

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
CADERNOS DE MATEMÁTICA E ESTATÍSTICA
SÉRIE B: TRABALHO DE APOIO DIDÁTICO

METODOLOGIA DE SUPERFÍCIE DE RESPOSTA:
