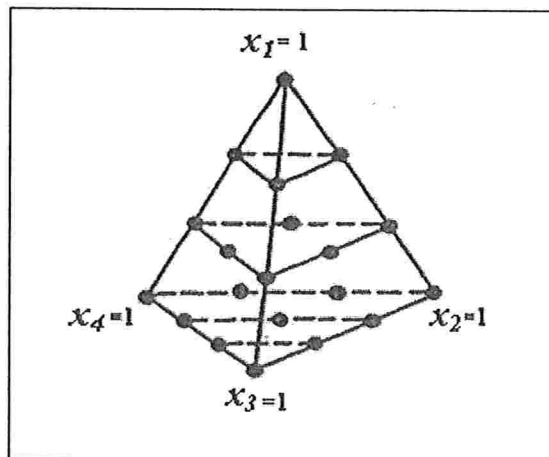


Figura 2.3: Malha{4,3}

Na Figura 2.2, temos uma malha{3,3} e seus 10 pontos uniformemente distribuídos no triângulo equilátero e na Figura 2.3, temos uma malha{4,3} e seus 20 pontos uniformemente distribuídos no tetraedro regular.

Será necessário modificar o planejamento simplex se por algum motivo as proporções de um ou mais componentes não puderem variar de 0 a 1, apresentando, por exemplo, um limite superior (S_j), um limite inferior (I_j) ou ambos, com

$$0 \leq I_j \leq x_j \leq S_j \leq 1, \quad 1 \leq j \leq k.$$

A propriedade estudada no experimento pode ser entendida como uma função real no simplex e será a variável resposta. Na análise dos experimentos do tipo mistura é utilizada uma apropriada forma de modelo de regressão polinomial. Uma primeira consideração é que na prática polinômios de baixo grau devem ser utilizados, para evitar um número muito grande de coeficientes a serem estimados.

2.3 Ajuste Polinomial

Uma função polinomial de grau p e k variáveis X_1, X_2, \dots, X_k será denominada polinomial $\{k, p\}$ se apresentar a seguinte forma funcional:

$$\begin{aligned} \eta_i = & \beta_0 + \sum_{1 \leq j \leq k} \beta_j x_{ij} + \sum_{1 \leq j \leq l \leq k} \beta_{jl} x_{ij} x_{il} + \sum_{1 \leq j \leq l \leq q \leq k} \beta_{j l q} x_{ij} x_{il} x_{iq} + \dots \\ & \dots \sum_{1 \leq j_1 \leq j_2 \leq \dots \leq j_p \leq k} \beta_{j_1 j_2 \dots j_p} x_{i j_1} x_{i j_2} \dots x_{i j_p}. \end{aligned} \quad (2.3)$$

O número de coeficientes da equação (2.3) é dado por $\binom{p+k}{p}$. No entanto, nos experimentos do tipo mistura, devido à restrição (2.2), uma das variáveis pode ser reescrita em função das demais, e conseqüentemente os coeficientes de regressão sofrerão alterações. Uma possível substituição pode ser efetuada através de

$$x_{ik} = 1 - \sum_{j=1}^{k-1} x_{ij}, \quad \forall i = 1, \dots, n. \quad (2.4)$$

No entanto, qualquer variável preditora pode ser substituída, a equação (2.4) consiste apenas em uma das k possibilidades. As conseqüências mais importantes da substituição de uma variável preditora em (2.3) seriam o surgimento de uma função polinomial com $k-1$ variáveis e a quantidade de coeficientes reduz-se a $\binom{p+k-1}{p}$.

Deste modo, temos a quantidade de coeficientes presentes na função polinomial coincidindo com o número de pontos de uma malha simplex $\{k, p\}$.

Segundo Scheffé (1958) e Cornell (1990), apesar da substituição de (2.4) em (2.3) gerar uma estrutura polinomial com um número menor de coeficientes, o efeito do k-ésimo componente seria desprezado. Portanto, para que este efeito não seja desconsiderado, uma alternativa seria a construção de polinômios em sua forma canônica que melhor se adaptam às restrições (2.1) e (2.2) e facilitam o ajuste do modelo de regressão e a interpretação de seus parâmetros. Conforme citação anterior, em termos práticos, o polinômio deve possuir um baixo grau, para evitar uma grande quantidade de coeficientes. Portanto, serão apresentados apenas os processos de obtenção das formas canônicas das funções polinomiais do primeiro, segundo e terceiro grau. Para que o efeito de nenhum componente seja desprezado na construção da forma canônica, Cornell (1990) sugere a multiplicação de alguns termos da função polinomial por $\sum_{j=1}^k x_{ij}$ (somatória de valor unitário, de acordo com a equação (2.2)), conforme verifica-se a seguir.

Da equação (2.3), quando $p = 1$, tem-se a função polinomial do primeiro grau:

$$\eta_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ij}, \quad i = 1, \dots, n. \quad (2.5)$$

Multiplicando β_0 por $\sum_{j=1}^k x_{ij}$, tem-se:

$$\eta_i = \beta_0 \left(\sum_{j=1}^k x_{ij} \right) + \sum_{j=1}^k \beta_j x_{ij}.$$

Logo,

$$\eta_i = \sum_{j=1}^k \beta_j^* x_{ij}, \quad (2.6)$$

portanto, os coeficientes são da forma β_j^* , onde $\beta_j^* = \beta_0 + \beta_j$, para todo $j=1,2,\dots,k$.

Da equação (2.3), quando $p = 2$, tem-se a função polinomial do segundo grau:

$$\eta = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \sum_{1 \leq j < l \leq k} \beta_{jl} x_{ij} x_{il}. \quad (2.7)$$

Na equação (2.7), multiplicando β_0 por $\sum_{j=1}^k x_{ij}$ e substituindo x_{ij}^2 por $x_{ij} \left(1 - \sum_{\substack{l=1 \\ l \neq j}}^k x_{il} \right)$ tem-se:

$$\eta_i = \beta_0 \left(\sum_{j=1}^k x_{ij} \right) + \sum_{j=1}^k \beta_j x_{ij} + \sum_{j=1}^k \beta_{jj} x_{ij} \left(1 - \sum_{\substack{l=1 \\ l \neq j}}^k x_{il} \right) + \sum_{1 \leq j < l \leq k} \beta_{jl} x_{ij} x_{il}.$$

Portanto,

$$\eta_i = \sum_{j=1}^k (\beta_0 + \beta_j + \beta_{jj}) x_{ij} - \sum_{j=1}^k \beta_{jj} x_{ij} \sum_{\substack{l=1 \\ l \neq j}}^k x_{il} + \sum_{1 \leq j < l \leq k} \beta_{jl} x_{ij} x_{il}.$$

Logo,

$$\eta_i = \sum_{j=1}^k \beta_j^* x_{ij} + \sum_{1 \leq j < l \leq k} \beta_{jl}^* x_{ij} x_{il}, \quad (2.8)$$

com coeficientes da forma β_j^* e β_{jl}^* , em que $\beta_j^* = \beta_0 + \beta_j + \beta_{jj}$ e $\beta_{jl}^* = \beta_{jl} - \beta_{jj} - \beta_{ll}$.

Deste modo, a estimação dos coeficientes da forma canônica passa a ser o objetivo do modelo. Quando $p = 3$, tem-se a forma canônica da função polinomial do terceiro grau :

$$\eta_i = \sum_{1 \leq j \leq k} \beta_j^* x_{ij} + \sum_{1 \leq j < l \leq k} \beta_{jl}^* x_{ij} x_{il} + \sum_{1 \leq j < l \leq k} \gamma_{jl} x_{ij} x_{il} (x_{ij} - x_{il}) + \sum_{1 \leq j < l < q \leq k} \beta_{jllq}^* x_{ij} x_{il} x_{iq}. \quad (2.9)$$

A partir deste momento, as funções polinomiais serão utilizadas somente em sua forma canônica, logo será omitido o asterisco dos coeficientes β^* e todo coeficiente representado por β corresponderá a um coeficiente da forma canônica.

Para interpretar os coeficientes destes polinômios, será utilizada a resposta média da função. Deste modo, a resposta média associada ao componente puro j (x_j é igual a 1 e os outros componentes são nulos) será denotada por η_j . Logo, substituindo os valores das variáveis preditoras associadas ao componente puro em uma função polinomial do primeiro, segundo ou terceiro grau tem-se:

$$\eta_j = \beta_j = E(Y | x_j = 1, x_l = 0, l \neq j), \quad (2.10)$$

sendo Y a variável resposta e $E(Y | \cdot)$ a média dessa variável condicional aos valores das variáveis preditoras.

Logo, os coeficientes β_j podem ser interpretados como a resposta média da função para os componentes puros.

A diferença encontrada na resposta média da função polinomial de grau p , em relação à resposta média obtida utilizando apenas as parcelas do primeiro grau do mesmo polinômio, é chamada em vários campos, como a farmacologia, de sinergismo se a diferença é positiva e antagonismo quando é negativa. Por exemplo, em uma mistura binária, ou seja, uma mistura composta por exatamente dois componentes, a resposta em uma função do segundo grau é dada por:

$$\eta_i = \beta_j x_{ij} + \beta_l x_{il} + \beta_{jl} x_{ij} x_{il} .$$

O sinergismo da mistura binária entre os componentes j e l é encontrado na terceira parcela e β_{jl} é o coeficiente quadrático binário do sinergismo dos componentes j e l .

Um segundo exemplo, em uma mistura binária, seria a resposta em uma função do terceiro grau:

$$\eta_i = \beta_j x_{ij} + \beta_l x_{il} + \beta_{jl} x_{ij} x_{il} + \gamma_{jl} x_{ij} x_{il} (x_{ij} - x_{il}).$$

Neste caso, o sinergismo seria a soma da terceira e quarta parcelas e γ_{jl} o coeficiente cúbico binário do sinergismo dos componentes j e l . Um último exemplo seria em uma mistura ternária (mistura composta por exatamente três componentes), a adoção de um polinômio de terceiro grau:

$$\eta_i = \sum_{1 \leq j \leq 3} \beta_j x_{ij} + \sum_{1 \leq j < l \leq 3} \beta_{jl} x_{ij} x_{il} + \sum_{1 \leq j < l \leq 3} \gamma_{jl} x_{ij} x_{il} (x_{ij} - x_{il}) + \beta_{j l q} x_{ij} x_{il} x_{iq} .$$

Assim, o sinergismo total da mistura ternária é dado por $\sum_{1 \leq j < l \leq 3} \beta_{jl} x_{ij} x_{il} + \sum_{1 \leq j < l \leq 3} \gamma_{jl} x_{ij} x_{il} (x_{ij} - x_{il}) + \beta_{j l q} x_{ij} x_{il} x_{iq}$, em que $\sum_{1 \leq j < l \leq 3} \beta_{jl} x_{ij} x_{il} + \sum_{1 \leq j < l \leq 3} \gamma_{jl} x_{ij} x_{il} (x_{ij} - x_{il})$ é o sinergismo binário na mistura ternária e $\beta_{j l q} x_{ij} x_{il} x_{iq}$ é o sinergismo ternário da mistura ternária.

2.4 Planejamento para o ajuste do modelo polinomial

No planejamento de um modelo polinomial cúbico ou quadrático para uma mistura de k componentes, as observações devem ser tomadas de uma malha simplex $\{k, m\}$. Os coeficientes das formas canônicas (2.6), (2.8) e (2.9) são unicamente determinados pelos valores da malha simplex $\{k, m\}$. Para exemplificar esta relação entre os coeficientes do polinômio e os valores da malha, serão apresentados os coeficientes β e γ em função das respostas de uma malha $\{3,2\}$ e posteriormente de uma malha $\{3,3\}$.

Uma notação específica será introduzida para viabilizar o equacionamento. Deste modo, para uma malha $\{3,2\}$, tem-se:

η_j : resposta para o componente puro j ($x_j = 1$),

η_l : resposta para o componente puro l ($x_l = 1$),

η_q : resposta para o componente puro q ($x_q = 1$),

η_{jl} : resposta para mistura binária dos componentes j e l ($x_j = \frac{1}{2}$ e $x_l = \frac{1}{2}$),

η_{jq} : resposta para mistura binária dos componentes j e q ($x_j = \frac{1}{2}$ e $x_q = \frac{1}{2}$)

e

η_{lq} : resposta para mistura binária dos componentes l e q ($x_l = \frac{1}{2}$ e $x_q = \frac{1}{2}$).

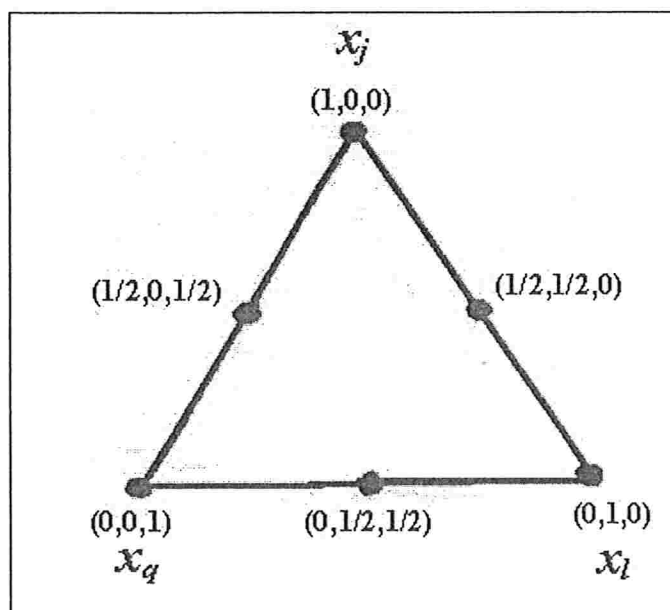


Figura 2.4: Malha $\{3,2\}$

Selecionando um dos polinômios (2.8) ou (2.9) e considerando a equação (2.10) para os componentes j, l e q , os coeficientes β_{jl} , β_{jq} e β_{lq} são obtidos de acordo com o procedimento a seguir:

Polinômio (2.8) para uma mistura de três componentes:

$$\eta = \beta_j x_j + \beta_l x_l + \beta_q x_q + \beta_{jl} x_j x_l + \beta_{jq} x_j x_q + \beta_{lq} x_l x_q .$$

Através da equação (2.10), tem-se:

$$\eta_j = \beta_j , \text{ para } x_j = 1 ,$$

$$\eta_l = \beta_l , \text{ para } x_l = 1 \text{ e}$$

$$\eta_q = \beta_q , \text{ para } x_q = 1 .$$

Logo, para obtenção de β_{jl} , deve-se utilizar uma mistura dos componentes j e l ($x_j = \frac{1}{2}$, $x_l = \frac{1}{2}$ e $x_q = 0$), em que:

$$\eta_{jl} = E(\eta | x_j = \frac{1}{2}, x_l = \frac{1}{2}, x_q = 0) .$$

Substituindo as condições da mistura dos componentes j e l ($x_j = \frac{1}{2}$, $x_l = \frac{1}{2}$ e $x_q = 0$) em (2.8), tem-se:

$$\eta_{jl} = \frac{1}{2}\beta_j + \frac{1}{2}\beta_l + \frac{1}{4}\beta_{jl} .$$

Logo,

$$\beta_{jl} = 4\eta_{jl} - 2\eta_j - 2\eta_l .$$

O mesmo procedimento deve ser utilizado para a obtenção dos coeficientes β_{jq} e β_{lq} , fornecendo

$$\beta_{jq} = 4\eta_{jq} - 2\eta_j - 2\eta_q \text{ e}$$

$$\beta_{lq} = 4\eta_{lq} - 2\eta_l - 2\eta_q .$$

Em uma malha{3,3}, tem-se:

η_j : resposta para o componente puro j ($x_j = 1$),

η_l : resposta para o componente puro l ($x_l = 1$),

η_q : resposta para o componente puro q ($x_q = 1$),

η_{jll} : resposta para a mistura dos componentes j e l ($x_j = \frac{2}{3}$ e $x_l = \frac{1}{3}$),

η_{jll} : resposta para a mistura dos componentes j e l ($x_j = \frac{1}{3}$ e $x_l = \frac{2}{3}$),

η_{jjq} : resposta para a mistura dos componentes j e q ($x_j = \frac{2}{3}$ e $x_q = \frac{1}{3}$),

η_{jjq} : resposta para a mistura dos componentes j e q ($x_j = \frac{1}{3}$ e $x_q = \frac{2}{3}$),

η_{llq} : resposta para a mistura dos componentes l e q ($x_l = \frac{2}{3}$ e $x_q = \frac{1}{3}$),

η_{llq} : resposta para a mistura dos componentes l e q ($x_l = \frac{1}{3}$ e $x_q = \frac{2}{3}$) e

η_{jll} : resposta para mistura dos componentes j, l e q ($x_j = \frac{1}{3}, x_l = \frac{1}{3}$ e $x_q = \frac{1}{3}$).

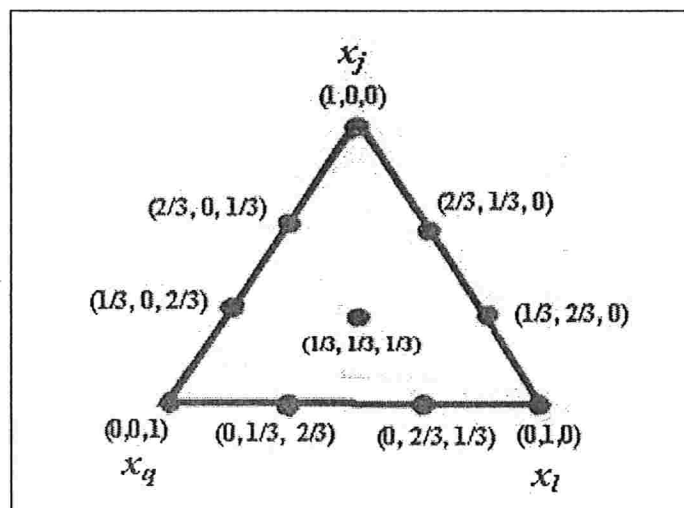


Figura 2.5: Malha{3,3}

Selecionando o polinômio (2.9) e considerando a equação (2.10) para os componentes j , l e q , os coeficientes β_{jl} , β_{jq} , β_{lq} , γ_{jl} , γ_{jq} , γ_{lq} e β_{jllq} são obtidos de acordo com o procedimento utilizado na malha $\{3,2\}$ e apresentados a seguir:

$$\beta_{jl} = \frac{9}{4}(\eta_{jll} + \eta_{jll} - \eta_j - \eta_l),$$

$$\beta_{jq} = \frac{9}{4}(\eta_{jjq} + \eta_{jjq} - \eta_j - \eta_q),$$

$$\beta_{lq} = \frac{9}{4}(\eta_{llq} + \eta_{llq} - \eta_l - \eta_q),$$

$$\gamma_{jl} = \frac{9}{4}(3\eta_{jll} - 3\eta_{jll} - \eta_j + \eta_l),$$

$$\gamma_{jq} = \frac{9}{4}(3\eta_{jjq} - 3\eta_{jjq} - \eta_j + \eta_q),$$

$$\gamma_{lq} = \frac{9}{4}(3\eta_{llq} - 3\eta_{llq} - \eta_l + \eta_q) e$$

$$\beta_{jllq} = 27\eta_{jllq} - \frac{27}{4}(\eta_{jll} + \eta_{jll} + \eta_{jjq} + \eta_{jjq} + \eta_{llq} + \eta_{llq}) + \frac{9}{2}(\eta_j + \eta_l + \eta_q).$$

Definidos a malha e o polinômio, caso este seja do segundo ou do terceiro grau, diversos caminhos podem ser adotados no planejamento do experimento, sendo um deles sugerido em Scheffé [1958]. Entre eles destaca-se a utilização de uma malha $\{k,p\}$, trabalhando possivelmente com mais que uma observação por ponto. A grande vantagem deste procedimento consiste na facilidade de obtenção dos estimadores de β e γ pelo método dos mínimos quadrados, pois as estimativas dos mesmos são calculadas através das médias das observações da variável resposta, $\hat{\eta}$, obtidas dos pontos da malha, e substituídas nas equações dos coeficientes β e γ em função de η .

2.5 Considerações Finais

Em algumas misturas, conforme veremos no próximo capítulo, devem ser consideradas restrições que limitam as proporções de alguns ou de todos os

componentes. Logo, sendo o limite superior da proporção do componente (S_j) e o limite inferior da proporção do componente (I_j), tem-se:

$$0 \leq I_j \leq x_j \leq S_j \leq 1, \quad 1 \leq j \leq k$$

Estas restrições adicionais, muitas vezes são responsáveis por um problema denominado de multicolinearidade, que gera efeitos indesejáveis principalmente na estimação dos parâmetros. A multicolinearidade pode ocorrer também em experimentos do tipo mistura que não apresentem restrições adicionais, assim uma vez identificada, seus efeitos devem ser minimizados. O próximo capítulo tratará do problema da multicolinearidade e dos seus efeitos nos experimentos do tipo mistura.

Capítulo 3

Multicolinearidade em experimentos do tipo mistura

3.1 Introdução

Os modelos de regressão das funções polinomiais em sua forma canônica podem ser representados na forma matricial como:

$$Y = X\beta + e,$$

sendo Y um vetor $n \times 1$ das observações da variável resposta, e é um vetor coluna $n \times 1$ dos erros, satisfazendo $E(e)=0$ e $V(e)=\sigma^2I$, em que I_n é a matriz identidade de ordem n . Na matriz X , de dimensões $n \times k$, cada linha possui os valores das k variáveis preditoras correspondentes a cada uma das observações e β é um vetor $k \times 1$ que contém os coeficientes de regressão

desconhecidos. De acordo com estas condições, o modelo quadrático na forma canônica proposto por Scheffé (1958) é dado por:

$$E(Y) = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \beta_{12} x_1 x_2 + \dots + \beta_{k-1,k} x_{k-1} x_k.$$

A estimação dos parâmetros pelo método dos mínimos quadrados fornece o seguinte estimador:

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

cujas respectivas matrizes de covariâncias é dada por:

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

Se a matriz X não é de posto completo, como consequência a matriz $X'X$ é singular, logo existe a necessidade da adoção de um novo modelo com matriz de planejamento de posto completo que satisfaça às condições. No entanto, um outro problema ocorre, quando a matriz é de posto completo, ou seja, existe a matriz inversa de $X'X$, porém há uma dependência linear aproximada entre as colunas da matriz X , isto é, existem constantes c_j ($j = 1, \dots, k$), tais que

$$\sum_{j=1}^k c_j x_j \approx 0 , \tag{3.1}$$

em que x_j é a j -ésima coluna da matriz X e 0 é o vetor nulo de dimensão n .

Esta condição apresentada em (3.1) é identificada como mau condicionamento ou multicolinearidade. A situação se torna mais complicada, neste caso, porque o cálculo da matriz inversa de $X'X$ é possível, mas muitas vezes os programas computacionais a realizam, sem qualquer indicação de um problema em potencial.

A presença da multicolinearidade origina alguns efeitos indesejados na estimação dos coeficientes de regressão pelo método dos mínimos quadrados. Por exemplo, em um modelo sem intercepto com apenas duas variáveis preditoras (W_1 e W_2) padronizadas :

$$z_i = \beta_1 w_{i1} + \beta_2 w_{i2} + e_i ,$$

em que o processo de padronização é realizado inclusive na variável resposta, representada por Z , e de acordo com a seguinte transformação:

$$w_{ij} = \frac{x_{ij} - \bar{x}_j}{\sqrt{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}}, \quad (3.2)$$

em que $i = 1, \dots, n$ e $j = 1, \dots, k$.

Por exemplo, no caso, $k = 2$, o modelo apresenta duas variáveis preditoras e n observações. As duas variáveis preditoras e a variável resposta padronizadas, são dadas por:

$$w_{i1} = \frac{x_{i1} - \bar{x}_1}{\sqrt{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2}},$$

$$w_{i2} = \frac{x_{i2} - \bar{x}_2}{\sqrt{\sum_{j=1}^n (x_{i2} - \bar{x}_2)^2}} \quad \text{e}$$

$$z_i = \frac{y_i - \bar{y}}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}.$$

O sistema de equações normais para a obtenção dos estimadores de mínimos quadrados de β_1 e β_2 é

$$(W'W)\hat{\beta} = W'Z,$$

$$\begin{pmatrix} 1 & r_{12} \\ r_{12} & 1 \end{pmatrix} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} r_{1z} \\ r_{2z} \end{pmatrix},$$

em que r_{12} é o coeficiente de correlação linear de Pearson entre W_1 e W_2 e r_{iz} é o correspondente coeficiente de correlação entre W_i e Z .

Assim,

$$(W'W)^{-1} = \begin{pmatrix} \frac{1}{(1-r_{12}^2)} & \frac{-r_{12}}{(1-r_{12}^2)} \\ \frac{-r_{12}}{(1-r_{12}^2)} & \frac{1}{(1-r_{12}^2)} \end{pmatrix}$$

e os estimadores de β_1 e β_2 são dados, respectivamente, por:

$$\hat{\beta}_1 = \frac{r_{1z} - r_{12}r_{2z}}{(1 - r_{12}^2)} \quad e$$

$$\hat{\beta}_2 = \frac{r_{2z} - r_{12}r_{1z}}{(1 - r_{12}^2)} .$$

Se existir forte multicolinearidade, o coeficiente de correlação r_{12} deve ser próximo de 1 ou -1. Portanto, através da matriz $(W'W)^{-1}$, verifica-se que se $|r_{12}| \rightarrow 1$, então $Var(\hat{\beta}_i) \rightarrow \infty$ e $Cov(\hat{\beta}_1, \hat{\beta}_2) \rightarrow \mp\infty$, dependendo se a correlação r_{12} tende a +1 ou -1.

Como consequência, dois efeitos devem ser destacados. O primeiro deles seria a geração de resultados imprecisos no cálculo da matriz inversa de $X'X$ pelos programas computacionais. O segundo efeito seria o aumento da variância dos estimadores dos parâmetros do modelo, mesmo quando a matriz inversa de $X'X$ é calculada corretamente. Este aumento na variância pode também gerar problemas para a significância das estimativas dos parâmetros ou interferir nos sinais das estimativas.

De acordo com St. John (1984), a multicolinearidade entre as variáveis preditoras pode conduzir a inconsistentes ou confusas conclusões quando comparados diferentes processos de seleção de variáveis do tipo “stepwise”. O processo de seleção de variáveis em presença da multicolinearidade apresenta um grande potencial para que equívocos sejam cometidos. Várias técnicas para detectar e identificar a multicolinearidade foram propostas e algumas serão discutidas a seguir.

Gorman (1970) sugeriu a utilização do coeficiente de determinação R_j^2 como medida de multicolinearidade. O coeficiente de determinação R_j^2 é obtido

através da regressão de X_j em função das variáveis preditoras restantes usando o modelo de regressão polinomial de Scheffé sem o termo constante. Grandes valores de R_j^2 indicam a presença de multicolinearidade. No entanto, segundo o autor, se $R_j^2 < 0,99$, então não se apresenta o problema.

É importante observar que, para modelos sem intercepto,

$$R_j^2 = \frac{x'_j X_{(j)} (X'_{(j)} X_{(j)})^{-1} X'_{(j)} x_j}{x'_j x_j} ,$$

em que $X_{(j)}$ é a matriz de planejamento obtida eliminando-se a coluna j e x'_j é o vetor linha, obtido através da transposição desta coluna.

Um segundo método de identificação da multicolinearidade foi sugerido por Marquardt (1970), em que o principal objetivo consistia no cálculo e análise do Fator de Inflação da Variância (FIV).

Os Fatores de Inflação da Variância são definidos como

$$FIV_j = \frac{1}{1-R_j^2} , \quad j = 1, \dots, k ,$$

em que, FIV_j é o Fator de Inflação da Variância relacionado ao estimador do parâmetro β_j .

Nota-se que em ambas as técnicas, a essência é a mesma, ou seja, a multicolinearidade indicada por valores de R_j^2 elevados também pode ser evidenciada através de altos valores dos Fatores de Inflação da Variância. De acordo com a definição, cada FIV_j é uma função do respectivo R_j^2 e conforme aumenta-se o coeficiente de determinação, tem-se como consequência um aumento do FIV. Portanto, evitar um grande R_j^2 é equivalente a evitar um grande FIV_j . De acordo com Marquardt (1970), para qualquer FIV com valor maior que dez, tem-se como consequência que os estimadores pelo método dos mínimos quadrados são imprecisos e diferentes modelos ou diferentes critérios de estimação são desejáveis. Por outro lado, Gorman (1970) sugere como critério de caracterização de multicolinearidade $R_j^2 > 0,99$, que é equivalente a $FIV > 100$, sendo, portanto, um critério mais brando.

A obtenção dos FIV pode ser feita através da matriz $(X'X)^{-1}$, em que os Fatores de Inflação da Variância correspondem aos elementos da diagonal principal desta matriz, de acordo com Montgomery, Peck e Vining (2006).

Na situação ideal de inexistência de multicolinearidade da variável X_j com as demais, $R_j^2 = 0$ e $FIV_j = 1$. Por outro lado, como $Var(\hat{\beta}) = (X'X)^{-1}\sigma^2$, segue que $FIV_j = \frac{var(\hat{\beta}_j)}{\sigma^2}$, e portanto, esta quantidade avalia quanto $Var(\hat{\beta}_j)$ é acrescida devido à multicolinearidade de X_j com as demais variáveis preditoras.

Um terceiro procedimento para identificação da multicolinearidade é através da observação dos autovalores da matriz $X'X$. Como o posto de uma matriz quadrada é igual ao número de autovalores diferentes de zero, então, se a matriz apresentar um autovalor igual a zero, trata-se de uma matriz de posto incompleto. No entanto, se a matriz apresentar todos os autovalores diferentes de zero e um ou mais autovalores próximos de zero, trata-se de uma matriz de posto completo, porém com indício de multicolinearidade.

O primeiro passo consiste na obtenção dos autovalores ou raízes características de $X'X$, que são representados por a_i sendo que em um modelo de k variáveis preditoras, a matriz $X'X$ é uma matriz quadrada de ordem k com k autovalores. Quando um ou mais autovalores apresentam magnitude reduzida, próxima de zero, tem-se um forte indício de uma dependência linear aproximada entre as colunas e muito provavelmente multicolinearidade. Alguns analistas preferem avaliar o parâmetro de condicionamento (c) de $X'X$ definido por:

$$c = \frac{a_{max}}{a_{min}}. \quad (3.3)$$

Este parâmetro mede a amplitude do espectro de autovalores $X'X$. De acordo com Montgomery, Peck e Vining (2006), se o parâmetro de condicionamento c é um número menor que 100, não há nenhum sério problema de multicolinearidade. No entanto, caso se encontre entre 100 e 1000, implica em multicolinearidade moderada ou forte e se o parâmetro exceder 1000, o problema é considerado severo.

Uma vez identificada a multicolinearidade, uma análise dos autovetores correspondentes associados aos autovalores próximos de zero possibilita a

determinação das restrições e da dependência linear aproximada que ocorre nas colunas da matriz X. Nos autovetores selecionados, os valores próximos de zero são descartados e como consequência são identificadas as restrições aproximadas, este procedimento está descrito com detalhes em Montgomery, Peck e Vining (2006, pág.302-303). Segundo Gunst e Mason (1980), esta análise de um autovetor associado a um autovalor próximo de zero é na essência similar ao processo de identificação de multicolinearidade através dos cálculos dos Fatores de Inflação da Variância. Um alto FIV só ocorre quando há um reduzido autovalor e as variáveis com altos valores do FIV estarão envolvidas na multicolinearidade.

Os experimentos do tipo mistura apresentam uma forte tendência à ocorrência da multicolinearidade, devido à restrição (2.2). Além disso, uma das características desses experimentos que tem como consequência o problema da multicolinearidade é o fato de que muitas vezes um dos componentes (x_k) apresentar seus possíveis valores limitados a um intervalo de variação muito reduzido e conseqüentemente a somatória das concentrações dos componentes restantes ($\sum_{i=1}^{k-1} x_i$) se aproxima de um valor constante, conduzindo o modelo à condição (3.1).

Logo, em um modelo de um experimento do tipo mistura com a função polinomial em sua forma canônica e alguns componentes com intervalos de variação reduzidos em relação a amplitude total de 0 até 1, a possibilidade de multicolinearidade deve ser considerada e analisada. O problema se agrava se a função polinomial na forma canônica for quadrática, pois a multiplicação entre duas variáveis desse tipo que originalmente apresentam intervalos de variação reduzidos gera um intervalo de menor amplitude para o produto.

Segundo Draper and Smith (1981), o condicionamento da matriz $X'X$ de uma forma geral apresenta melhorias quando as variáveis preditoras são padronizadas, como por exemplo, a padronização ocorrida em (3.2).

A padronização inicial das variáveis preditoras é necessária e ocorre de acordo com o procedimento destacado na equação (3.2). Logo, as novas variáveis preditoras padronizadas são definidas como W_1, W_2, \dots e W_k . Deste modo, o modelo escrito em função destas variáveis é:

$$Y = \alpha_w \cdot 1_n + W\beta_w + e ,$$

sendo 1_n um vetor $n \times 1$ com todos os elementos iguais a 1 e W a matriz associada às variáveis padronizadas.

Os novos parâmetros são representados pelo vetor β_w e um intercepto α_w é incluído no modelo desde que o vetor Y não esteja padronizado. Uma consequência importante deste modelo é que a matriz $W'W$ é a matriz de correlação, que apresenta as correlações entre as variáveis preditoras duas a duas. O vetor dos estimadores dos parâmetros β_w é dado por

$$\hat{\beta} = (W'W)^{-1}W'Y , \quad (3.4)$$

com

$$V(\hat{\beta}) = \sigma^2(W'W)^{-1} . \quad (3.5)$$

Como a forma canônica do modelo de regressão polinomial de Scheffé não possui intercepto, a padronização das variáveis preditoras é dada simplesmente por:

$$w_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^n x_{ij}^2}} , i = 1,2 \dots n \text{ e } j = 1,2 \dots k. \quad (3.6)$$

Logo, o modelo para as variáveis padronizadas é dado por:

$$Y = W\beta_w + \varepsilon .$$

Segundo Cornell (1990), na forma canônica do modelo polinomial quadrático, os produtos entre as variáveis preditoras podem ser padronizados de duas formas. A primeira consiste na padronização das variáveis preditoras, assim depois de padronizadas, as mesmas seriam multiplicadas gerando os produtos. A segunda é definida pela multiplicação das variáveis preditoras na forma original e os produtos são padronizados de acordo com a equação (3.7):

$$w_{ij}w_{iu} = \frac{x_{ij}x_{iu}}{\sqrt{\sum_{i=1}^n x_{ij}^2 \cdot x_{iu}^2}} , i = 1,2 \dots n \text{ e } j = 1,2 \dots k. \quad (3.7)$$

Uma diferença importante entre estas padronizações na forma canônica e a anterior é que a matriz $W'W$, neste caso do formato canônico, não é a matriz de correlação. No entanto, esta matriz, mesmo não sendo a matriz de correlação, é de fundamental importância para o cálculo do Fator de Inflação da Variância (FIV). Os elementos da diagonal principal da matriz $(W'W)^{-1}$ são definidos como os Fatores de Inflação da Variância.

Se a matriz $(W'W)^{-1}$ é resultante de um planejamento ortogonal, ou seja, não há dependência linear entre as variáveis preditoras, então $(W'W)^{-1}$ é a matriz identidade. Logo, o valor mínimo para o FIV é 1. Portanto, quanto maior o valor do FIV, em relação ao mínimo 1, maior o aumento da variância dos estimadores dos parâmetros β_{w_j} , como consequência da falta de ortogonalidade do modelo.

De acordo com Marquardt e Snee (1975), as padronizações removem a multicolinearidade não essencial, portanto os autores entendem que ainda devem analisar os autovalores da matriz $W'W$, que permitem identificar as condições da forma (3.1) que conduzem à multicolinearidade.

Gorman (1970) sugeriu que os experimentos do tipo mistura com alguns componentes apresentando intervalos muito reduzidos deveriam ser modelados através de termos denominados pseudocomponentes. O objetivo seria melhorar o condicionamento da matriz $X'X$, através da transformação algébrica nas variáveis preditoras definida a seguir:

$$x'_{ij} = \frac{(x_{ij} - l_j)}{(1 - \sum l_j)}, \text{ com } j=1,2,\dots,k,$$

em que l_j é o limite inferior do intervalo de variação das proporções do componente j .

Esta transformação frequentemente melhora o condicionamento da matriz $X'X$ e apresenta a vantagem de satisfazer a restrição de mistura da equação (2.2), ou seja, se originalmente a somatória das variáveis preditoras era constante e igual a 1, após a transformação, os pseudocomponentes mantêm a soma igual a 1, a demonstração deste fato encontra-se no Apêndice 2.

Os exemplos a seguir ilustram o processo de identificação do problema da multicolinearidade nos experimentos do tipo mistura e as primeiras medidas para saná-lo. Entre elas, destacam-se o processo de padronização e a utilização de pseudocomponentes.

3.2 Exemplos

Exemplo 1

O primeiro exemplo encontra-se em Snee e Marquardt (1976), nele destacam-se algumas estratégias para reduzir os efeitos do problema da multicolinearidade.

O estudo envolve um novo produto alimentício que pode ser produzido a partir de oito ingredientes, sendo que quatro deles (X_1 , X_2 , X_5 , X_6) devem estar presentes obrigatoriamente, por razões econômicas.. A variável resposta não foi especificada no artigo.

O modelo proposto consiste em uma função polinomial linear na forma canônica com oito componentes. Na Tabela 3.1, verifica-se que a amostra é composta por 20 misturas envolvendo os oito componentes e que suas respectivas proporções são limitadas de acordo com os intervalos a seguir:

$$0,10 \leq x_1 \leq 0,45$$

$$0,05 \leq x_2 \leq 0,50$$

$$0,00 \leq x_3 \leq 0,10$$

$$0,00 \leq x_4 \leq 0,10$$

$$0,10 \leq x_5 \leq 0,60$$

$$0,05 \leq x_6 \leq 0,20$$

$$0,00 \leq x_7 \leq 0,05$$

$$0,00 \leq x_8 \leq 0,05$$

Tabela 3.1: Amostra de 20 misturas com oito componentes

<i>Mistura</i>	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	y
1	0,100	0,500	0,000	0,000	0,100	0,200	0,050	0,050	30
2	0,100	0,050	0,000	0,000	0,550	0,200	0,050	0,050	113
3	0,100	0,500	0,000	0,100	0,100	0,200	0,000	0,000	17
4	0,150	0,050	0,000	0,100	0,600	0,050	0,050	0,000	94
5	0,100	0,050	0,100	0,000	0,550	0,200	0,000	0,000	89
6	0,100	0,500	0,100	0,100	0,100	0,050	0,000	0,050	18
7	0,100	0,050	0,100	0,100	0,550	0,050	0,000	0,050	90
8	0,400	0,050	0,100	0,100	0,100	0,200	0,050	0,000	20
9	0,350	0,050	0,100	0,100	0,100	0,200	0,050	0,050	21
10	0,300	0,500	0,000	0,000	0,100	0,050	0,000	0,050	15
11	0,100	0,500	0,100	0,000	0,200	0,050	0,050	0,000	28
12	0,450	0,050	0,000	0,000	0,450	0,050	0,000	0,000	48
13	0,450	0,200	0,000	0,100	0,100	0,050	0,050	0,050	18
14	0,450	0,150	0,000	0,100	0,100	0,200	0,000	0,000	7
15	0,450	0,250	0,100	0,000	0,100	0,050	0,050	0,000	16
16	0,450	0,100	0,100	0,000	0,100	0,200	0,000	0,050	19
17	0,259	0,222	0,050	0,050	0,244	0,125	0,025	0,025	38
18	0,259	0,222	0,050	0,050	0,244	0,125	0,025	0,025	30
19	0,259	0,222	0,050	0,050	0,244	0,125	0,025	0,025	35
20	0,259	0,222	0,050	0,050	0,244	0,125	0,025	0,025	40

Os componentes X_7 e X_8 apresentam intervalos de variação bastante reduzidos, que são indicadores de provável multicolinearidade neste modelo. Através do modelo linear polinomial na forma canônica, os autovalores da matriz $X^T X$ são calculados e os valores de máximo e mínimo são apresentados a seguir:

$$a_{\text{máx}} = 3,98281 \quad e$$

$$a_{\text{mín}} = 0,00977.$$

O cálculo do parâmetro de condicionamento da matriz $X'X$ é dado por:

$$c = \frac{3,98281}{0,00977} = 407,66$$

Logo, segundo Montgomery, Peck e Vining (2006), trata-se de um valor que indica moderada multicolinearidade, o modelo deve ser investigado e alterado para que o parâmetro de condicionamento seja reduzido.

Conforme comentado anteriormente, segundo Gorman (1970), os experimentos do tipo mistura com alguns componentes apresentando intervalos muito reduzidos deveriam ser modelados através de termos denominados pseudocomponentes. No modelo proposto por Snee e Marquardt (1976), as variáveis preditoras foram transformadas em pseudocomponentes e os valores extremos dos autovalores foram:

$$a_{\text{máx}} = 3,63688 \quad e$$

$$a_{\text{mín}} = 0,01994.$$

O parâmetro de condicionamento da matriz $X'X$ é igual a 182,39, indicando uma sensível melhora no condicionamento, logo uma redução na multicolinearidade.

No entanto, a maioria dos pacotes computacionais no processo de estimação por mínimos quadrados padroniza as variáveis preditoras antes de gerar a matriz $X'X$. Deste modo, para um modelo sem intercepto, a padronização usual é a proposta em (3.6). Logo, a análise dos autovalores da matriz $W'W$ torna-se a maneira mais apropriada para se avaliar o problema da multicolinearidade.

Serão analisados os FIV e os autovalores, após as variáveis preditoras serem padronizadas. Os valores da diagonal principal da matriz $(W'W)^{-1}$ são os Fatores de Inflação da Variância (FIV) e serão apresentados na Tabela 3.2.

Tabela 3.2: Fatores de Inflação da Variância do modelo linear na forma canônica após padronização

<i>Variável</i>	<i>FIV</i>
w ₁	3,13
w ₂	2,05
w ₃	2,08
w ₄	2,09
w ₅	2,04
w ₆	2,61
w ₇	2,16
w ₈	2,21

Os autovalores máximo e mínimo da matriz $W^T W$ são:

$$a_{\text{máx}} = 5,18040 \quad \text{e}$$

$$a_{\text{mín}} = 0,21113.$$

Verifica-se que a padronização promove uma nítida redução do problema da multicolinearidade. Um indicador é o parâmetro de condicionamento que foi reduzido para 24,536, que, conforme discutido anteriormente, quando abaixo de 100, sugere que não haja problemas de multicolinearidade. Além disso, todos os valores dos FIV reforçam essa conclusão.

Tabela 3.3: Fatores de Inflação da Variância do modelo linear utilizando a transformação em pseudocomponentes, antes da padronização

<i>Variável</i>	<i>FIV</i>
w'_1	1,97
w'_2	1,76
w'_3	2,01
w'_4	2,02
w'_5	1,48
w'_6	1,94
w'_7	2,11
w'_8	2,17

Na análise da Tabela 3.3, em que as variáveis foram transformadas em pseudocomponentes antes da padronização, a principal conclusão é que não ocorre o problema da multicolinearidade, ou seja, a mesma do modelo anterior, em que houve apenas o processo de padronização.

Logo, torna-se evidente que o problema da multicolinearidade era originado pela falta de padronização das variáveis preditoras. Uma vez padronizadas, a multicolinearidade é solucionada e a transformação em pseudocomponentes é desnecessária, pois, segundo os autores, não há nenhuma melhoria significativa no condicionamento da matriz $W'W$. No entanto, através de um estudo mais detalhado do modelo com pseudocomponentes, percebe-se uma sutil melhoria no condicionamento, já que todos os respectivos FIV são menores na Tabela 3.3, quando comparados com os da Tabela 3.2 e o parâmetro de condicionamento reduz-se a 14,922.

Exemplo 2

O segundo exemplo foi publicado por Piepel (1982). Os dados apresentados na Tabela 3.4 correspondem a uma mistura de onze componentes químicos de uma particular formulação de vidro em que a variável resposta é a lixiviabilidade do vidro, medida através da perda de massa percentual do vidro.

O teste de lixiviação é o mais utilizado na determinação da estabilidade química e do potencial poluidor de um material destinado a resíduo. Entende-se por lixiviabilidade a capacidade de transferência dos contaminantes do resíduo para a solução lixivante. A solução lixivante é a solução utilizada no ensaio de lixiviação e o lixiviado é a solução contaminada pelo contato com o resíduo (CHAMIE, 1994). A lixiviabilidade é determinada expondo-se o resíduo, tratado ou não, a uma solução com características conhecidas e determinando-se o grau de dissolução dos contaminantes. O ensaio de lixiviação procura reproduzir, em laboratório, os fenômenos de arraste, de diluição e de dessorção que ocorrem pela passagem de água através do resíduo. O método de ensaio é descrito pela Norma da ABNT NBR 10005 – Lixiviação de Resíduos.

Tabela 3.4: Amostra de 44 misturas com onze componentes químicos de uma particular formulação de vidro, variável resposta: lixiviabilidade do vidro.

X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	Y
SiO ₂	B ₂ O ₃	Al ₂ O ₃	CaO	MgO	Na ₂ O	ZnO	TiO ₂	Cr ₂ O ₃	Fe ₂ O ₃	NiO	
0,448	0,063	0,142	0,000	0,000	0,165	0,059	0,069	0,000	0,029	0,025	8,10
0,544	0,063	0,000	0,138	0,000	0,154	0,000	0,071	0,030	0,000	0,000	0,14
0,565	0,120	0,000	0,000	0,000	0,153	0,061	0,070	0,031	0,000	0,000	0,01
0,581	0,061	0,012	0,000	0,079	0,108	0,061	0,070	0,000	0,028	0,000	0,01
0,432	0,126	0,000	0,098	0,079	0,158	0,000	0,070	0,000	0,009	0,028	7,40
0,486	0,064	0,000	0,140	0,000	0,159	0,063	0,000	0,028	0,031	0,029	3,90
0,429	0,060	0,145	0,139	0,013	0,152	0,062	0,000	0,000	0,000	0,000	91,4
0,429	0,124	0,144	0,138	0,000	0,105	0,060	0,000	0,000	0,000	0,000	49,0
0,443	0,120	0,000	0,099	0,080	0,101	0,061	0,070	0,000	0,026	0,000	0,11
0,447	0,114	0,148	0,012	0,077	0,150	0,000	0,000	0,027	0,025	0,000	33,40
0,427	0,121	0,145	0,010	0,078	0,161	0,000	0,000	0,029	0,029	0,000	39,50
0,550	0,060	0,146	0,000	0,076	0,111	0,000	0,030	0,000	0,000	0,027	0,42
0,434	0,058	0,148	0,127	0,000	0,109	0,000	0,069	0,030	0,025	0,000	9,60
0,580	0,122	0,000	0,135	0,000	0,110	0,000	0,000	0,000	0,025	0,028	0,20
0,575	0,123	0,000	0,134	0,000	0,112	0,000	0,000	0,000	0,028	0,028	0,24
0,498	0,066	0,000	0,134	0,080	0,105	0,059	0,000	0,030	0,000	0,028	6,20
0,435	0,118	0,145	0,000	0,031	0,102	0,060	0,051	0,031	0,000	0,027	2,20
0,496	0,092	0,072	0,067	0,038	0,125	0,031	0,035	0,015	0,015	0,014	0,86
0,497	0,091	0,072	0,066	0,039	0,129	0,029	0,034	0,014	0,015	0,014	0,80
0,497	0,091	0,072	0,067	0,039	0,127	0,030	0,034	0,014	0,015	0,014	0,86
0,565	0,125	0,000	0,010	0,036	0,163	0,000	0,070	0,031	0,000	0,000	0,20
0,428	0,068	0,073	0,117	0,084	0,108	0,000	0,071	0,000	0,027	0,024	1,47
0,433	0,064	0,155	0,047	0,000	0,159	0,000	0,073	0,028	0,022	0,019	16,4
0,550	0,141	0,000	0,000	0,000	0,156	0,063	0,079	0,000	0,011	0,000	0,15
0,577	0,127	0,023	0,000	0,086	0,094	0,060	0,000	0,033	0,000	0,000	0,15
0,586	0,126	0,024	0,000	0,079	0,093	0,058	0,000	0,034	0,000	0,000	0,08
0,585	0,128	0,023	0,000	0,082	0,091	0,059	0,000	0,032	0,000	0,000	0,15
0,579	0,068	0,000	0,000	0,088	0,146	0,061	0,000	0,000	0,031	0,027	0,18
0,541	0,064	0,153	0,072	0,000	0,113	0,000	0,000	0,031	0,026	0,000	2,04
0,417	0,127	0,152	0,010	0,080	0,162	0,000	0,000	0,031	0,000	0,021	51,8
0,429	0,133	0,031	0,130	0,000	0,144	0,059	0,074	0,000	0,000	0,000	9,04
0,454	0,067	0,142	0,000	0,000	0,145	0,060	0,077	0,033	0,000	0,022	2,09
0,420	0,130	0,138	0,000	0,083	0,155	0,000	0,074	0,000	0,000	0,000	24,7
0,574	0,066	0,000	0,114	0,000	0,108	0,000	0,077	0,034	0,000	0,027	0,02
0,584	0,067	0,031	0,128	0,000	0,165	0,000	0,000	0,000	0,025	0,000	0,08
0,464	0,067	0,142	0,126	0,083	0,118	0,000	0,000	0,000	0,000	0,000	37,2
0,433	0,134	0,142	0,038	0,000	0,093	0,058	0,072	0,000	0,030	0,000	7,39
0,432	0,134	0,012	0,129	0,083	0,096	0,059	0,000	0,000	0,030	0,025	96,6
0,434	0,068	0,127	0,131	0,000	0,153	0,061	0,000	0,000	0,000	0,026	77,8
0,432	0,067	0,138	0,131	0,000	0,146	0,060	0,000	0,000	0,000	0,026	75,6
0,431	0,067	0,139	0,132	0,000	0,149	0,061	0,000	0,000	0,000	0,021	76,4
0,413	0,127	0,151	0,011	0,080	0,165	0,000	0,000	0,030	0,000	0,023	43,3
0,430	0,133	0,144	0,127	0,000	0,111	0,000	0,000	0,031	0,000	0,024	25,4
0,596	0,068	0,000	0,011	0,000	0,136	0,063	0,077	0,024	0,000	0,025	0,07

Na Tabela 3.4, a coluna da variável resposta (Y) correspondente às medidas de lixiviabilidade contém os valores das perdas de massa em porcentagem das respectivas formulações dos vidros, após o ensaio de lixiviação.

As proporções dos componentes nas misturas especificadas para o modelo são limitadas de acordo com os intervalos a seguir:

$$0,41 \leq x_1 \leq 0,60$$

$$0,055 \leq x_2 \leq 0,15$$

$$0 \leq x_3 \leq 0,16$$

$$0 \leq x_4 \leq 0,14$$

$$0 \leq x_5 \leq 0,09$$

$$0,09 \leq x_6 \leq 0,17$$

$$0 \leq x_7 \leq 0,065$$

$$0 \leq x_8 \leq 0,080$$

$$0 \leq x_9 \leq 0,035$$

$$0 \leq x_{10} \leq 0,035$$

$$0 \leq x_{11} \leq 0,035$$

$$0,54 \leq x_1 + x_2 + x_3 \leq 0,80$$

$$0,13 \leq x_4 + x_5 + x_6 \leq 0,35.$$

O fato de todos os componentes possuírem limites inferior e superior para a suas respectivas proporções já seria um bom indicativo da existência do problema da multicolinearidade.

A Tabela 3.5 apresenta os FIV para a matriz padronizada W^*W gerada a partir dos dados da Tabela 3.4.

Tabela 3.5: Fatores de Inflação da Variância do modelo linear na forma canônica após padronização

<i>Variável</i>	<i>FIV</i>
w_1	30,68
w_2	10,65
w_3	2,69
w_4	3,04
w_5	2,48
w_6	24,45
w_7	2,44
w_8	2,26
w_9	2,61
w_{10}	1,98
w_{11}	2,26

Os FIV das variáveis w_1 , w_2 e w_6 , razoavelmente elevados, e o parâmetro de condicionamento, igual a 320,363, indicam a presença de multicolinearidade moderada no modelo.

Nota-se, na Tabela 3.5, que os maiores valores para o parâmetro FIV não estão associados às variáveis que possuem os intervalos mais restritos (x_9 , x_{10} e x_{11}), mas sim às variáveis que possuem o limite inferior afastado do zero (x_1 , x_2 e x_6). Este fato justifica-se, pois no processo de padronização, as variáveis não padronizadas que possuem intervalos estreitos e limite inferior igual ou próximo de zero apresentam a soma dos quadrados não corrigida muito reduzida. Conseqüentemente, estas mesmas variáveis padronizadas terão um aumento em seus intervalos de variação. Por outro lado, as variáveis não padronizadas que possuem o limite inferior afastado do zero apresentam a soma dos quadrados não corrigida muito maior quando comparada à correspondente

soma para as variáveis limitadas e próximas de zero. Logo, com a padronização, seus intervalos de variação podem ser reduzidos. Como, quanto maior o intervalo de variação menor o FIV, pode-se concluir que houve uma redução nos intervalos de x_1 , x_2 e x_6 e seus respectivos FIV são indicadores de multicolinearidade.

Portanto, neste caso, a transformação em pseudocomponentes pode ser interessante, e o mais importante para a melhoria do condicionamento da matriz W^TW seria que todos os componentes após a transformação tivessem limite inferior igual a zero. A Tabela 3.6 fornece os FIV do modelo utilizando a transformação em pseudocomponentes antes da padronização.

Tabela 3.6: Fatores de Inflação da Variância do modelo linear utilizando a transformação em pseudocomponentes, antes da padronização

<i>Variável</i>	<i>FIV</i>
w'_1	3,00
w'_2	2,77
w'_3	2,70
w'_4	2,21
w'_5	2,08
w'_6	3,44
w'_7	2,07
w'_8	1,90
w'_9	2,37
w'_{10}	1,88
w'_{11}	2,20

Na Tabela 3.6, verifica-se que todos os FIV são menores que 3,50, já os autovalores geram um parâmetro de condicionamento igual a 35,877. Logo, de acordo com Piepel (1982), para o modelo polinomial linear na forma canônica, verificou-se que a transformação em pseudocomponentes antes da padronização, gerou um ajuste em que a multicolinearidade é bastante reduzida.

No entanto, o autor não optou pelo modelo polinomial linear e após o processo de seleção de variáveis do tipo “stepwise”, decidiu por um modelo do segundo grau, em que foram adicionados produtos de variáveis, conforme a equação a seguir:

$$E(Y) = \sum_{j=1}^{11} \beta_j x'_j + \beta_{6,8} x'_6 x'_8 + \beta_{3,11} x'_3 x'_{11} + \beta_{7,9} x'_7 x'_9 + \beta_{5,6} x'_5 x'_6 + \beta_{7,11} x'_7 x'_{11} + \beta_{1,5} x'_1 x'_5 + \beta_{6,10} x'_6 x'_{10} + \beta_{5,8} x'_5 x'_8 . \quad (3.8)$$

As variáveis indicadas por x'_i , na equação (3.8), encontram-se transformadas em pseudocomponentes. Os FIV associados a este modelo são dados na Tabela 3.7, após a padronização.

Tabela 3.7: Fatores de Inflação da Variância para as variáveis predictoras no modelo do segundo grau com a transformação em pseudocomponentes antes da padronização

<i>Variável</i>	<i>FIV</i>
w'_1	15,02
w'_2	8,82
w'_3	8,06
w'_4	14,28
w'_5	22,70
w'_6	67,19
w'_7	18,25
w'_8	15,39
w'_9	8,09
w'_{10}	11,81
w'_{11}	20,24
$w'_6 w'_8$	21,82
$w'_3 w'_{11}$	21,71
$w'_7 w'_9$	7,17
$w'_5 w'_6$	45,40
$w'_7 w'_{11}$	8,21
$w'_1 w'_5$	12,60
$w'_6 w'_{10}$	10,08
$w'_5 w'_9$	7,46

Nota-se, na Tabela 3.7, que o maior FIV está associado ao pseudocomponente padronizado w'_6 e que este encontra-se em três dos oito produtos de variáveis do modelo. Logo, é bastante compreensível que o segundo maior FIV também envolva w'_6 . Um dos mecanismos para tentar solucionar o problema seria remover um dos produtos que envolvem w'_6 . A sugestão de retirada recai sobre o produto $w'_6 \cdot w'_8$, pois apesar de não apresentar o maior FIV, verifica-se que o seu coeficiente não é significativo a um nível de 0,05. Quando este termo é removido do modelo, o parâmetro de condicionamento melhora, passa de 1718,06 para 766,94 e o FIV máximo é reduzido de 67,189 para 28,169, ambos correspondentes ao pseudocomponente w'_6 . Portanto, segundo Piepel (1982), embora o problema da multicolinearidade não tenha sido totalmente eliminado, houve uma sensível melhora no nível de condicionamento do modelo, considerando a natureza dos dados iniciais.

Exemplo 3

O terceiro exemplo foi proposto por Snee (1975). O objetivo do estudo foi determinar a quantidade de aditivo (X_1) necessária em uma mistura de três lubrificantes (X_2 , X_3 e X_4), de modo que uma determinada propriedade física da mistura atingisse um determinado nível. A medida dessa propriedade é a variável resposta do modelo (Y).

Os dados estão apresentados na Tabela 3.8 e correspondem a uma amostra de 18 misturas de quatro componentes e as proporções dos componentes nas misturas especificadas para o modelo são limitadas de acordo com os intervalos a seguir:

$$0,07 \leq x_1 \leq 0,18$$

$$0 \leq x_2 \leq 0,30$$

$$0,37 \leq x_3 \leq 0,70$$

$$0 \leq x_4 \leq 0,15$$

Tabela 3.8: Dados amostrais de 18 misturas com quatro componentes

<i>Mistura</i>	x_1	x_2	x_3	x_4	y
1	0,15	0,00	0,70	0,15	13,89
2	0,18	0,30	0,37	0,15	13,99
3	0,07	0,23	0,70	0,00	7,6
4	0,07	0,08	0,70	0,15	9,45
5	0,18	0,12	0,70	0,00	12,93
6	0,07	0,30	0,63	0,00	7,38
7	0,07	0,30	0,48	0,15	8,58
8	0,18	0,00	0,67	0,15	15,65
9	0,18	0,30	0,52	0,00	11,94
10	0,18	0,00	0,70	0,12	15,24
11	0,07	0,2275	0,6275	0,075	8,24
12	0,18	0,144	0,592	0,084	13,84
13	0,125	0,30	0,50	0,075	10,08
14	0,13	0,086	0,70	0,084	11,48
15	0,125	0,2375	0,6375	0,00	9,64
16	0,13	0,136	0,584	0,15	11,94
17	0,133	0,163	0,617	0,087	11,25
18	0,18	0,15	0,52	0,15	14,65

O modelo selecionado consiste em uma função polinomial do segundo grau, conforme a equação a seguir:

$$E(Y) = \sum_{j=1}^4 \beta_j x_j + \beta_{1,2} x_1 x_2 + \beta_{1,3} x_1 x_3 + \beta_{1,4} x_1 x_4 + \beta_{2,3} x_2 x_3 + \beta_{2,4} x_2 x_4 + \beta_{3,4} x_3 x_4.$$

De acordo com as conclusões do exemplo anterior, feita a padronização, são esperados altos valores para os FIV das variáveis X_1 e X_3 e também,

possivelmente, para os produtos de variáveis em que pelo menos uma delas esteja envolvida. Na Tabela 3.9, encontram-se os FIV do modelo.

Tabela 3.9: FIV para os dez termos no modelo do segundo grau padronizado

<i>Variável</i>	<i>FIV</i>
w_1	7142,86
w_2	250,00
w_3	196,08
w_4	1587,30
w_1w_2	411,52
w_1w_3	4761,90
w_1w_4	178,57
w_2w_3	256,41
w_2w_4	65,36
w_3w_4	819,67

Os maiores FIV correspondem realmente aos coeficientes de w_1 e w_1w_3 . Nota-se também que este fator para a variável X_3 não é tão elevado, como era inicialmente esperado. Este fato pode ser justificado devido ao intervalo de variação de X_3 ser relativamente grande quando comparado aos outros componentes, compensando seu limite inferior afastado do zero. O parâmetro de condicionamento igual a 95886,07 reforça o diagnóstico de que existe um severo problema de multicolinearidade.

Neste modelo, como duas das quatro variáveis apresentam seus limites inferiores afastados do zero, uma boa estratégia para a solução do problema da multicolinearidade é a transformação em pseudocomponentes.

Aplicada esta transformação, os resultados obtidos encontram-se na Tabela 3.10.

Tabela 3.10: FIV para as variáveis preditoras no modelo do segundo grau com transformação em pseudocomponentes antes da padronização

<i>Variável</i>	<i>FIV</i>
w'_1	761,53
w'_2	34,45
w'_3	41,56
w'_4	369,89
$w'_1w'_2$	119,23
$w'_1w'_3$	234,81
$w'_1w'_4$	56,65
$w'_2w'_3$	35,20
$w'_2w'_4$	65,57
$w'_3w'_4$	136,64

A partir desta tabela, observa-se que a transformação em pseudocomponentes provocou uma grande melhora no condicionamento da matriz W^TW , mas muitos FIV ainda indicam multicolinearidade, mesmo quando são analisados de forma mais branda. O parâmetro de condicionamento reduziu-se bruscamente a 7645,78, porém indica ainda alta multicolinearidade.

Em seu artigo, Snee (1975) não trabalha com a transformação em pseudocomponentes e as previsões são realizadas sempre dentro da região experimental. St. John (1984) mostrou que a transformação em pseudocomponentes neste modelo traria uma grande melhoria na matriz W^TW a um custo muito reduzido. No entanto, neste modelo, novas técnicas para mitigar o problema da multicolinearidade devem ser introduzidas, pois os recursos da padronização e a transformação em pseudocomponentes não foram suficientes para que se atingisse um nível de condicionamento aceitável.

Caso a multicolinearidade não seja totalmente eliminada, os autores apontam alguns cuidados a serem tomados na modelagem, como tentar

interpretar coeficientes individualmente, principalmente aqueles que possuem elevado FIV e não fazer previsões além da região experimental.

Snee e Rayner (1982) explicam como obter o polinômio na forma canônica, a partir de seu modelo com intercepto. No entanto, neste processo, pode-se apontar duas deficiências. A primeira é que existe uma maior afinidade com a interpretação das estimativas geradas pelo modelo polinomial na forma canônica proposto por Scheffé. Logo, os próprios autores recomendam a transformação para o modelo de Scheffé, a partir do modelo com intercepto. A segunda é que o modelo com intercepto tem como principal função eliminar os erros de arredondamento das rotinas computacionais e não o problema da multicolinearidade, que muitas vezes é uma sintomática característica dos dados que não será sanada apenas com a proposta de Snee e Rayner (1982).

Antes de tratar de novas técnicas para solucionar o problema da multicolinearidade, é importante destacar o modelo de uma malha simplex com quatro componentes quando adotada uma função polinomial do segundo grau na forma canônica. Segundo Kiefer (1961), este planejamento é D-ótimo para a função polinomial do segundo grau na forma canônica.

De acordo com Ryan (1997, p.459), um planejamento é denominado D-ótimo se o determinante da matriz $X'X$ é maximizado para uma dada região experimental. Esta característica é importante porque tal planejamento minimiza o volume do elipsóide de confiança para os parâmetros do modelo de regressão.

Nas Tabelas 3.11 e 3.12 são apresentados os dados da malha simplex para quatro componentes para uma função polinomial do segundo grau na forma canônica e os valores dos FIV, respectivamente.

No entanto, apesar das características ideais, o mínimo teórico do FIV não é atingido neste modelo, conforme verifica-se na Tabela 3.12. Logo, esta constatação sugere que a classificação de um FIV como elevado, deve possuir como referencial os FIV de modelos com boas propriedades, como o da malha simplex de quatro componentes, e não simplesmente levar em consideração o mínimo teórico.

Tabela 3.11: Malha Simplex para uma função polinomial do segundo grau com quatro componentes

x_1	x_2	x_3	x_4
1,0	0,0	0,0	0,0
0,0	1,0	0,0	0,0
0,0	0,0	1,0	0,0
0,0	0,0	0,0	1,0
0,5	0,5	0,0	0,0
0,5	0,0	0,5	0,0
0,5	0,0	0,0	0,5
0,0	0,5	0,5	0,0
0,0	0,5	0,0	0,5
0,0	0,0	0,5	0,5

Tabela 3.12: FIV do modelo polinomial do segundo grau para o planejamento simplex

<i>Variável</i>	<i>FIV</i>
x_1	1,75
x_2	1,75
x_3	1,75
x_4	1,75
x_1x_2	1,50
x_1x_3	1,50
x_1x_4	1,50
x_2x_3	1,50
x_2x_4	1,50
x_3x_4	1,50

Apesar das tentativas apresentadas até o momento para solucionar o problema da multicolinearidade, verificou-se que em alguns modelos a transformação em pseudocomponentes e a padronização das variáveis preditoras não foram suficientes para se atingir um condicionamento satisfatório na matriz de planejamento. Logo, ainda há a necessidade de novas técnicas, para que alguns modelos atinjam níveis aceitáveis nos parâmetros que medem a multicolinearidade. Na próxima seção, será discutida a técnica de regressão em cristas aplicada à análise estatística dos experimentos do tipo mistura.

3.3 Uso do procedimento de regressão em cristas nos experimentos do tipo mistura

O uso do procedimento de regressão em cristas nos experimentos do tipo mistura tem como objetivo melhorar as estimativas dos parâmetros do modelo diante do problema da multicolinearidade e das possibilidades já discutidas neste capítulo.

Uma consequência importante do método dos mínimos quadrados é que o estimador de β é não viciado. A propriedade de Gauss-Markoff garante que o estimador de mínimos quadrados possui variância mínima dentro da classe dos estimadores lineares não viciados de β , no entanto, não há garantia que esse valor mínimo seja de magnitude reduzida. Como já discutido anteriormente, uma variância elevada inviabiliza uma boa interpretação das estimativas dos parâmetros do modelo e torna os estimadores mais instáveis, suscetíveis a grandes variações, diante de diferentes amostras do mesmo tamanho ou pequenas alterações nas variáveis preditoras diferentes.

Um dos recursos para mitigar este problema consiste na adoção de um estimador viciado de β , porém com menor variância e um viés relativamente reduzido.

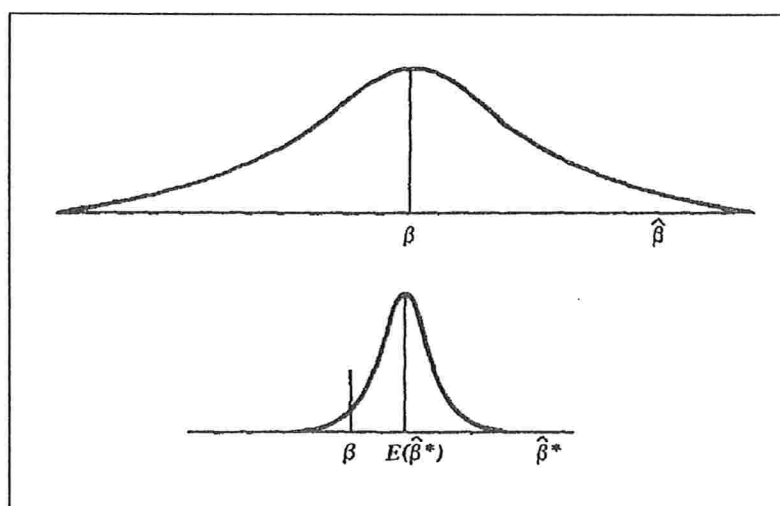


Figura 3.1

Na Figura 3.1, temos a situação de um estimador viciado ($\hat{\beta}^*$) com uma variância relativamente menor quando comparada ao estimador não viciado.

O erro quadrático médio de $\hat{\beta}^*$ é a esperança do quadrado da distância entre $\hat{\beta}^*$ e β , com

$$EQM(\hat{\beta}^*) = E(\hat{\beta}^* - \beta)^2 = V(\hat{\beta}^*) + [E(\hat{\beta}^*) - \beta]^2 . \quad (3.9)$$

Na equação 3.9, verifica-se que o EQM é a soma da variância do estimador viciado com o quadrado do viés de $\hat{\beta}^*$. Logo, se o viés não possuir um valor alto, EQM ($\hat{\beta}^*$) pode ser consideravelmente menor que o EQM do estimador não viciado, assim a sua utilização torna-se prioritária. Deve-se ressaltar que EQM do estimador não viciado coincide com a própria variância.

Alguns procedimentos foram desenvolvidos para a obtenção de estimadores viciados dos coeficientes de regressão com reduzidos EQM. Um destes procedimentos é a regressão em cristas, originalmente proposta por Hoerl e Kennard (1970).

O estimador do método de regressão em cristas é encontrado através da solução de um conjunto de equações levemente modificado das equações normais do método dos mínimos quadrados,

$$(X'X + \lambda I) \hat{\beta}_C = X'Y$$

de modo que

$$\hat{\beta}_C = (X'X + \lambda I)^{-1} X'Y , \quad (3.10)$$

sendo $\hat{\beta}_C$ o estimador em cristas. Verifica-se, através da equação (3.10), que $\hat{\beta}_C$ é um estimador viciado. Este fato pode ser verificado, pois através da equação matricial do estimador não viciado de β tem-se:

$$(X'X) \hat{\beta} = X'Y. \quad (3.11)$$

Substituindo (3.11) em (3.10) tem-se:

$$\hat{\beta}_C = (X'X + \lambda I)^{-1} (X'X) \hat{\beta}.$$

Logo,

$$E(\hat{\beta}_C) = (X'X + \lambda I)^{-1} (X'X) \beta.$$

Deste modo, torna-se evidente que $\hat{\beta}_C$ é um estimador viciado.

A constante λ é um parâmetro de viés e quando assume o valor zero, o estimador em cristas coincide com o estimador de mínimos quadrados.

Na obtenção do erro quadrático médio de $\hat{\beta}_C$, que é uma quantidade $(k+1)$ dimensional, a literatura utiliza o “Erro Quadrático Médio Total”, definido como

$$EQM(\hat{\beta}_C) = E [(\hat{\beta}_C - \beta)'(\hat{\beta}_C - \beta)].$$

De acordo com Montgomery, Peck e Vining (2006), o erro quadrático médio total do estimador em cristas é dado pela expressão:

$$EQM(\hat{\beta}_C) = \sigma^2 \sum_{j=1}^k \frac{a_j}{(a_j + \lambda)^2} + \lambda^2 \beta'(X'X + \lambda I)^{-2} \beta \quad (3.12)$$

Na equação (3.12), o primeiro termo corresponde à soma das variâncias das componentes de $\hat{\beta}_C$ e o segundo termo é o quadrado do viés. Também destacam-se os valores de a_j , que são os autovalores da matriz $X'X$. Da expressão (3.12) verifica-se que se $\lambda > 0$, quanto maior o valor de λ , menor a soma das variâncias e maior o quadrado do viés. Logo, a essência do método de regressão em cristas consiste em escolher um valor de λ , para o qual a redução da soma das variâncias seja maior que o aumento do quadrado do viés. Se isto puder ser feito, o erro quadrático médio do estimador em cristas $\hat{\beta}_C$ será menor que a variância do estimador $\hat{\beta}$ obtido pelo método dos mínimos quadrados.

Hoerl e Kennard (1970) provaram a existência de um valor de λ , diferente de zero, em que o EQM do estimador em cristas é menor que a variância do estimador obtido pelo método dos mínimos quadrados.

Um outro aspecto importante refere-se ao coeficiente de determinação (R^2), pois o aumento do parâmetro λ provoca um aumento na soma dos quadrados dos resíduos. Logo, como a soma de quadrados total é fixa, R^2 decresce com o aumento de λ . Deste modo, estimar através do processo de regressão em cristas não consiste simplesmente em selecionar um modelo com alto coeficiente de determinação, o interesse maior encontra-se na obtenção de um conjunto estável de estimativas dos parâmetros.

Um apropriado valor de λ pode ser encontrado através do estudo do gráfico denominado traço de crista, de acordo com Hoerl e Kennard [1970]. Trata-se de um gráfico de $\hat{\beta}_C$ contra λ , para valores de λ usualmente entre 0 e 1. Marquardt e Snee (1975) sugerem o uso de cerca de 25 valores, espaçados em escala logarítmica dentro do intervalo (0,1), como candidatos a “estimar” λ .

Nos casos em que a multicolinearidade é severa, a instabilidade dos coeficientes de regressão será óbvia neste gráfico. No entanto, através dele, também percebe-se que conforme o valor de λ aumenta, os coeficientes tendem a se estabilizar. Portanto, deve ser selecionado, aproximadamente, o menor valor de λ de modo que os coeficientes sejam estáveis. Logicamente, espera-se que o conjunto de estimativas selecionado apresente EQM menor que o dos mínimos quadrados.

Os FIV, além de atuarem como parâmetros de condicionamento, também podem ser interpretados como medidas da estabilidade das estimativas. Na regressão em cristas, quando as variáveis encontram-se padronizadas, o estimador em cristas é formulado por

$$\hat{\beta}_C = (W^T W + \lambda I)^{-1} W^T Y$$

e os FIV são os elementos da diagonal principal da matriz

$$(W^T W + \lambda I)^{-1} W^T W (W^T W + \lambda I)^{-1}.$$

De acordo com Gunst e Mason (1980), quando β é estimado por mínimos quadrados, os FIV correspondem aos elementos da diagonal principal de Q , em que $Var(\hat{\beta}) = Q\sigma^2$ para as variáveis preditoras padronizadas.

No caso da regressão ridge,

$$\text{Var}(\hat{\beta}_C) = (W^T W + \lambda I)^{-1} W^T W (W^T W + \lambda I)^{-1} \sigma^2$$

e portanto, por analogia, os FIV são os elementos da diagonal principal de

$$(W^T W + \lambda I)^{-1} W^T W (W^T W + \lambda I)^{-1}.$$

No terceiro exemplo deste capítulo, de acordo com Snee (1975), a proposta era determinar o comportamento de uma certa propriedade física em função da composição de uma mistura de três lubrificantes e um aditivo, verificou-se que a padronização e a transformação em pseudocomponentes não eliminaram o problema da multicolinearidade.

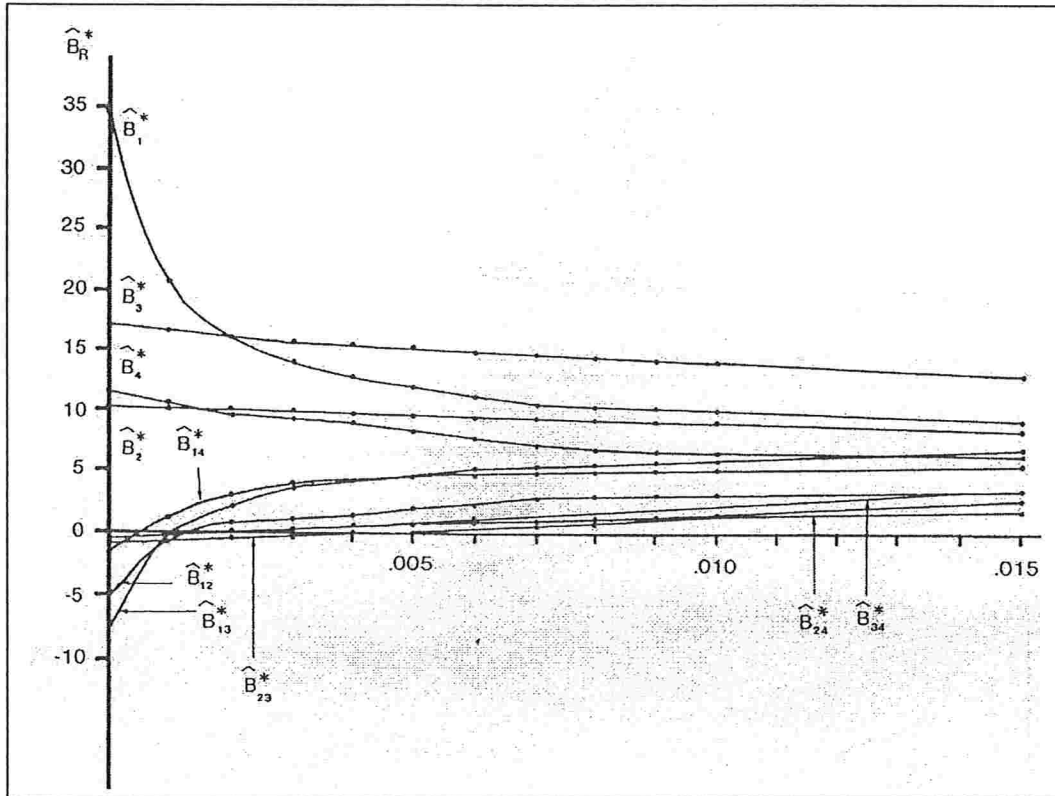
Deste modo, o processo de regressão em cristas foi introduzido com o intuito principal de reduzir a sensibilidade das estimativas dos parâmetros em relação a diferentes amostras.

O modelo polinomial do segundo grau definido para este experimento será rerepresentado a seguir :

$$E(Y) = \sum_{j=1}^4 \beta_j x_j + \beta_{1,2} x_1 x_2 + \beta_{1,3} x_1 x_3 + \beta_{1,4} x_1 x_4 + \beta_{2,3} x_2 x_3 + \beta_{2,4} x_2 x_4 + \beta_{3,4} x_3 x_4.$$

A obtenção do valor mais adequado de λ ocorreu através da análise do gráfico denominado de traço de crista.

**Figura 3.2: Traço de Crista correspondente aos dados da
mistura de lubrificantes**



Fonte: St. John (1984)

Uma análise cuidadosa do traço de crista (Figura 3.2) sugere a escolha de um valor de λ no intervalo que varia de 0,004 a 0,005, pois estes correspondem aos primeiros valores da constante de crista que promovem uma maior estabilidade para as estimativas dos parâmetros.

A Tabela 3.14 apresenta os FIV correspondentes às simulações para alguns valores de λ , realizadas com os dados da mistura de lubrificantes no modelo, utilizando a transformação em pseudocomponentes e a padronização respectivamente.

Tabela 3.14: FIV do processo de regressão em cristas correspondentes aos dados da mistura de lubrificantes

$\lambda = 0$	$\lambda = 0,001$	$\lambda = 0,002$	$\lambda = 0,003$	$\lambda = 0,004$	$\lambda = 0,005$
761,53	163,12	70,17	39,13	25,03	17,45
34,45	31,15	27,89	25,21	22,94	20,98
41,56	36,24	32,06	28,64	25,77	23,33
369,89	146,55	80,26	50,87	35,2	25,86
119,23	29,44	15,38	10,58	8,32	7,04
234,81	54,41	26,46	17,01	12,62	10,18
56,65	21,36	13,28	9,95	8,21	7,17
35,2	29,40	25,41	22,30	19,79	17,71
65,57	29,39	18,60	13,68	10,95	9,24
136,64	56,92	33,88	23,55	17,9	14,43

A primeira coluna da Tabela 3.14 repete os valores da Tabela 3.10, pois quando $\lambda=0$, as estimativas correspondem às obtidas pelo método dos mínimos quadrados. Uma outra observação importante consiste perceber, através dos FIV, que mesmo para valores de λ muito pequenos, como $\lambda=0,001$, há uma melhoria grande no condicionamento, apesar de ainda não se atingir os níveis aceitáveis.

Gunst e Hua (1978) sugeriram utilizar os FIV para selecionar a constante de crista. Logo, através da Tabela 3.14, verifica-se que um valor adequado para a constante de crista (λ) estaria entre 0,004 e 0,005, onde os FIV são satisfatórios tendo como referência o modelo inicial.

Vários autores observam ainda que, nos experimentos do tipo mistura, baixos valores de λ provocam sensível melhora. Este fato sugeriria que a estabilidade das estimativas dos parâmetros seria conseguida com apenas pequenos aumentos no vício do estimador.

Alguns comentários devem ser feitos, sendo o primeiro deles referente ao fato do valor de λ , selecionado através do traço de crista (o que pode ser notado através da análise do gráfico), encontrar-se bem próximo do valor sugerido pelo estudo dos FIV. Portanto, os FIV podem ser bons indicadores de estabilidade das estimativas dos parâmetros do modelo, ou seja, quanto menor a magnitude do FIV maior a estabilidade das estimativas do parâmetro correspondente.

Um segundo fato, também observado na Figura 3.2, refere-se a alta sensibilidade das estimativas do método dos mínimos quadrados. Analisando os valores dessas estimativas que estão no eixo das ordenadas ($\lambda=0$) e pode-se verificar que um pequeno aumento no valor de λ provoca grandes oscilações de magnitude em algumas estimativas dos parâmetros. Pode-se destacar $\hat{\beta}_1^*$ que apresenta magnitude 35 quando λ vale zero e passa a 13 quando trabalha-se com λ entre 0,004 e 0,006. Outros parâmetros que também merecem destaque são β_{12}^* e β_{13}^* , pois pelo método dos mínimos quadrados, apresentaram estimativas negativas e pela regressão em cristas, suas estimativas são positivas. Assim, quando há multicolinearidade, as estimativas pelo método dos mínimos quadrados podem ser muito instáveis e pequenas variações na amostra podem acarretar grandes alterações nas estimativas dos parâmetros. Logo, torna-se muito claro neste tipo de experimento a importância do estimador em cristas para gerar estimativas estáveis.

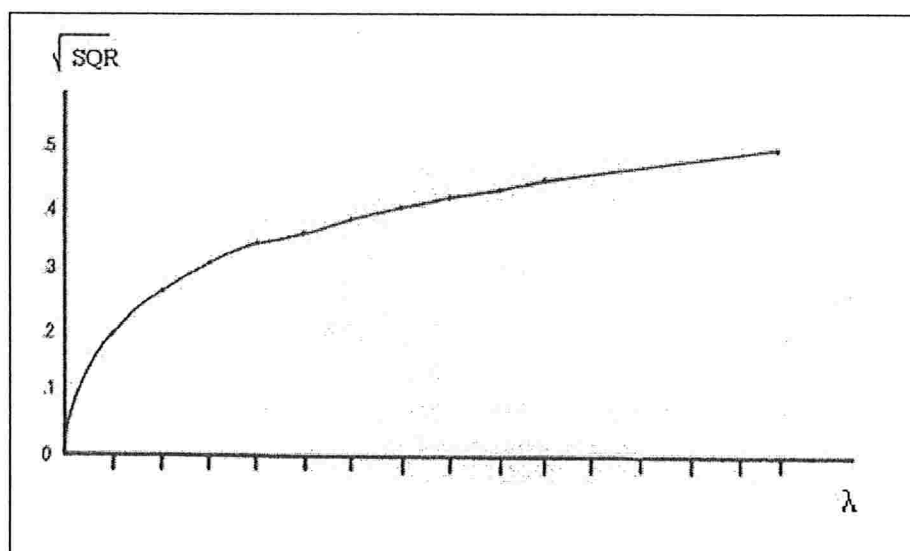
3.4 Considerações Finais

Uma das verificações mais importantes no processo de regressão em cristas, quando aplicado a experimentos do tipo mistura, é a grande melhoria apresentada no condicionamento da matriz de planejamento mesmo quando utilizados valores muito reduzidos da constante de crista.

No exemplo de Snee (1975), o primeiro valor de λ utilizado foi de 0,001 (Tabela 3.14) e o mesmo provocou uma considerável redução nos FIV, em relação àqueles obtidos quando λ vale zero.

É importante destacar também que o aumento da constante λ , que traz como benefício uma maior estabilidade para as estimativas, também provoca um aumento na soma dos quadrados dos resíduos (SQR) e conseqüentemente uma redução no coeficiente de determinação (R^2). Na Figura 3.3, apresenta-se o gráfico da raiz quadrada de SQR em função de λ referente à mistura de lubrificantes do exemplo.

Figura 3.3: Gráfico \sqrt{SQR} em função de λ do modelo da mistura de lubrificantes



Fonte: St. John (1984)

Através da Figura 3.3, tem-se uma verificação prática de que um aumento de λ provoca obrigatoriamente um aumento na SQR. Logo, deve-se contrabalançar o ganho em estabilidade nas estimativas e as perdas na explicação do modelo, ou seja, a redução de R^2 .

Deste modo, diante dos experimentos do tipo mistura, em que problemas de multicolinearidade são muito frequentes, a obtenção dos FIV é de extrema importância para se dimensionar a gravidade do problema e a regressão em cristas contribui principalmente para atingir uma maior estabilidade dos estimadores dos parâmetros do modelo.

O problema da inferência quando se utiliza o estimador em cristas não está solucionado e há divergências entre os autores quanto à sua utilização, estes diferentes pontos de vista podem ser encontrados em Montgomery, Peck e Vining (2006, p.357), Gunst e Mason (1980, p.347) e Ryan (1997, p. 408).

Uma vez definido o modelo, um interesse muito comum por parte dos pesquisadores é de maximizar ou minimizar a variável resposta. No Capítulo 4, será apresentado o processo de Análise de Cristas, cuja finalidade é justamente determinar estes valores de máximo e de mínimo.

Capítulo 4

Análise de Cristas nos experimentos do tipo mistura

4.1 Introdução

Este capítulo tem como propósito discutir a técnica de análise de cristas aplicada aos modelos gerados a partir dos experimentos do tipo mistura. Em um primeiro momento, o método de análise de cristas será minuciosamente discutido, incluindo a sua adaptação e utilidade nos experimentos do tipo mistura. Posteriormente serão apresentados exemplos que ilustram e permitem uma visão mais prática do método proposto.

Antes do início do desenvolvimento do método de análise de cristas, é importante reforçar que este consiste em um processo de otimização, maximização ou minimização da variável resposta, dependendo do interesse do pesquisador. Já a regressão em cristas, que foi discutida no capítulo anterior,

envolve um método de estimação, logo é crucial para o entendimento deste capítulo, que o leitor tenha enfatizadas as diferenças entre esses dois procedimentos.

Os modelos que melhor se adaptam aos experimentos do tipo mistura, segundo Scheffé (1958), são funções polinomiais de baixo grau. De acordo com Draper e Pukelsheim (2000, 2002) e Hoerl (1987), os modelos de segunda ordem são os mais adequados em experimentos do tipo mistura. Tais modelos serão utilizados no presente capítulo com o propósito de investigar o comportamento da variável resposta, através do estudo de superfícies de resposta.

A técnica de superfície de resposta constrói uma relação entre uma variável resposta quantitativa com variáveis quantitativas que podem influenciá-la. Alguns dos objetivos desse ajuste consistem em analisar como a resposta é afetada pelas variáveis explicativas em uma particular região de interesse ou ainda localizar o ponto, formado pelo conjunto de valores das variáveis explicativas, que produziria a resposta máxima (ou mínima).

Segundo Oishi (1983), a análise de cristas foi desenvolvida por Hoerl em 1959, na tentativa de desenvolver uma técnica alternativa de otimização a ser aplicada em problemas industriais.

Este procedimento de otimização tem como objetivo maximizar ou minimizar uma função, quando certas restrições são impostas. A proposta é de realizar uma varredura na região definida para o problema e fazer um levantamento dos pontos de máximo ou mínimo locais que serão encontrados. Estes pontos serão denominados de cristas (*ridges*) e através deles constrói-se uma sequência de pontos que conduzirá ao máximo ou ao mínimo absoluto, sendo este definido como o caminho de cristas.

Nos experimentos do tipo mistura, a proposta será maximizar ou minimizar a variável resposta dentro da região de mistura, que será definida pelos limites inferior e superior das proporções dos componentes. Para uma melhor compreensão do método de análise de cristas, a idéia inicial será

desenvolvê-lo, sem considerar as restrições de mistura, que serão introduzidas gradativamente ao longo deste capítulo.

O método consiste na realização de uma varredura, a partir da origem, do espaço k -dimensional $x = (x_1, x_2, \dots, x_k) = (0, 0, \dots, 0)$, utilizando a restrição

$$x'x = x_1^2 + x_2^2 + \dots + x_k^2 = R^2, \quad (4.1)$$

em que o valor de R é aumentado gradativamente e os pontos de máximo ou mínimo são determinados para cada valor de R selecionado. Logo, se o problema for de maximização, todos estes pontos de máximo local definirão um caminho para a obtenção do máximo absoluto. No entanto, se o problema for de minimização, a determinação do mínimo absoluto ocorrerá através do caminho definido por todos os mínimos locais determinados de acordo com a variação de R .

Neste processo também podem ser obtidos caminhos que conduzem a pontos estacionários intermediários, que não são pontos nem de máximo absoluto nem de mínimo absoluto. Estes pontos podem ter interesse prático em determinados problemas, porém nesta dissertação eles serão irrelevantes, já que a proposta é que esta técnica atue como uma ferramenta de otimização (maximização ou minimização) da resposta dos experimentos do tipo mistura.

Para ilustrar o processo graficamente, os dois exemplos mais simples envolvem experimentos do tipo mistura, em que no primeiro caso tem-se uma mistura binária (dois componentes) e no segundo caso uma mistura ternária (três componentes).

Quando a mistura possui exatamente dois componentes (x_1 e x_2), a região a ser estudada é um segmento de reta de equação

$$x_1 + x_2 = 1, \text{ em que}$$

$$0 \leq x_1 \leq 1 \text{ e } 0 \leq x_2 \leq 1.$$

Na Figura 4.1, o sistema de coordenadas cartesiano bidimensional possui os eixos definidos por x_1 e x_2 e o segmento de reta unindo os pontos $(0,1)$ e $(1,0)$ representa os pontos que satisfazem à região de mistura.

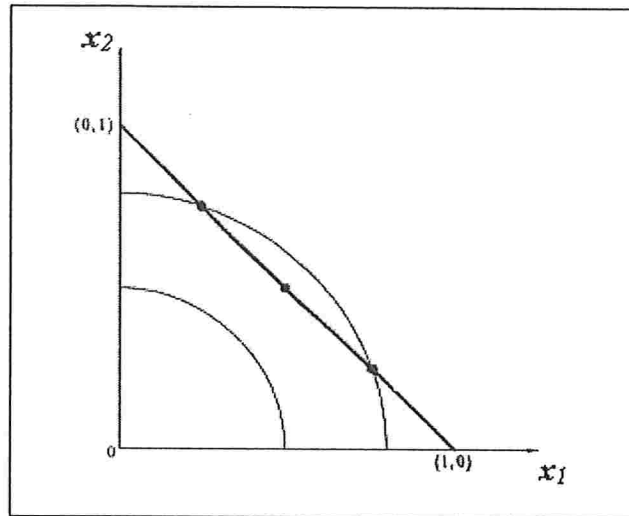


Figura 4.1: Representação gráfica do procedimento de Análise de Cristas para uma mistura com dois componentes

Quando a mistura é formada por dois componentes, a restrição da equação (4.1) será representada graficamente por circunferências. Na Figura 4.1, são ilustradas duas das infinitas circunferências de equação

$$x'x = x_1^2 + x_2^2 = R^2 \quad (4.2)$$

que possuem centro na origem e raio R , em que R é uma constante que inicialmente possui valor nulo e vai sendo aumentada continuamente até que todo o espaço de mistura tenha sido varrido.

Nota-se que se o valor de R é menor que $\frac{1}{\sqrt{2}}$ não há solução, pois não há intersecção entre as circunferências e o espaço de mistura. Quando o valor de R é igual a $\frac{1}{\sqrt{2}}$, há um único ponto de interesse, que é o ponto de tangência da reta com a circunferência, e quando R apresentar um valor maior que $\frac{1}{\sqrt{2}}$ e menor ou igual a 1, haverá dois pontos de intersecção, ou seja, duas soluções para cada valor de R neste intervalo.

Quando a mistura apresenta três componentes (x_1 , x_2 e x_3), a restrição de mistura define uma região que corresponde a um triângulo equilátero de equação

$x_1 + x_2 + x_3 = 1$, em que

$0 \leq x_1 \leq 1$, $0 \leq x_2 \leq 1$ e $0 \leq x_3 \leq 1$,

conforme ilustra a Figura 4.2.

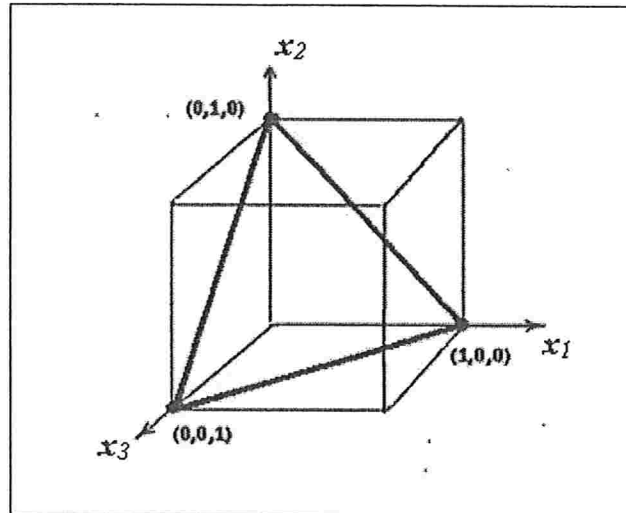


Figura 4.2: Representação gráfica da região de mistura para um experimento com três componentes

Em presença de três componentes, a equação (4.1) representará diferentes superfícies esféricas a partir do contínuo aumento de R , tendo como ponto de partida a origem:

$$x'x = x_1^2 + x_2^2 + x_3^2 = R^2.$$

Neste caso, percebe-se que se o valor de R é menor que $\frac{1}{\sqrt{3}}$ não há solução, pois não há intersecção entre as superfícies esféricas e o espaço de mistura. Deste modo, quando o valor de R é igual a $\frac{1}{\sqrt{3}}$, há um único ponto de interesse, que é o ponto de tangência entre a superfície esférica e o triângulo equilátero. Logo, quando R apresentar um valor maior que $\frac{1}{\sqrt{3}}$ e menor que 1, as intersecções corresponderão a circunferências ou arcos de circunferências, todos com o centro no baricentro do triângulo que representa a região de mistura

e finalmente quando R é igual a 1 tem-se exatamente três pontos de intersecção que correspondem aos vértices do triângulo equilátero.

Estas representações geométricas são de extrema importância para uma melhor visualização e entendimento do procedimento. No entanto, o mais importante é destacar a validade do processo algébrico em experimentos do tipo mistura com qualquer quantidade de componentes.

Conforme explicado anteriormente, o modelo de segunda ordem é o que melhor se adapta aos experimentos do tipo mistura e o método de análise de cristas apresentado no presente capítulo será desenvolvido exclusivamente para este modelo.

Segundo Draper e Pukelsheim (2002), o modelo de segunda ordem padrão que ainda não envolve as restrições dos experimentos do tipo mistura é dado por:

$$\hat{y} = b_0 + x'b + x'Bx,$$

em que

$$x' = (x_1, x_2, \dots, x_k),$$

$$b' = (b_1, b_2, \dots, b_k) \text{ e}$$

$$B = \begin{pmatrix} b_{11} & \frac{1}{2}b_{12} & \dots & \frac{1}{2}b_{1k} \\ \frac{1}{2}b_{21} & b_{22} & \dots & \frac{1}{2}b_{2k} \\ \vdots & & \ddots & \\ \frac{1}{2}b_{k1} & \frac{1}{2}b_{k2} & \dots & b_{kk} \end{pmatrix}.$$

Através desta equação matricial, obtém-se a equação de regressão,

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k + b_{11}x_1^2 + b_{22}x_2^2 + \dots + b_{kk}x_k^2 + b_{12}x_1x_2 + b_{13}x_1x_3 + \dots + b_{k-1,k}x_{k-1}x_k, \quad (4.3)$$

em que $b_1, b_2 \dots b_k$ e $b_{11}, \dots b_{k-1,k}$ são as estimativas dos parâmetros do modelo de regressão de segunda ordem.

O procedimento de análise de cristas é baseado nos pontos estacionários do modelo, que podem ser de máximo, mínimo ou pontos de sela. O processo de maximização, por exemplo, consiste em um levantamento de todos os máximos locais e através deles determinar o máximo absoluto. A mesma idéia tem-se para o processo de minimização e, de acordo com Draper e Pukelsheim (2002), estes pontos estacionários podem ser obtidos pela função de Lagrange.

A equação da função de Lagrange, com a restrição (4.1) incorporada, é dada por:

$$F = b_0 + x'b + x'Bx - \alpha(x'x - R^2), \quad (4.4)$$

em que α é o multiplicador de Lagrange.

Diferenciando a equação (4.4) em relação a $x = (x_1, x_2, \dots, x_k)'$, tem-se a seguinte equação na forma matricial

$$\frac{\partial F}{\partial x} = b + 2Bx - 2\alpha x,$$

que quando igualada a zero gera o seguinte vetor solução,

$$x = -\frac{1}{2}(B - \alpha I_k)^{-1}b. \quad (4.5)$$

Portanto, os vetores x obtidos da equação (4.5) serão pontos estacionários do modelo, ou seja, para um determinado valor de α , o conjunto de variáveis preditoras resultante corresponderá a um ponto de máximo, mínimo ou sela. Deste modo, para se determinar um ponto estacionário, pode-se selecionar um valor qualquer para α , desde que $(B - \alpha I_k)^{-1}$ exista. Assim, os valores de α que corresponderem aos autovalores da matriz B, devem ser excluídos.

Segundo Draper (1963), os pontos estacionários são obtidos de acordo com os seguintes passos:

- a) Adota-se um valor possível para α .
- b) Determina-se o valor de x_1, x_2, \dots, x_k a partir das equações (4.5).
- c) Calcula-se o valor de R, através da equação (4.1).
- d) Calcula-se \hat{y} , através da equação (4.3).

Após determinado o conjunto de valores $(\alpha, x_1, x_2, \dots, x_k, R, \hat{y})$, é importante entender que eles correspondem a um ponto estacionário de valor \hat{y} obtido através de (x_1, x_2, \dots, x_k) , que se encontra sobre a “superfície esférica” de raio R com centro na origem. Este ponto estacionário pode ser classificado em máximo local, mínimo local, máximo absoluto, mínimo absoluto ou um ponto de sela.

A necessidade do cálculo de R ainda não foi justificada, porém pode-se adiantar que este parâmetro será importante na classificação dos pontos estacionários.

Algumas propriedades dos pontos estacionários demonstradas em Draper (1963) são de fundamental importância para a classificação quanto à natureza (máximo, mínimo ou sela). Estas propriedades serão apresentadas, pois estes resultados fundamentam a classificação destes pontos em função de α .

Resultado 4.1

Se R_1 e R_2 são valores de R e α_1 e α_2 valores de α , com

$$R_1 = R_2 \text{ e } \alpha_1 > \alpha_2, \text{ então } \hat{y}_1 > \hat{y}_2.$$

Resultado 4.2

Se α é menor que todos os autovalores da matriz B , então o ponto estacionário obtido é um ponto de mínimo local.

Se α é maior que todos os autovalores da matriz B , então o ponto estacionário obtido é um ponto de máximo local.

Logo, de acordo com Draper (1963), através dos resultados 4.1 e 4.2, verifica-se que o máximo absoluto é obtido quando adotado um específico valor para α , que deve respeitar inicialmente a condição de ser maior que todos os

autovalores da matriz B. Deste modo, o autor sugere estudar os valores de α , a partir de $+\infty$ até o maior autovalor da matriz B, pois tem-se a garantia que o máximo absoluto será obtido a partir de um valor de α deste intervalo. Analogamente, conclui-se que para minimizar a função devem ser estudados os os valores de α , a partir de $-\infty$ até o menor autovalor da matriz B. Uma discussão final refere-se aos critérios a serem adotados para selecionar este valor de α específico dentro do intervalo proposto. Para isto, dois outros resultados, extraídos de Draper (1963), serão apresentados a seguir:

Resultado 4.3

Conforme o aumento de R, a variação de \hat{y} dos pontos estacionários apresenta uma das características a seguir:

- a) Monotonicamente decrescente.
- b) Monotonicamente crescente.
- c) Passa pelo máximo e então decresce monotonicamente.
- d) Passa pelo mínimo e então cresce monotonicamente.

Resultado 4.4

Se $\alpha \rightarrow \pm\infty$, então $R \rightarrow 0$.

O autor acrescenta ainda que quando α tende a um dos autovalores da matriz B, o valor de R tende a $+\infty$. Ainda dentro deste contexto, o parâmetro R varia de zero para $+\infty$ quando a variação de α ocorre de $-\infty$ para o menor autovalor da matriz B e também quando a variação de α ocorre de $+\infty$ para o maior autovalor da matriz B.

Logo, através dos resultados 4.3 e 4.4, pode-se afirmar que na obtenção do máximo absoluto, conforme α varia de $+\infty$ para o maior autovalor da matriz B, o parâmetro R varia de zero a $+\infty$. Segundo o resultado 4.3, se o valor de R aumenta, a evolução de \hat{y} pode ocorrer apenas das quatro formas distintas destacadas anteriormente. Deste modo, se a variação de \hat{y} em função de R for monotonicamente crescente, o máximo absoluto é obtido quando se atinge um limite experimental, ou seja, quando a proporção máxima ou mínima de um dos componentes da mistura, estabelecida nas condições iniciais do problema, é

ultrapassada. Se \hat{y} passar pelo máximo antes de ser extrapolado um limite experimental, este fato será diagnosticado, pois após a obtenção do máximo absoluto, a variação de \hat{y} é monotonicamente decrescente. Em caso, de comportamentos decrescentes de \hat{y} , as variáveis preditoras correspondentes ao máximo absoluto ou serão as coordenadas do ponto de partida do procedimento ou serão coordenadas de um limite experimental.

Portanto, de acordo com os resultados acima, serão definidos de forma prática os processos de obtenção dos pontos de máximo e mínimo absolutos.

Se o objetivo é encontrar o máximo absoluto de \hat{y} , os valores de α devem variar de $+\infty$ para o maior autovalor da matriz B e para cada valor de α deve-se obter os valores de x_1, x_2, \dots, x_k resolvendo o sistema de equações gerado por (4.5), o valor de R, através da equação (4.2) e \hat{y} , através da equação (4.3). Deste modo o máximo absoluto é o maior \hat{y} obtido, desde que sejam respeitadas todas as restrições experimentais.

Se o objetivo é encontrar o mínimo absoluto, a única mudança na sequência de passos é que os valores de α devem variar de $-\infty$ para o menor autovalor da matriz B.

4.2 Análise de cristas em experimentos do tipo mistura

Nos experimentos do tipo mistura são adicionadas restrições lineares à função de Lagrange, no caso mais simples a única restrição é a de mistura

$$x_1 + x_2 + \dots + x_k = 1, \quad (4.6)$$

no entanto, um número maior de restrições lineares podem ser associadas ao modelo, conforme veremos ainda nesta seção.

Draper e Pukelsheim (2002) sugerem uma mudança no ponto de partida da análise de cristas nos experimentos do tipo mistura, que será denotado por f. Segundo os autores, este ponto inicial pode ser um ponto qualquer, no entanto, em experimentos que envolvem a restrição (4.6), sugere-se o ponto central da

região de mistura (centróide) como ponto de partida da análise. Logo, quando $f \neq 0$, uma transformação deve ser feita na equação (4.2):

$$(x - f)'(x - f) = R^2, \quad (4.7)$$

em que f é dado por:

$$f' = \left(\frac{1}{k}, \frac{1}{k}, \dots, \frac{1}{k} \right).$$

Esta transformação, em termos práticos, traz apenas uma redefinição do valor do raio no modelo original de análise de cristas e uma mudança no centro das “esferas” que devem interceptar o espaço de mistura, ou seja, este novo centro passa a ser f , o centróide do espaço de mistura. Pode-se verificar, através da equação (4.7) esta alteração no centro das “esferas” e a redefinição do raio a partir da demonstração a seguir:

$$\begin{aligned} (x - f)'(x - f) &= R^2, \\ x'x - 2f'x + f'f &= R^2, \\ x'x - 2/k + 1/k &= R^2, \\ x'x &= R^2 + 1/k. \end{aligned} \quad (4.8)$$

Deste modo, a equação (4.8) corresponde exatamente à restrição (4.2) com uma alteração na magnitude do raio, portanto a introdução de f no modelo traz apenas uma mudança de referenciais.

Além da introdução de f como ponto de partida, será discutida a adição das restrições lineares dos experimentos do tipo mistura na função de Lagrange, ressaltando que nos casos mais simples a única restrição é a de mistura, dada pela equação (4.6).

O conjunto de restrições lineares podem ser reescritas da seguinte forma:

$$Mx = N, \quad (4.9)$$

em que M é uma matriz $m \times k$ e as m linhas são linearmente independentes e normalizadas de modo que a soma dos quadrados dos elementos de cada linha é igual a 1. O vetor N é um vetor coluna de m linhas.

Por exemplo, quando $m=1$, a única restrição será a de mistura e as matrizes M e N podem ser escritas da seguinte forma:

$$M = (1/k^{1/2}, 1/k^{1/2}, \dots, 1/k^{1/2}) \quad e \quad (4.10)$$

$$N = (1/k^{1/2}). \quad (4.11)$$

Se houver uma restrição a mais, neste caso, $m=2$ e as matrizes M e N serão dadas por:

$$M = \begin{pmatrix} 1/k^{1/2} & 1/k^{1/2} & \dots & 1/k^{1/2} \\ \gamma_1 & \gamma_2 & \dots & \gamma_3 \end{pmatrix} \quad e$$

$$N = \begin{pmatrix} 1/k^{1/2} \\ \gamma \end{pmatrix},$$

em que

$$\gamma_1 + \gamma_2 + \dots + \gamma_3 = \gamma.$$

O número de linhas da matriz M é normalmente menor que k e quando este número de linhas é igual a k , tem-se um único ponto do espaço de mistura que satisfaz ao sistema $Mx = N$.

Em uma função polinomial na forma canônica, sob as condições (4.7) e (4.9), a função de Lagrange é dada por:

$$G = x'b + x'Bx - \alpha[(x-f)'(x-f) - R^2] - \theta'(Mx - N), \quad (4.12)$$

em que α e $(\theta_1, \theta_2, \dots, \theta_m)$ que formam θ' são os multiplicadores de Lagrange. Logo, derivando a equação (4.12) em relação a x , obtém-se:

$$\frac{\partial G}{\partial x} = b + 2Bx - 2\alpha(x-f) - M'\theta. \quad (4.13)$$

Igualando a equação (4.13) ao vetor nulo, tem-se a seguinte solução para o vetor x :

$$x = \frac{1}{2}(B - \alpha I)^{-1}(M'\theta - b - 2\alpha f). \quad (4.14)$$

Devido a equação (4.9),

$$N = \frac{1}{2}M(B - \alpha I)^{-1}M'\theta - \frac{1}{2}M(B - \alpha I)^{-1}(b + 2\alpha f)$$

e como consequência,

$$\theta = \{M(B - \alpha I)^{-1}M'\}^{-1}\{2N + M(B - \alpha I)^{-1}(b + 2\alpha f)\}. \quad (4.15)$$

Logo, a sequência de passos para a obtenção do máximo ou do mínimo absoluto é estruturada por:

- a) Escolha dos valores apropriados para α , de acordo com o caminho desejado (máximo ou mínimo).
- b) Para cada α adotado, obter o vetor θ , solucionando a equação (4.15)
- c) Obter o vetor x , solucionando a equação (4.14) para α e θ correspondentes.
- d) Calcula-se R e \hat{y} , através das equações (4.7) e (4.3), respectivamente.

Segundo Draper e Pukelsheim (2002), as análises de cristas que não apresentam restrições lineares, em que a função de Lagrange é representada pela equação (4.4),

$$F = b_0 + x'b + x'Bx - \alpha(x'x - R^2),$$

a escolha dos valores de α apropriados para os caminhos de máximo e mínimo absoluto é orientada pela matriz de segundas derivadas:

$$\begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_1 \partial x_k} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 F}{\partial x_k \partial x_1} & \frac{\partial^2 F}{\partial x_k \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_k^2} \end{pmatrix} = 2(B - \alpha I_k). \quad (4.16)$$

Logo, são os autovalores da matriz B que atuam como referências para discriminação entre os caminhos, conforme explicado anteriormente.

Nos problemas em que são acrescidas restrições lineares, verifica-se que os autovalores da matriz B tornam-se não apropriados para os procedimentos de maximização e minimização, logo há a necessidade da geração de uma matriz de

menor dimensão que permita a consideração destas restrições. A seguir, será discutido o processo de transformação que viabiliza a introdução das restrições lineares.

Quando o conjunto de restrições lineares é incorporado ao modelo, os autores demonstram que para uma matriz M com m linhas normalizadas e k colunas, ou seja, contendo m restrições lineares, pode-se construir uma outra matriz T com $k-m$ linhas e k colunas, em que qualquer linha de T é ortogonal em relação a todas as linhas de M e que $TT' = I_{k-m}$.

Portanto, as colunas de M' formam uma base para o espaço de restrições e as colunas de T' formam uma base ortonormal para o espaço ortogonal a M' .

Assim, é importante destacar que a matriz T a ser construída não é única e conforme as considerações acima, deve respeitar ao conjunto de equações:

$$TM' = 0 \text{ de dimensão } (k-m) \times m \quad (4.17)$$

$$MT' = 0 \text{ de dimensão } m \times (k-m) \quad (4.18)$$

$$TT' = I_{k-m} \quad (4.19)$$

Após construída a matriz T , para iniciar a transformação deve ser gerada a seguinte matriz $H = \begin{pmatrix} M \\ T \end{pmatrix}$.

A matriz quadrada H de ordem k promove uma transformação do vetor $x = (x_1, x_2, \dots, x_k)$ em um vetor $z = (z_1, z_2, \dots, z_k)$, através de $z = Hx$.

Por meio da partição do vetor $z = (z_1, z_2, \dots, z_m, z_{m+1}, \dots, z_k)$ em $z' = (e', f')$, com $e' = (z_1, z_2, \dots, z_m)$ e $f' = (z_{m+1}, \dots, z_k)$, tem-se

$$z = \begin{pmatrix} e \\ f \end{pmatrix} = \begin{pmatrix} M \\ T \end{pmatrix} x = \begin{pmatrix} Mx \\ Tx \end{pmatrix} = \begin{pmatrix} N \\ Tx \end{pmatrix}.$$

De acordo com Draper e Pukelsheim (2002), a matriz inversa de H é dada por:

$$H^{-1} = [M'(MM')^{-1} \quad T'].$$

A matriz MM' é não singular e a verificação deste fato pode ser feita através do cálculo:

$$\begin{aligned} HH^{-1} &= \begin{bmatrix} M \\ T \end{bmatrix} [M'(MM')^{-1} \quad T'] = \begin{bmatrix} I_m & 0 \\ TM'(MM')^{-1} & TT' \end{bmatrix} = \\ &= \begin{bmatrix} I_m & 0 \\ 0 & I_{k-m} \end{bmatrix} = I_k. \end{aligned}$$

Analogamente, $H^{-1}H = I_k$. Assim, utilizando $x = H^{-1}z$ no primeiro termo quadrático da função de Lagrange representada na equação (4.12), tem-se

$$\begin{aligned} x'Bx &= z'(H^{-1})BH^{-1}z = \\ &= [N' \quad f'] \begin{bmatrix} (MM')^{-1}M \\ T \end{bmatrix} B [M'(MM')^{-1} \quad T'] \begin{bmatrix} N \\ f \end{bmatrix} = \\ &= [N'(MM')^{-1}M + f'T] B [M'(MM')^{-1}N + T'f] = \\ &= f'TBT'f + 2f'TBM'(MM')^{-1}N + N'(MM')^{-1}MBM'(MM')^{-1}N. \end{aligned}$$

Na equação do primeiro termo quadrático da função de Lagrange, quando adota-se $B = I_k$, obtém-se uma expressão para $x'x$ muito útil para o desenvolvimento do segundo termo quadrático da função de Lagrange (4.12), dado por:

$$\alpha x'x = \alpha f'f + \alpha N'(MM')^{-1}N.$$

Deste modo, diferenciando duas vezes a função de Lagrange da equação (4.12) em relação a f , tem-se:

$$\begin{pmatrix} \frac{\partial^2 G}{\partial v_1^2} & \frac{\partial^2 G}{\partial v_1 \partial v_2} & \cdots & \frac{\partial^2 G}{\partial v_1 \partial v_{k-m}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 G}{\partial v_{k-m} \partial v_1} & \frac{\partial^2 G}{\partial v_{k-m} \partial v_2} & \cdots & \frac{\partial^2 G}{\partial v_{k-m}^2} \end{pmatrix} = 2(TBT' - \alpha I_{k-m}). \quad (4.20)$$

No procedimento de análise de cristas, a equação (4.20) tem a função de substituir a equação (4.16) quando existem restrições lineares e os autovalores da matriz TBT' tornam-se as referências do processo de maximização ou minimização.

Deste modo, Draper e Pukelsheim (2002) demonstraram que com a inclusão das restrições lineares, pode-se, uma vez determinadas as matrizes B e T , calcular a matriz quadrada TBT' de ordem $k-m$. Os $k-m$ autovalores de TBT' serão as referências dos processos de maximização ou minimização.

Logo, em um experimento do tipo mistura, com uma ou mais restrições lineares, tem-se as seguintes condições:

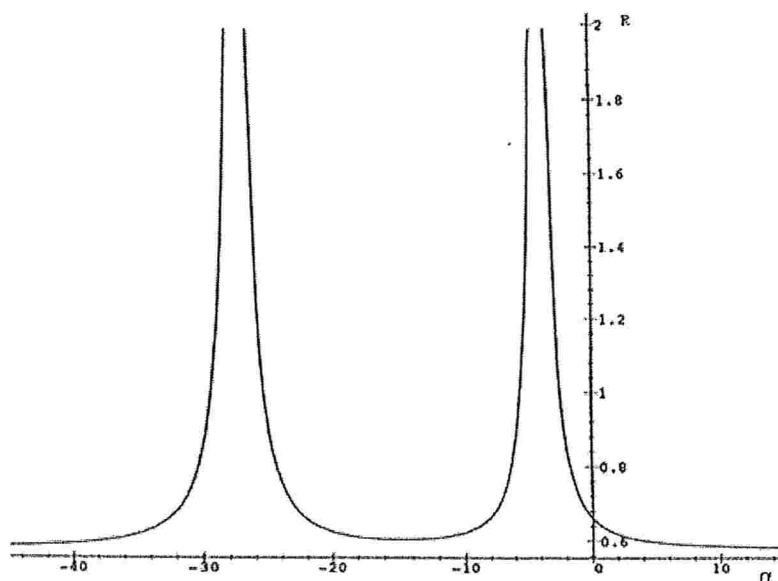
- a) Escolhendo-se valores de α maiores que o maior autovalor de TBT' , são obtidos pontos de máximo local.
- b) Escolhendo-se valores de α menores que o menor autovalor de TBT' , são obtidos pontos de mínimo local.

Em termos práticos, a obtenção do máximo absoluto consiste em uma varredura dos valores de α , a partir de $+\infty$ ao maior autovalor de TBT' , respeitando todas as restrições experimentais. O processo de minimização segue a mesma lógica, com a diferença da varredura partir de $-\infty$ até o menor autovalor de TBT' .

Destaca-se que o espaçamento entre os valores de α estudados dependem da precisão requisitada e da proximidade em relação ao ponto de máximo ou de mínimo absoluto, ou seja, quanto mais próximo do ponto de interesse, mais cuidadosa deve ser a análise e menor o espaçamento entre os valores do parâmetro α .

Para ilustrar a variação de R , de acordo com a evolução de α , encontra-se na Figura 4.3, um gráfico característico de R em função de α , extraído de um exemplo prático de Draper e Pukelsheim (2000), em que os autovalores da matriz TBT' são $-27,28$ e $-3,86$.

Figura 4.3: Gráfico característico de R em função de α .



Fonte: Draper e Pukelsheim (2000)

Algumas considerações finais sobre a matriz T e seu respectivo processo de construção devem ser feitas. Por exemplo, em um experimento do tipo mistura, em que $k=3$ e há apenas a restrição de mistura, ou seja, $m=1$, a matriz M e a matriz N poderão ser:

$$M = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix} \quad e \quad N = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Deste modo, a matriz T será da forma

$$T = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}_{2 \times 3}.$$

Portanto, resolvendo o sistema com as equações (4.17), (4.18) e (4.19), o processo de resolução agrega a informação que a somatória dos elementos de cada linha da matriz T deve ser nula, devido às equações (4.17) e (4.18). Já o

conjunto solução corresponde às matrizes T , que satisfazem ao sistema, deve-se evidenciar o fato da matriz T não ser única.

Assim, adotando $b=0$, tem-se uma possibilidade para a matriz T , representada a seguir:

$$T = \begin{pmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix}.$$

A construção da matriz T pode ser de forma empírica dentro das condições apresentadas. Segundo Draper e Pukelsheim (2000), para este propósito, podem também ser utilizados os polinômios ortogonais, pois satisfazem ao conjunto de equações (4.17), (4.18) e (4.19).

4.3 Exemplo

Para ilustrar a aplicação da análise de cristas, utilizaremos o exemplo de Draper e Pukelsheim (2002), um estudo realizado a partir do conjunto de dados de um experimento do tipo mistura da área farmacêutica extraído de Anik e Sukumar (1981). O experimento consistiu em variar as concentrações de cinco ingredientes polietileno glicol 400 (x_1), glicerina (x_2), polisorbato 60 (x_3), água (x_4) e poloxamer 407 (x_5) com o objetivo de avaliar a solubilidade da mistura que seria a variável resposta.

Neste estudo, uma das considerações dos autores é que um dos ingredientes, poloxamer 407 (x_5), tem a sua concentração mantida constante ao longo do experimento, deste modo a restrição linear de mistura representada pela equação (4.6), neste caso particular, passa a ser:

$$x_1 + x_2 + x_3 + x_4 = 0,9. \quad (4.21)$$

Portanto, o modelo polinomial quadrático na forma canônica extraído de Draper e Pukelsheim (2002), possui 4 variáveis preditoras ($k=4$) e a restrição de mistura particularizada na equação (4.21).

O processo de regressão através do método dos mínimos quadrados, apresenta o seguinte modelo:

$$\hat{y} = 49,716x_1 + 8,414x_2 + 29,95x_3 + 4,3365x_4 - 58,671x_1x_2 - 27,83x_1x_3 - 74,902x_1x_4 + 10,20x_2x_3 + 33,81x_3x_4 . \quad (4.22)$$

Neste exemplo, não será discutido o problema da multicolinearidade dos dados experimentais nem as estratégias que foram necessárias para mitigá-lo. O foco desta discussão é a maximização da variável resposta do modelo representado pela equação (4.22), através da análise de cristas.

Tabela 4.1: Limites inferiores e superiores da região experimental do modelo de Anik e Sukumar [1981].

Variável	x_1	x_2	x_3	x_4
Limite inferior	0,10	0,10	0,00	0,30
Limite superior	0,40	0,40	0,08	0,70

A Tabela 4.1 apresenta os limites inferior e superior para cada variável de acordo com a região experimental definida. Logo, neste modelo, tem-se a seguinte notação matricial:

$$b = \begin{pmatrix} 49,716 \\ 8,414 \\ 29,95 \\ 4,3365 \end{pmatrix},$$

$$B = \begin{pmatrix} 0 & -29,3355 & -13,915 & -37,451 \\ -29,3355 & 0 & 5,1 & 0 \\ -13,915 & 5,1 & 0 & 16,905 \\ -37,451 & 0 & 16,905 & 0 \end{pmatrix},$$

$$M=(0,5 \ 0,5 \ 0,5 \ 0,5) \text{ e}$$

$$N = (0,5).$$

Apesar da sugestão da escolha do centróide como ponto de partida, neste caso particular o ponto escolhido pelos autores foi $f' = (0,21, 0,21, 0,04, 0,44)$.

Uma forma apropriada para a matriz T , de acordo com as condições (4.17), (4.18) e (4.19), é:

$$T = \begin{pmatrix} -0,6708 & -0,2236 & 0,2236 & 0,6708 \\ 0,5 & -0,5 & -0,5 & 0,5 \\ -0,2236 & 0,6708 & -0,6708 & 0,2236 \end{pmatrix}.$$

Deste modo a matriz TBT' é calculada e os seus três autovalores são 20,04 , 2,52 e 46,87. Assim, o caminho para o máximo absoluto por meio da análise de cristas é definido pelo estudo dos valores de α , a partir de $+\infty$, em que $x = f$ e $R = 0$ até o maior autovalor, no caso, 46,87. Conforme os valores de α se aproximam deste maior autovalor, R tende ao infinito.

A Tabela 4.2 apresenta a evolução dos valores de α , assim como os valores das quatro variáveis preditoras, R e \hat{y} , que são calculados de acordo com as diretrizes apresentadas na Seção 4.2.

Tabela 4.2: Caminho de cristas para obtenção do máximo absoluto do modelo proposto sob a restrição $x_1 + x_2 + x_3 + x_4 = 0,9$

α	x_1	x_2	x_3	x_4	R	\hat{y}
$+\infty$	0,210	0,210	0,040	0,440	0	6,27
2000	0,209	0,207	0,048	0,436	0,010	6,64
1000	0,208	0,204	0,056	0,432	0,020	7,02
750	0,207	0,202	0,062	0,429	0,026	7,27
500	0,206	0,199	0,072	0,423	0,038	7,75
400	0,205	0,196	0,080	0,419	0,048	8,10
300	0,204	0,191	0,092	0,413	0,062	8,66
250	0,203	0,187	0,102	0,408	0,074	9,10

Pode-se perceber através da evolução das quatro variáveis preditoras que a variável x_3 move-se rapidamente para o limite superior ($x_3 = 0,08$), atingindo este contorno quando $\alpha=400$. As outras variáveis (x_1, x_2 e x_4) sofrem alterações mais lentamente, o que ressalta a importância da variável x_3 no processo, logo as futuras explorações devem inserir

$$x_3 = 0,08, \quad (4.23)$$

como mais uma restrição. Conseqüentemente, esta nova condição promoverá um estudo restrito da região de mistura, selecionando apenas os pontos em que a equação (4.23) é satisfeita.

Deste modo, o processo deve ser reiniciado, inserindo na matriz M mais uma linha correspondente à equação (4.23). Logo, as matrizes M e N são dadas por:

$$M = \begin{pmatrix} 0,5 & 0,5 & 0,5 & 0,5 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ e } N = \begin{pmatrix} 0,45 \\ 0,08 \end{pmatrix}.$$

A inclusão de novas restrições inseridas na matriz M redefinem a matriz T , que deve ser recalculada, e neste exemplo uma das possibilidades é:

$$T = \begin{pmatrix} 0,2672 & 0,5345 & 0 & -0,8017 \\ 0,7715 & -0,6172 & 0 & -0,1543 \end{pmatrix}.$$

A seguir, os cálculos dos autovalores da matriz TBT' são realizados. Os resultados obtidos são -0,49 e 45,01 e neste estágio basta definir um foco (f) e construir a tabela de análise de cristas, variando α de $+\infty$ ao maior autovalor, repetindo o procedimento descrito anteriormente.

Os autores selecionaram como foco $f' = (0,203, 0,203, 0,08, 0,30)$ e a Tabela 4.3 foi construída para o modelo, que neste caso apresenta duas restrições lineares.

Tabela 4.3: Caminho de cristas para obtenção do máximo absoluto do modelo proposto sob as restrições $x_1 + x_2 + x_3 + x_4 = 0,9$ e $x_3 = 0,08$

α	x_1	x_2	x_3	x_4	R	\hat{y}
$+\infty$	0,203	0,203	0,08	0,414	0	8,12
1000	0,207	0,203	0,08	0,410	0,005	8,16
500	0,211	0,202	0,08	0,407	0,010	8,21
200	0,225	0,200	0,08	0,395	0,029	8,41
150	0,236	0,197	0,08	0,387	0,042	8,57
125	0,246	0,194	0,08	0,380	0,055	8,74
100	0,265	0,189	0,08	0,366	0,079	9,10
90	0,279	0,184	0,08	0,357	0,097	9,39
80	0,301	0,177	0,08	0,342	0,124	9,90
75	0,317	0,171	0,08	0,332	0,144	10,32
70	0,341	0,162	0,08	0,317	0,173	10,97
66	0,367	0,152	0,08	0,301	0,205	11,80
66,95	0,368	0,152	0,08	0,300	0,206	11,82
60	0,433	0,127	0,08	0,260	0,287	14,31
55	0,549	0,081	0,08	0,190	0,429	20,13

Na Tabela 4.3, são apresentadas as evoluções das quatro variáveis preditoras, R e \hat{y} , de acordo com os valores de α sugeridos. Nota-se que quando $\alpha = 66,95$, o limite inferior da variável x_4 é atingido, assim conforme as diretrizes para maximização do processo de análise de cristas, sempre que um limite é atingido, o estudo passa a ser realizado ao longo deste limite, ou seja, o processo é reiniciado com a inclusão da restrição $x_4 = 0,3$.

Portanto, as novas matrizes M e N serão dadas por:

$$M = \begin{pmatrix} 0,5 & 0,5 & 0,5 & 0,5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad e \quad N = \begin{pmatrix} 0,45 \\ 0,08 \\ 0,30 \end{pmatrix}.$$

Redefinida a matriz M , a matriz T passa, neste caso, a ser um vetor, em que uma das possibilidades é:

$$T = (0,7071, -0,7071, 0, 0)$$

Neste estágio, o único autovalor da matriz TBT' é 29,335 e o máximo absoluto será encontrado variando α a partir de $+\infty$ até este autovalor. Sendo $f' = (0,26, 0,26, 0,08, 0,30)$, o foco sugerido, o próximo passo é a construção da tabela.

Tabela 4.4: Caminho de cristas para obtenção do máximo absoluto do modelo proposto sob as restrições $x_1 + x_2 + x_3 + x_4 = 0,9$, $x_3 = 0,8$ e $x_4 = 0,30$

α	x_1	R	\hat{y}
$+\infty$	0,260	0	9,45
1000	0,264	0,006	9,51
500	0,268	0,012	9,58
250	0,278	0,025	9,75
100	0,316	0,079	10,51
90	0,325	0,092	10,72
80	0,338	0,110	11,03
75	0,346	0,122	11,25
70	0,357	0,137	11,53
65	0,371	0,157	11,91
60	0,389	0,182	12,45
57,5	0,400	0,198	12,81

Na Tabela 4.4, apenas a variável x_1 é apresentada, isto justifica-se pelo fato das variáveis x_3 e x_4 serem constantes e a restrição da equação (4.21) poder ser escrita como $x_1 + x_2 = 0,52$. Logo, a variável x_2 encontra-se em função de x_1 .

Com relação a evolução dos valores de α , R , x_1 e \hat{y} , a resposta máxima é dada quando a variável x_1 atinge seu limite superior de 0,40. Deste modo, a tripla aplicação do processo de análise de cristas promoveu uma melhoria na

identificação da resposta máxima de $\hat{y} = 8,09$ (primeira aplicação do processo) para $\hat{y} = 12,81$ (terceira aplicação do processo).

Destaca-se que a sequência de aplicações do processo de análise de cristas deve-se encerrar quando for atingido um limite de uma variável ainda não fixada e o modelo não apresentar mais variáveis livres ou termina quando a variável \hat{y} atingir o máximo absoluto sem que um limite tenha sido rompido, ou seja, na prática os valores de \hat{y} crescem de acordo com a redução gradativa de α e para um determinado valor α , a variável resposta atinge o máximo absoluto e posteriormente começa a decrescer.

No próximo capítulo, apresentaremos uma aplicação prática dos experimentos do tipo mistura, em que além de definir o modelo, também haverá o interesse de maximizar a variável resposta através do procedimento de análise de cristas.

Capítulo 5

Aplicações

5.1 Introdução

O objetivo deste capítulo é a aplicação das técnicas apresentadas anteriormente a um conjunto de dados reais.

A aplicação consiste na obtenção de um modelo para prever o ganho de massa corporal em frangos, de acordo com a composição da dieta fornecida. O conjunto de dados foi extraído de Cornell (1990, p.526-531).

Neste experimento 30 grupos, cada um com 4 frangos, todos com massa e dimensões corporais uniformizadas, receberam uma dieta especial que consistia em um suplemento energético composto por três componentes: proteína, gordura e carboidrato.

Os frangos foram alimentados três vezes por dia, durante 10 dias, e a variável resposta medida em cada frango foi o ganho de massa corporal.

5.2. Descrição das variáveis

Este experimento envolve três variáveis independentes e uma variável resposta que serão analisadas em dois modelos diferentes. A descrição de cada uma das variáveis será apresentada a seguir:

Variáveis Independentes

- P: proporção de proteína, em gramas, contida no suplemento energético.
- G: proporção de gordura, em gramas, contida no suplemento energético.
- C: proporção de carboidrato, em gramas, contida no suplemento energético.

Variável Resposta

- MC: ganho de massa corporal médio por frango, em gramas, em que as médias foram obtidas para cada grupo com 4 frangos.

Há ainda uma variável complementar que seria o total de suplemento, em gramas, ingerido por cada grupo de 4 frangos, durante os 10 dias. Esta variável será denominada TSI. Segundo Cornell (1990), uma proposta inicial seria a adoção de um modelo sem levar em consideração a TSI, no entanto, após este primeiro estudo, o autor sugere a verificação da importância e influência desta variável na variável resposta original (MC).

Todas as variáveis deste experimento encontram-se na Tabela 5.1, em que cada linha corresponde aos dados de cada um dos 30 grupos com 4 frangos do experimento.

Tabela 5.1 : Dados do experimento Evolução da Massa Corporal em Frangos.

P	G	C	MC	TSI
0,05	0,06	0,89	59	288
0,05	0,22	0,73	60	262
0,05	0,86	0,09	29	162
0,05	0,70	0,25	35	188
0,05	0,38	0,57	53	237
0,05	0,54	0,41	48	212
0,12	0,06	0,82	82	288
0,12	0,22	0,66	90	262
0,12	0,86	0,02	63	161
0,12	0,70	0,18	73	187
0,12	0,38	0,50	100	237
0,12	0,54	0,34	80	212
0,19	0,06	0,75	105	288
0,19	0,22	0,59	105	262
0,19	0,38	0,43	120	238
0,19	0,54	0,27	109	213
0,19	0,70	0,11	110	187
0,26	0,06	0,68	136	288
0,26	0,70	0,04	125	188
0,26	0,54	0,20	133	212
0,26	0,38	0,36	130	237
0,26	0,22	0,52	141	263
0,33	0,06	0,61	141	289
0,33	0,22	0,45	156	263
0,33	0,38	0,29	157	238
0,33	0,54	0,13	153	213
0,40	0,06	0,54	156	289
0,40	0,22	0,38	179	263
0,40	0,38	0,22	163	238
0,40	0,54	0,06	143	213

De acordo com Cornell [1990], as proporções dos três componentes do suplemento energético são restritas aos limites descritos, a seguir:

$$0,05 \leq P \leq 0,40,$$

$$0,06 \leq G \leq 0,86 \text{ e}$$

$$0,02 \leq C \leq 0,89.$$

5.3 Análises

5.3.1 Modelo polinomial quadrático com variável resposta MC

Este experimento encontra-se nos padrões dos experimentos do tipo mistura, logo, segundo Cornell, o modelo polinomial quadrático na forma canônica deve ser adotado, conforme a equação a seguir:

$$MC = \beta_P P + \beta_G G + \beta_C C + \beta_{P,G} P \cdot G + \beta_{P,C} P \cdot C + \beta_{G,C} G \cdot C + e \quad (5.1)$$

Em seu estudo, Cornell obteve as estimativas dos parâmetros do modelo através do método de mínimos quadrados, além de calcular as estimativas dos respectivos desvios padrão. No entanto, o autor não abordou a possível existência do problema da multicolinearidade. Assim, nesta aplicação, temos basicamente dois objetivos, o primeiro consiste na verificação da existência da multicolinearidade e no seu eventual tratamento, o segundo é a maximização da variável resposta do modelo descrito na equação (5.1).

Deve-se destacar que a variável TSI não será considerada em um primeiro estudo.

Verificação da existência e tratamento da multicolinearidade para o modelo com variável resposta MC.

O modelo proposto encontra-se na equação (5.1) e os dados das variáveis predictoras (P, G e C) e os correspondentes valores da variável resposta (MC) são provenientes da Tabela 5.1.

De acordo com os aspectos teóricos já destacados, o cálculo dos FIV e do parâmetro de condicionamento (c) são bons indicadores para avaliar a presença da multicolinearidade.

Na Tabela 5.2 são apresentados os FIV das três variáveis independentes (P, G e C) e de seus respectivos produtos dois a dois, é importante destacar que a matriz de planejamento continha os dados originais do experimento.

Tabela 5.2: FIV das variáveis no modelo do segundo grau do experimento da Evolução da Massa Corporal em Frangos.

<i>Variável</i>	<i>FIV</i>
P	81,897
G	1,303
C	1,239
PG	241,758
PC	237,347
GC	13,209

Através da análise dos FIV, verifica-se a presença da multicolinearidade, caracterizada através dos valores elevados da variável P e dos produtos PG e PC. Quando calculado o parâmetro de condicionamento ($c = 6118,64$), pode-se avaliá-la como severa, segundo Montgomery, Peck e Vining (2006).

Uma primeira medida para reduzir a multicolinearidade é a padronização das variáveis preditoras. A Tabela 5.3 apresenta os valores dos FIV, após a padronização sugerida nas equações (3.6) e (3.7).

Tabela 5.3: FIV dos termos no modelo do segundo grau do experimento da Evolução da Massa Corporal em Frangos, após a padronização.

<i>Variável</i>	<i>FIV</i>
P	138,856
G	8,298
C	8,233
PG	63,261
PC	66,883
GC	5,728

Após a padronização, nota-se uma boa redução dos FIV dos produtos das variáveis independentes, no entanto, os FIV das variáveis independentes aumentam, destacando-se o valor obtido para a variável P. O problema da multicolinearidade ainda não está sanado, porém pode-se dizer que houve uma melhora no condicionamento, principalmente quando a referência é o parâmetro de condicionamento que neste caso foi reduzido a 1142,09.

Segundo St. John (1984), a transformação em pseudocomponentes frequentemente traz uma boa melhoria no condicionamento da matriz de planejamento a um custo muito reduzido. Deste modo, na Tabela 5.4 são apresentados os valores dos FIV, após a transformação em pseudocomponentes e, em seguida, a padronização das variáveis preditoras e seus respectivos produtos, de acordo com as equações (3.6) e (3.7).

Tabela 5.4: FIV dos termos no modelo do segundo grau do experimento da Evolução da Massa Corporal em Frangos, após a transformação em pseudocomponentes e padronização

<i>Variável</i>	<i>FIV</i>
P	80,3310
G	4,3635
C	3,8643
PG	30,9023
PC	39,8070
GC	3,4841

A transformação em pseudocomponentes e a padronização foram responsáveis por uma melhora no condicionamento, justificada através da comparação das Tabelas 5.3 e 5.4, sendo que todos os FIV da Tabela 5.4 são menores que os correspondentes na Tabela 5.3.

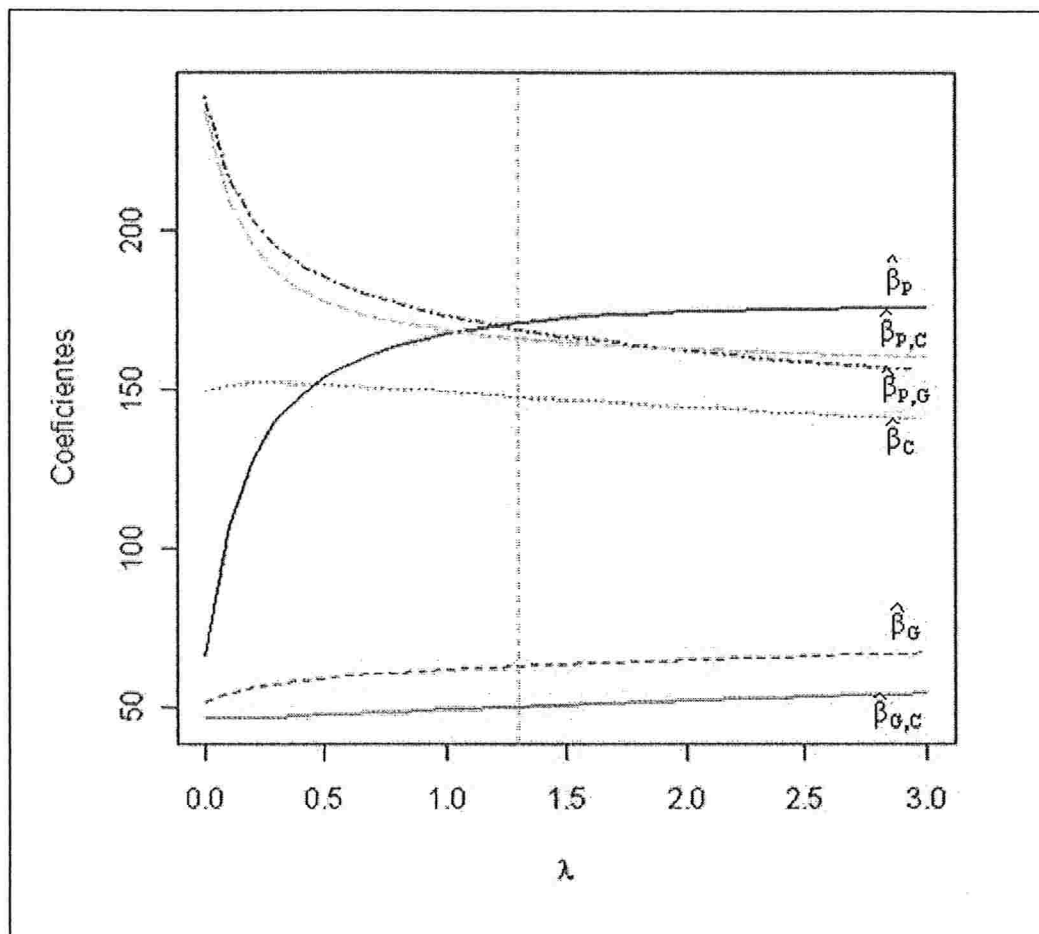
O parâmetro de condicionamento foi reduzido a 567,16, logo houve uma sensível redução da multicolinearidade. No entanto, segundo Marquardt (1970), quando os FIV são maiores que 10, existe problema com a multicolinearidade. Além disso, como o parâmetro de condicionamento encontra-se entre 100 e 1000, há indicação de multicolinearidade moderada, de acordo com Montgomery, Peck e Vining (2006).

Existe portanto a necessidade do emprego de novas técnicas para eliminar o problema da multicolinearidade, neste caso, a regressão em cristas torna-se uma opção bastante viável.

Deste modo, considerando todos os aspectos teóricos apresentados no Capítulo 3 sobre regressão em cristas, o traço de crista, neste caso, foi considerado como o método mais adequado para a obtenção da constante λ e os valores dos FIV foram utilizados apenas para verificar se o problema da multicolinearidade foi realmente extinto.

Na Figura 5.1, tem-se o traço de crista para o modelo polinomial quadrático apresentado na equação (5.1) com as variáveis transformadas em pseudocomponentes e padronizadas.

Figura 5.1: Traço de Crista para o modelo da equação (5.1)



Através da análise do traço de crista, verifica-se que para valores de λ entre 1,0 e 1,5, todas as estimativas dos parâmetros do modelo tornam-se mais estáveis, portanto o valor adotado para λ foi de 1,3. Assim, através da regressão em cristas tem-se:

$$\widehat{MC} = 170,902W_P + 62,640W_G + 147,727W_C + 168,974W_P \cdot W_G + 165,966W_P \cdot W_C + 49,922W_G \cdot W_C \quad (5.2)$$

Um estudo dos FIV mostra que através da técnica de regressão em cristas eliminou-se o problema da multicolinearidade.

Tabela 5.5: FIV dos termos no modelo do experimento da Evolução da Massa Corporal em Frangos, obtidos através de técnica de regressão em cristas.

<i>Variável</i>	<i>FIV</i>
P	0,074
G	0,117
C	0,125
PG	0,114
PC	0,107
GC	0,127

Na Tabela 5.5, nota-se que todos os FIV tornaram-se muito reduzidos, solucionando o problema de mau condicionamento que havia sido originalmente detectado. Torna-se importante destacar que os valores dos FIV obtidos encontram-se todos abaixo do mínimo teórico para os estimadores não viciados dos parâmetros, que seria de 1,0. Isto justifica-se, segundo Marquardt (1970), pelo fato do estimador da regressão em cristas ser viciado .

Análise de Cristas para o modelo com variável resposta MC

Uma vez definido o modelo, através da regressão em cristas, um segundo objetivo deste estudo é a maximização da resposta em função das variáveis preditoras. Desta forma, deseja-se determinar os valores das variáveis preditoras que maximizam o ganho de massa corporal. A análise de cristas será o recurso utilizado e a intenção é determinar quais proporções dos componentes do suplemento geram o maior ganho de massa corporal a partir do modelo pré-definido.

Inicialmente o modelo foi escrito na forma matricial proposta no Capítulo 4, dada por:

$$\hat{y} = b_0 + x'b + x'Bx,$$

em que b_0 é o vetor nulo, pois trata-se do modelo polinomial na forma canônica que não apresenta intercepto.

O modelo proposto na equação (5.2) encontra-se com as variáveis preditoras transformadas em pseudocomponentes e padronizadas, a padronização implica que as variáveis padronizadas W_P, W_G e W_C , quando somadas, não obrigatoriamente resultarão em um inteiro. Logo, para a análise de cristas foi retirada a padronização e manteve-se a forma de pseudocomponentes, que por definição satisfaz a condição de mistura da equação (2.2).

Por tratar-se do modelo polinomial quadrático na forma canônica, a padronização havia sido realizada de acordo com as equações (3.6) e (3.7), logo para que seja retirada, propõe-se a seguinte transformação:

$$\hat{\beta}_j = \frac{\hat{\beta}w_j}{\sqrt{\sum_{i=1}^n x_{ij}^2}}, \text{ em que } \hat{\beta}_j \text{ é o estimador do coeficiente de regressão para}$$

o modelo com as variáveis apenas transformadas em pseudocomponentes e $\hat{\beta}w_j$ é o correspondente estimador quando as variáveis encontram-se em pseudocomponentes e padronizadas.

Deste modo o modelo com as variáveis preditoras transformadas em pseudocomponentes e não padronizadas é dado por:

$$\widehat{MC} = 138,606P' + 24,205G' + 51,743C' + 357,772P' \cdot G' + 306,776P' \cdot C' + 73,575G' \cdot C'. \quad (5.3)$$

A matrizes b e B são apresentadas a seguir:

$$b = \begin{pmatrix} 138,606 \\ 24,205 \\ 51,743 \end{pmatrix} \text{ e } B = \begin{pmatrix} 0 & 178,886 & 153,388 \\ 178,886 & 0 & 36,787 \\ 153,388 & 36,787 & 0 \end{pmatrix}.$$

As matrizes (M e N) que compõem a equação matricial (4.9) de restrições inicialmente são matrizes linha. Apresentam como única restrição aquela que é característica dos experimentos do tipo mistura, em que a soma das proporções

de todos os componentes é constante e igual a 1, representada pela equação (4.6).

As matrizes M e N , de acordo com as equações (4.10) e (4.11), são dadas por:

$$M = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \quad \text{e} \quad N = \begin{pmatrix} \frac{1}{\sqrt{3}} \end{pmatrix}.$$

A matriz T foi construída respeitando o conjunto de equações (4.17), (4.18) e (4.19).

$$T = \begin{pmatrix} -0,707 & 0 & 0,707 \\ 0,408 & -0,816 & 0,408 \end{pmatrix}$$

O ponto inicial da análise de cristas, denominado por f , é o vetor coluna apresentado, a seguir, na forma transposta:

$$f' = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

Os próximos passos da análise de cristas foram apresentados no Capítulo 4, logo, neste momento, a proposta é a apresentação e discussão dos resultados obtidos e não uma descrição detalhada do procedimento. Deste modo a matriz TBT' é calculada e seus dois autovalores são -210,368 e -35,500. Assim, o caminho para o máximo absoluto por meio da análise de cristas é definido pelo estudo dos valores de α , a partir do infinito positivo, em que $x = f$ e $R = 0$ até o maior autovalor da matriz TBT' , no caso, -35,500. Destacando que para cada valor de α , serão apresentados como saída os valores correspondentes de P , G , C , R e \hat{y} (em que \hat{y} representa o previsor da variável resposta do modelo MC), que serão obtidos através do algoritmo de análise de cristas. De acordo com as diretrizes detalhadas no Capítulo 4, este algoritmo foi criado e desenvolvido no *software* R exclusivamente para esta dissertação e encontra-se no Apêndice 3.

Um comentário importante com relação ao processo de análise de cristas nesta aplicação envolve o fato de que as variáveis preditoras (P , G e C) quando inseridas no algoritmo da análise de cristas encontram-se transformadas em

pseudocomponentes, conforme o modelo da equação (5.3). No entanto, a saída do algoritmo apresenta as variáveis preditoras na forma original, ou seja, a transformação em pseudocomponentes foi retirada pelo algoritmo. Esta forma de apresentação é justificada pelo fato de um dos critérios de parada ser através dos limites superior e inferior da variável e estes se encontram na forma original. Uma outra razão, além da facilidade de comparação com os limites das variáveis, é possibilitar ao pesquisador uma visualização imediata do resultado obtido.

Tabela 5.6: Caminho de cristas para obtenção do máximo absoluto do modelo do experimento da Evolução da Massa Corporal em Frangos.

α	P	G	C	R	\hat{y}
964,500	0,386	0,325	0,288	0,0650	162,57
864,500	0,390	0,323	0,287	0,0711	163,32
764,500	0,396	0,320	0,284	0,0784	164,21
664,500	0,402	0,317	0,281	0,0873	165,26
564,500	0,410	0,313	0,277	0,0986	166,54
464,500	0,421	0,307	0,272	0,1132	168,13
364,500	0,434	0,300	0,266	0,1329	170,12
264,500	0,454	0,290	0,256	0,1610	172,67
164,500	0,485	0,275	0,240	0,2040	175,95
64,500	0,538	0,253	0,209	0,2789	179,81
-35,500	NC	NC	NC	NC	NC

NC: Não consta

Deste modo, foram estudados 10 valores de α , em que o ponto de partida teórico seria $+\infty$, no entanto, o procedimento foi iniciado com um valor de α de 964,500, ou seja, 1000 unidades a mais que o maior autovalor da matriz TBT' e os valores de α foram obtidos a partir de uma redução sequencial de 100 unidades, conforme verificado na Tabela 5.6.

A Tabela 5.6 apresenta além dos valores de α , os valores correspondentes de P, G, C, R e \hat{y} obtidos conforme a sequência de passos a) a d) descrita na Seção 4.2. A conclusão mais importante, extraída da Tabela 5.6, é

que o limite superior da variável P que é de 0,40 foi extrapolado e isto pode ser verificado nas linhas em destaque na Tabela 5.6. Logo, uma nova restrição ($P = 0,40$) será adicionada às matrizes M e N , que serão dadas por:

$$M = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 1 & 0 & 0 \end{pmatrix} \quad \text{e} \quad N = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ 0,4023 \end{pmatrix}.$$

Destaca-se que o valor 0,4023, presente na segunda linha da matriz N , corresponde à variável P , que quando originalmente vale 0,40, ao ser transformada em pseudocomponentes, resulta em 0,4023.

A nova matriz T é dada por:

$$T = (0 \quad -0,707 \quad 0,707).$$

O ponto inicial da análise de cristas, denominado por f , é o vetor coluna necessário para iniciar o processo de análise de cristas. Conforme dito anteriormente, uma sugestão de Draper e Pukelsheim (2002) é que este ponto seja o centro da região de mistura, logo o vetor transposto deve ser construído através da forma $\left(\frac{1}{k} \quad \frac{1}{k} \quad \frac{1}{k}\right)$. No entanto, quando uma ou mais variáveis encontram-se fixadas, como a variável $P=0,40$, a sugestão é centralizar as variáveis restantes, que neste caso são duas (G e C) e a soma de suas proporções é de 0,60. Logo, para construir f , dividimos 0,60 por dois e obtivemos o seguinte vetor na forma transposta:

$$f' = (0,40 \quad 0,30 \quad 0,30).$$

A matriz TBT' foi calculada e o seu único autovalor é -36,769. Logo, o caminho para o máximo absoluto por meio da análise de cristas é definido pelo estudo dos valores de α , a partir $+\infty$ até o único autovalor, no caso, -36,769.

Assim, foi novamente utilizado o algoritmo do *software* R que se encontra no Apêndice 3, em que os valores de α estudados se iniciaram em 1000

e foram reduzidos de 0,5 unidade de forma sequencial até -36,5, que consiste em uma aproximação para o autovalor da matriz TBT' .

Neste momento, o intervalo entre os valores de α devem ser reduzidos, no entanto a amplitude deste intervalo depende da precisão exigida para a obtenção do valor máximo do modelo.

Em função de terem sido estudados muitos valores para α , a apresentação de uma tabela completa com todos estes valores seria inviável, logo será apresentada uma parte da tabela original (Tabela 5.7), logicamente aquela correspondente aos valores de α que levaram à maximização da variável resposta do modelo.

Tabela 5.7: Caminho de cristas para obtenção do máximo absoluto do modelo do experimento da Evolução da Massa Corporal em Frangos com a restrição $P = 0,4$.

α	P	G	C	R	\hat{y}
2,5	0,4000	0,2811	0,3189	0,0633	165,08831
2,0	0,4000	0,2806	0,3194	0,0641	165,08855
1,5	0,4000	0,2801	0,3199	0,0649	165,08874
1,0	0,4000	0,2796	0,3204	0,0658	165,08888
0,5	0,4000	0,2790	0,3210	0,0666	165,08896
0,0	0,4000	0,2785	0,3215	0,0675	165,08899
-0,5	0,4000	0,2779	0,3221	0,0685	165,08896
-1,0	0,4000	0,2773	0,3227	0,0694	165,08886
-1,5	0,4000	0,2767	0,3233	0,0704	165,08869
-2,0	0,4000	0,2761	0,3239	0,0714	165,08844
-2,5	0,4000	0,2755	0,3245	0,0725	165,08810

Através da Tabela 5.7, verifica-se que a maximização ocorre quando α é igual a zero, pois as estimativas da variável resposta, que vinham crescendo com a diminuição dos valores de α , atingem o máximo quando $\alpha = 0$ e posteriormente começam a decrescer. Logo, um dos critérios de parada foi atingido que consiste nas estimativas de \hat{y} que vinham crescendo de acordo com a redução de α , apresentarem um máximo e a partir dele decrescem continuamente.

Logo, de acordo com o procedimento de análise de cristas, a estimativa do valor máximo da variável resposta do modelo da equação (5.2) é 165,08899g , obtido para proporção de proteína (P) de 0,4000, proporção de gordura (G) de 0,2785 e a proporção de carboidrato (C) de 0,3215.

5.3.2 Introdução da variável complementar TSI no modelo.

Após este primeiro estudo foi realizado um segundo, considerando a variável complementar que seria o total de suplemento, em gramas, ingerido por cada grupo de 4 frangos, durante os 10 dias (TSI).

A introdução da variável TSI no modelo não poderia descaracterizar o experimento do tipo mistura, ou seja, a restrição da equação (2.2) não deveria ser violada. Logo, optou-se por introduzir a variável TSI na variável resposta.

Assim, apresentamos um novo modelo, não considerado em Cornell (1990), em que a nova variável resposta (Y_N) foi gerada através da equação a seguir:

$$Y_N = \frac{MC}{\left(\frac{TSI}{4}\right)}. \quad (5.4)$$

Deste modo, verifica-se que o ganho de massa corporal médio por frango (MC) foi dividido por um quarto do TSI, sendo que a divisão por quatro justifica-se pelo fato da variável TSI corresponder ao total de suplemento consumido pelos quatro frangos do grupo nos 10 dias de experimento e a intenção foi construir uma variável que apresentasse como unidade o ganho, em gramas, de massa corporal por grama de suplemento consumido.

A equação (5.5) apresenta o modelo definido por esta nova proposta:

$$Y_N = \beta_P P + \beta_G G + \beta_C C + \beta_{P,G} P \cdot G + \beta_{P,C} P \cdot C + \beta_{G,C} G \cdot C + e \quad (5.5)$$

Este modelo origina-se do modelo da equação (5.1) e a única alteração realizada ocorreu na variável resposta, conforme descrito anteriormente.

A Tabela 5.8 apresenta os valores das variáveis preditoras originais do experimento, que não sofreram nenhuma alteração, associadas à nova variável resposta definida pela equação (5.4):

Tabela 5.8 : Dados do experimento Evolução da Massa Corporal em Frangos associados à variável resposta Y_N

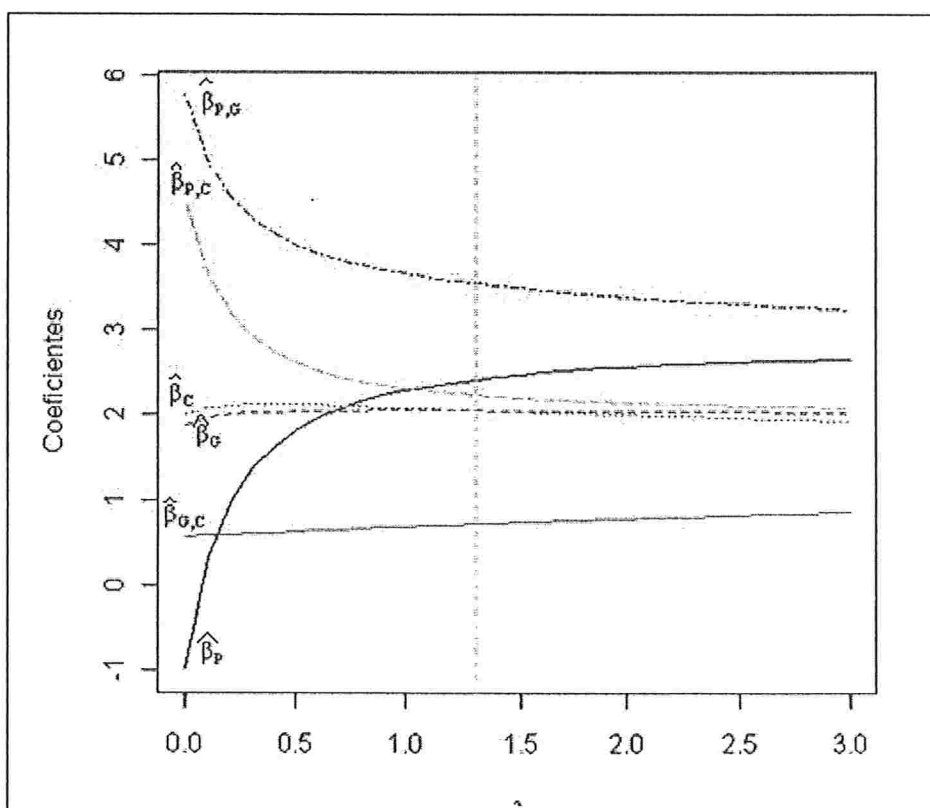
P	G	C	Y_N
0,05	0,06	0,89	0,81944
0,05	0,22	0,73	0,91603
0,05	0,86	0,09	0,71605
0,05	0,70	0,25	0,74468
0,05	0,38	0,57	0,89451
0,05	0,54	0,41	0,90566
0,12	0,06	0,82	1,13889
0,12	0,22	0,66	1,37405
0,12	0,86	0,02	1,56522
0,12	0,70	0,18	1,56150
0,12	0,38	0,50	1,68776
0,12	0,54	0,34	1,50943
0,19	0,06	0,75	1,45833
0,19	0,22	0,59	1,60305
0,19	0,38	0,43	2,01681
0,19	0,54	0,27	2,04695
0,19	0,70	0,11	2,35294
0,26	0,06	0,68	1,88889
0,26	0,70	0,04	2,65957
0,26	0,54	0,20	2,50943
0,26	0,38	0,36	2,19409
0,26	0,22	0,52	2,14449
0,33	0,06	0,61	1,95156
0,33	0,22	0,45	2,37262
0,33	0,38	0,29	2,63866
0,33	0,54	0,13	2,87324
0,40	0,06	0,54	2,15917
0,40	0,22	0,38	2,72243
0,40	0,38	0,22	2,73950
0,40	0,54	0,06	2,68545

Verificação da existência e tratamento da Multicolinearidade para o modelo com a variável resposta Y_N

Neste caso, como não houve alteração nas variáveis preditoras, os problemas relacionados à multicolinearidade são os mesmos do modelo da equação (5.1). Portanto, as medidas tomadas para corrigir o problema também foram as mesmas. Assim, será destacado, nesta seção, apenas o procedimento de regressão em cristas, pois como o estimador em cristas é definido em função da variável resposta, as estimativas naturalmente sofrerão alterações.

De acordo com esclarecimentos anteriores, o traço de crista foi avaliado como o método mais adequado para a obtenção da constante λ . Portanto, na Figura 5.2, apresenta-se o traço de crista para o modelo polinomial quadrático da equação (5.5) com as variáveis preditoras transformadas em pseudocomponentes e padronizadas.

Figura 5.2: Traço de Crista para o modelo da equação (5.2)



Analisando o traço de crista, verifica-se que, novamente como no modelo anterior, para valores de λ entre 1,0 e 1,5, as estimativas dos parâmetros do modelo tornam-se mais estáveis, portanto o valor adotado para λ foi de 1,3. Assim, através da regressão em cristas tem-se:

$$\hat{Y}_N = 2,4163W_P + 2,0561W_G + 2,0584W_C + 3,5560W_P \cdot W_G + 2,2375W_P \cdot W_C + 0,7185W_G \cdot W_C \quad (5.6)$$

Os Fatores de Inflação da Variância foram os mesmos apresentados na Tabela 5.5, pois os FIV dependem apenas dos valores da matriz de planejamento e do valor adotado para λ . Logo, com base na Tabela 5.5, conclui-se que o problema da multicolinearidade foi solucionado.

Análise de Cristas para o modelo com variável resposta Y_N

De acordo com as orientações anteriores, no procedimento de análise de cristas, encontra-se a seguir o modelo ajustado com os pseudocomponentes sem padronização:

$$\hat{Y}_N = 1,9596P' + 0,7945G' + 0,7210C' + 7,5293P' \cdot G' + 4,1359P' \cdot C' + 1,0590G' \cdot C' \quad (5.7)$$

A matrizes b e B são apresentadas a seguir:

$$b = \begin{pmatrix} 1,9596 \\ 0,7945 \\ 0,7210 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 0 & 3,7646 & 2,0679 \\ 3,7646 & 0 & 0,5295 \\ 2,0679 & 0,5295 & 0 \end{pmatrix}.$$

As matrizes (M e N) que compõem a equação matricial (4.9) de restrições são dadas por:

$$M = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \quad \text{e} \quad N = \begin{pmatrix} \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Deste modo, tem-se a matriz T :

$$T = \begin{pmatrix} -0,707 & 0 & 0,707 \\ 0,408 & -0,816 & 0,408 \end{pmatrix}.$$

O ponto inicial da análise de cristas, na forma transposta, é dado por:

$$f' = \left(\frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3} \right).$$

Tabela 5.9: Caminho de cristas para obtenção do máximo absoluto do modelo da equação (5.7)

α	P	G	C	R	\hat{y}
13,1143	0,390	0,347	0,263	0,0787	2,7485
12,1143	0,393	0,348	0,259	0,0843	2,7600
11,1143	0,397	0,348	0,255	0,0908	2,7731
10,1143	0,401	0,349	0,250	0,0983	2,7883
9,1143	0,406	0,350	0,244	0,1073	2,8060
8,1143	0,412	0,351	0,237	0,1183	2,8272
7,1143	0,419	0,353	0,228	0,1318	2,8529
6,1143	0,428	0,357	0,215	0,1491	2,8849
5,1143	0,440	0,362	0,198	0,1722	2,9262
4,1143	0,455	0,370	0,175	0,2046	2,9821
3,1143	0,476	0,386	0,138	0,2547	3,0639
2,1143	0,509	0,419	0,072	0,3443	3,1998
1,1143	NC	NC	NC	NC	NC

NC: Não consta

A matriz TBT' foi calculada e os seus dois autovalores são -2,9264 e 1,1143. Assim, o caminho para o máximo absoluto por meio da análise de cristas é definido pelo estudo dos valores de α , a partir de $+\infty$, quando $x = f$ e $R = 0$ até o maior autovalor da matriz TBT' , no caso, 1,1143.

Deste modo, foram estudados 100 valores de α , com ponto de partida teórico $+\infty$, no entanto, o procedimento foi iniciado com um valor para α de 100 e os outros valores de α foram obtidos a partir de uma redução sequencial de 1 unidade. Na Tabela 5.9, destaca-se o trecho de maior relevância do caminho de cristas.

A conclusão mais importante, extraída da Tabela 5.9, é que o limite superior da variável P que é de 0,40 foi ultrapassado e isto pode ser verificado nas linhas em destaque. Logo, uma nova restrição ($P = 0,40$) será adicionada às matrizes M e N , que serão dadas por:

$$M = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 1 & 0 & 0 \end{pmatrix} \quad \text{e} \quad N = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ 0,4023 \end{pmatrix}.$$

A nova matriz T é dada por:

$$T = (0 \quad -0,707 \quad 0,707)$$

O ponto inicial da análise de cristas, f , é tal que

$$f' = (0,40 \quad 0,30 \quad 0,30).$$

A matriz TBT' foi calculada e o seu único autovalor é -0,5293. Logo, o caminho para o máximo absoluto por meio da análise de cristas é definido pelo estudo dos valores de α , a partir $+\infty$ até o único autovalor, no caso, -0,5293.

Os valores de α estudados, se iniciaram em 10 e foram reduzidos de 0,01 de forma sequencial até -0,5, que consiste em uma aproximação para o autovalor da matriz TBT' .

Tabela 5.10: Caminho de cristas para obtenção do máximo absoluto do modelo da equação (5.7) com a restrição $P=0,4$

α	P	G	C	R	\hat{y}
0,74	0,4000	0,5665	0,0335	0,4007	3,0609
0,73	0,4000	0,5684	0,0316	0,4039	3,0628
0,72	0,4000	0,5704	0,0296	0,4071	3,0647
0,71	0,4000	0,5724	0,0276	0,4104	3,0666
0,70	0,4000	0,5745	0,0255	0,4137	3,0685
0,69	0,4000	0,5766	0,0234	0,4171	3,0705
0,68	0,4000	0,5787	0,0213	0,4206	3,0725
0,67	0,4000	0,5809	0,0191	0,4241	3,0745
0,66	0,4000	0,5831	0,0169	0,4276	3,0765
0,65	0,4000	0,5853	0,0147	0,4312	3,0785
0,64	0,4000	0,5876	0,0124	0,4349	3,0806
0,63	0,4000	0,5899	0,0101	0,4387	3,0827

Através da Tabela 5.10, em que encontram-se os valores para as conclusões finais do processo de análise de cristas, verifica-se que a maximização ocorre quando α está entre 0,67 e 0,68, pois a variável C ultrapassa o seu limite inferior de 0,02. Logo, um dos critérios de parada foi atingido e a estimativa do máximo da variável resposta do modelo da equação (5.7) é 3,0737, sendo que as proporções dos componentes que geram a maximização são de 0,40 para a proteína (P), 0,58 para a gordura (G) e 0,02 para o carboidrato (C). Destaca-se que a variável resposta (Y_N) é adimensional, pois trata-se de uma razão entre massas, ou seja, o ganho de massa corporal médio, em gramas, dividido pela massa média de suplemento consumida por cada frango, também em gramas.

5.4 Considerações finais

Nesta aplicação, analisou-se um experimento do tipo mistura, cuja idéia principal era prever e maximizar o ganho de massa corporal em frangos, em função da composição do suplemento alimentar fornecido.

Em ambos os modelos, as técnicas apresentadas nesta dissertação foram bastante eficientes. A multicolinearidade presente inicialmente foi tratada principalmente através do processo de regressão em cristas.

O fato das variáveis preditoras estarem padronizadas não inviabiliza o processo de maximização, pois transformações podem ser realizadas, conforme verificamos neste capítulo.

A maximizações das variáveis respostas de ambos os modelos foram satisfatórias e comparadas com as maximizações obtidas através de algoritmos de maximização fechados como a ferramenta *Solver* do *Excel*. Os resultados obtidos foram similares, mas pode-se destacar como vantagem do processo de análise de cristas, a possibilidade de acompanhar de forma mais detalhada o processo de otimização e a facilidade de intervir no processo quando trabalha-se com conjuntos de dados mais problemáticos.

CAPÍTULO 6

Considerações Finais

O objetivo do presente trabalho foi utilizar os modelos de regressão nos experimentos do tipo mistura como o intuito de prever uma determinada variável resposta através de um conjunto de variáveis preditoras que são componentes de uma mistura.

No presente trabalho discutimos o processo de construção dos modelos para os experimentos do tipo mistura, destacando a importância dos modelos polinomiais na forma canônica desenvolvidos por Scheffé (1958).

Abordamos também um problema muito comum nos modelos que envolvem mistura que é a multicolinearidade e discutimos as técnicas para solucioná-lo.

Devido ao grande interesse dos pesquisadores em encontrar o valor máximo ou mínimo da variável resposta depois do modelo definido, apresentamos, como recurso para este propósito, o procedimento de Análise de Cristas.

Todos esses aspectos, foram exemplificados em uma aplicação prática da utilização dos modelos de regressão nos experimentos do tipo mistura, conforme verificado no Capítulo 5.

Dada a extensão do tema, procuramos abordar os aspectos que acreditamos serem os mais importantes, tanto do ponto de vista teórico quanto aplicado. Finalizando, gostaríamos de salientar que as técnicas expostas na presente dissertação têm amplas aplicações. Isto pode ser observado na grande variedade de exemplos que descrevemos e nos demais exemplos encontrados na literatura relativa ao assunto.

Vários temas de pesquisa podem ser sugeridos como trabalhos futuros. Existe extensa literatura na área de planejamentos amostrais para experimentos do tipo mistura. Este ponto foi brevemente comentado no Capítulo 2, mas encontra-se discutido com detalhes em vários artigos, dentre os quais destacamos Goos e Donev (2006), Piepel (1983) e Vuchkov, Damgaliev e Yontchev(1981).

Destaca-se ainda o artigo de Montgomery e Voth (1994), que seria um excelente ponto de partida para análise de influência e diagnóstico em modelos ajustados para experimentos do tipo mistura.

Apêndice 1

O número de pontos no espaço simplex em uma malha $\{k, m\}$ é dado por $\binom{m+k-1}{m}$. Esta combinação de $m+k-1$ elementos tomados m a m é justificada pelo fato de cada ponto da malha simplex apresentar como condição a equação (2.2), ou seja, cada ponto do simplex consiste em um conjunto de k variáveis onde cada variável x_j é um múltiplo de $\frac{1}{m}$ de modo que a somatória de todas as variáveis x_i em cada ponto seja sempre igual a 1. Logo, ao efetuar a multiplicação da equação (2.2) por m , tem-se:

$$x'_1 + x'_2 + \dots + x'_k = m, \text{ em que } x'_j = 0, 1, 2, \dots, m. \quad (\text{A.1})$$

A obtenção da quantidade de soluções inteiras não negativas da equação (A.1) consiste em um clássico problema de análise combinatória onde a estratégia consiste em codificar as diferentes soluções em sequências que trabalhem apenas com o algarismo 1 e o sinal de adição. Por exemplo, a equação

$$x_1 + x_2 + x_3 = 5,$$

apresenta diferentes soluções inteiras e não negativas, entre elas:

$$S = \{(1, 2, 2)\}.$$

O código desta solução é:

$$1+11+11.$$

Logo, o total de soluções consiste em encontrar o total de sequências distintas para m algarismos 1 e $k-1$ sinais de adição. Desta forma, o número de pontos do espaço simplex é igual à quantidade de soluções inteiras não negativas da equação (A.1) que é dada por $\binom{m+k-1}{m}$.

Apêndice 2

Os pseudocomponentes são definidos a partir dos componentes e seus respectivos limites inferiores (l_j), em que $l_j \geq 0$.

A transformação em pseudocomponentes nos experimentos do tipo mistura é realizada a partir da restrição (2.2):

$$x_{ij} + x_{ip} + \dots + x_{ik} = 1 .$$

Assim,

$$x_{ij} - l_j + x_{ip} - l_p + \dots + x_{ik} - l_k = 1 - \sum_1^k l_j.$$

Logo,

$$\frac{x_{ij}-l_j}{1-\sum_1^k l_j} + \frac{x_{ip}-l_p}{1-\sum_1^k l_j} + \dots + \frac{x_{ik}-l_k}{1-\sum_1^k l_j} = 1.$$

Deste modo, cada pseudocomponente é dado por

$$x'_{ij} = \frac{(x_{ij}-l_j)}{(1-\sum l_j)}, \text{ com } j=1,2,\dots,k.$$

Destaca-se que mesmo com a transformação, a restrição de mistura continua a ser respeitada, ou seja, a soma dos pseudocomponentes de todos os componentes é sempre igual a 1.

Antes da utilização desta transformação deve-se verificar se a condição $\sum_1^k l_j < 1$ é satisfeita.

Apêndice 3

Neste apêndice encontram-se os recursos e os programas computacionais elaborados no *software* R para esta dissertação.

Matriz de Planejamento original

```
Dados <- cbind(P,G,C)
Dados1 <- cbind(Dados, P*G, P*C, G*C)
```

Cálculo do FIV da matriz de planejamento original

```
Matriz1 <- solve( t(Dados1) %*% Dados1 )
FIV1 <- diag(Matriz1);FIV1
```

Cálculo dos autovalores da matriz de planejamento original

```
AV1 <- eigen(t(Dados1) %*% Dados1)
```

Padronização das variáveis preditoras na matriz de planejamento original

```
W <- matrix(0, 30, 6)
for(i in 1:30) {
  for(j in 1:6) {
    W[i,j] <- Dados1[i,j] / sqrt( sum( Dados1[, j]^2 ) )}
```

Cálculo do FIV da matriz de planejamento padronizada

```
MatrizW <- solve( t(W) %*% W )
FIVW <- diag(MatrizW); FIVW
```

Cálculo dos autovalores da matriz de planejamento padronizada

```
AVW <- eigen(t(W) %*% W)
```

Transformação em pseudocomponentes das variáveis preditoras na matriz de planejamento original

```
LInf <- numeric()
for(j in 1:3){
  LInf[j] <- min(Dados[, j])}
```



```

Pseudo <- matrix(0, 30, 3)
for(i in 1:30){
  for(j in 1:3){
    Pseudo[i,j] <- (Dados[i, j] - LInf[j])/(1 - sum(LInf)) } }
Pseudo1 <- cbind(Pseudo, Pseudo[,1]*Pseudo[,2], Pseudo[,1]*Pseudo[,3],
Pseudo[,2]*Pseudo[,3])

```

Padronização dos pseudocomponentes das variáveis preditoras na matriz de planejamento original

```

WPseudo <- matrix(0, 30, 6)
for(i in 1:30){
  for(j in 1:6){
    WPseudo[i,j] <- Pseudo1[i,j] / sqrt( sum( Pseudo1[, j]^2 ) ) } }

```

Cálculo do FIV da matriz de planejamento com as variáveis preditoras transformadas em pseudocomponentes e posteriormente padronizadas

```

MatrizWPseudo <- solve( t(WPseudo) %*% WPseudo )
FIVWPseudo <- diag(MatrizWPseudo); FIVWPseudo

```

Cálculo dos autovalores da matriz de planejamento com as variáveis preditoras transformadas em pseudocomponentes e posteriormente padronizadas

```

AVWPseudo <- eigen(t(WPseudo) %*% WPseudo)

```

Procedimento de Regressão em Cristas

```

lm.ridge(Y~WPseudo[,1]+WPseudo[,2]+WPseudo[,3]+WPseudo[,4]+
WPseudo[,5]+ WPseudo [,6]-1, lambda=seq(0,3,0.1))

```

Cálculo dos FIV no Procedimento de Regressão em Cristas

```

Id <- diag(1,6)
Lambda <- seq(0, 1.4, 0.1)
N <- length(Lambda)
FIVRC <- matrix(0, 6, N)
for(i in 1:N)

```

```
{FIVRC[,i] <- diag( solve( t(WPseudo) %*% WPseudo + Lambda[i]*Id ) %*%
t(WPseudo) %*% WPseudo %*% solve( t(WPseudo) %*% WPseudo +
Lambda[i]*Id ) ) }
```

Procedimento de Análise de Cristas

```
Funcao <- function(b, B, M, Ene, T, f)
```

```
{Av <- T %*% B %*% t(T)
```

```
MAv <- max(eigen(Av)$values)
```

```
min <- (1 - sum(LInf))
```

```
Alfa <- seq(MAv, Limite Superior, Intervalo de variação)
```

Comentário: O limite superior de α e o respectivo intervalo de variação devem ser especificados em cada nova execução do procedimento.

```
Repet <- length(Alfa)
```

```
Res <- matrix(0,1,6)
```

```
for(i in 1:Repet)
```

```
{
```

```
Theta <- solve(M %*% solve(B - Alfa[i]*Id) %*% t(M)) %*% ( 2*Ene +
( M %*% solve(B - Alfa[i]*Id) ) %*% (b + 2*Alfa[i]*f) )
```

```
PseudoRC <- 0.5*solve(B - Alfa[i]*Id) %*% (t(M) %*% Theta - b - 2*Alfa[i]*f)
```

```
R <- sqrt(t(PseudoRC -f) %*% (PseudoRC -f))
```

```
Y <- t(PseudoRC) %*% b + t(PseudoRC) %*% B %*% PseudoRC
```

```
X <- min* PseudoRC + LInf
```

```
Res <- rbind(Res, c(Alfa[i], X, R, Y))
```

```
return(Res)}
```

```
Funcao(b, B, M, Ene, T, f)
```

Referências Bibliográficas

ANIK, S. T. e SUKUMAR, L. (1981), "Extreme Vertexes Design in Formulation Development: Solubility of Butoconazole Nitrate in a Multicomponent System", *Journal of Pharmaceutical Sciences*, 70, p. 897–900.

CHAMIE, S. (1994). *Encapsulamento de resíduos de lamas galvânicas através da solidificação em matriz de cimento*. Dissertação de Mestrado – Escola Politécnica da Universidade de São Paulo, São Paulo, SP, p.246.

CORNELL, J. A. (1990). *Experiments with Mixtures: Designs, Models, and the Analysis of Mixture Data*, Segunda Edição, John Wiley & Sons, New York. p.632

DRAPER, N.R. (1963), "Ridge Analysis of Response Surfaces", *Technometrics* 5, p. 469-479.

DRAPER, N. R. e PUKELSHEIM, F. (2000), "Ridge Analysis of Mixture. Response Surfaces". *Statistics & Probability Letters*, 48, p. 131–140.

DRAPER, N. R. e PUKELSHEIM, F. (2002), "Generalized Ridge Analysis Under Linear Restrictions, with Particular Applications to Mixture Experiments Problems". *Technometrics* 44, p. 250-255.

DRAPER, N.R. e SMITH, H. (1981). *Applied Regression Analysis*, Segunda Edição, John Wiley & Sons, New York, p. 407.

GOOS, P. e DONEV, A.N. (2006), "The D-optimal design of blocked experiments with mixture components" , *Journal of Quality Technology*, 38, n. 4, p. 319-332.

GORMAN, J. W. (1970), "Fitting Equations to Mixture Data With. Restraints on Compositions," *Journal of Quality Technology*, 2, p. 186-194.

GUNST R. F. e HUA T.A. (1978) " A Study of techniques for selecting ridge parameter values". *Joint Annual Meetings of the American Statistical Association and the Biometric Society*, San Diego, CA.

- GUNST, R.F. e MASON, R.L. (1980). *Regression Analysis and Its Application*. Marcel Dekker, New York, p. 402.
- HOERL, A. E. (1959). "Optimum Solution of Many Variables Equations", *Chemical Engineering Progress*, 55, p. 67-78
- HOERL, A. E. e KENNARD, R. W. (1970). "Ridge regression: biased estimation for nonorthogonal problems". *Technometrics*, 12, p.55-67.
- HOERL, R.W. (1987). "The application of ridge techniques to mixture data: ridge analysis". *Technometrics* 29, p. 161–172
- KIEFER, J. (1961). "Optimum designs in regression problems, II". *Annals of Mathematical Statistics* 32, p. 298-325.
- MARQUARDT, D. W. (1970). "Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation." *Technometrics* 12, p. 591-612.
- MARQUARDT, D.W. e SNEE, R.D. (1975), "Ridge Regression in. Practice," *The American Statistician*, 29, p. 3-20.
- MCLEAN, R.A. e ANDERSON V.L. (1966), "Extreme vertices design of mixture experiments", *Technometrics* 8 , p. 447-454.
- MONTGOMERY D.C., PECK, E.A. e VINING, G.G. (2006). *Introduction to Linear Regression. Analysis*, 4th edition, Wiley, New York, p.612.
- MONTGOMERY D. C. e VOTH S.R. (1994). "Multicollinearity and Leverage in Mixture Experiments". *Journal of Quality Technology*, 26, n.2, p.96-108.
- OISHI, J. (1983). *Regressão sobre cristas*. Dissertação de Mestrado – Instituto de Matemática e Estatística da Universidade de São Paulo, São Paulo, SP, p.100.
- PIEPEL, G. F. (1982) "Measuring component effects in constrained mixture experiments", *Technometrics*, 24, p. 29–39.
- PIEPEL, G. F. (1983). "Defining Consistent Constraint Regions in Mixture Experiments," *Technometrics* , 25, p. 97-101.

RYAN, T.P. (1997). *Modern Regression Methods*, John Wiley & Sons, New York, p.515.

SCHEFFÉ, H. (1958). "Experiments with mixtures". *Journal of the Royal Statistics Society B20*, p. 344-360.

SCHEFFÉ, H. (1963). "The simplex-centroid design for experiments with mixtures". *Journal of the Royal. Statistical Society B*. 25, p. 235-263

SNEE, R. D. (1975). "Experimental designs for quadratic models in constrained mixture spaces", *Technometrics*, 17, p. 149-159.

SNEE, R. D. e MARQUARDT, D. W. (1976). "Screening concepts and designs for experiments with mixtures". *Technometrics* 18, p. 19-29.

SNEE, R. D., e RAYNER, A. A. (1982), "Assessing the Accuracy of Mixture Model Regression Calculations". *Journal of Quality. Technology*, 14, p.67-79.

ST. JOHN, R. C. (1984). "Experiments with Mixtures, Ill-Conditioning and Ridge Regression". *Journal of Quality Technology*, 16, n. 2, p. 81-96.

STEINBERG, M.D. e HUNTER. W.G. (1984). "Experimental design: review and comment". *Technometrics*, 26, n.2, p.71-130.

VUCHKOV, I. N., DAMGALIEV, D. L. e YONTCHEV, C. A. (1981). "Sequentially generated second order quasi D-optimal designs for experiments with mixture and process variables", *Technometrics*, 23, p. 233-238.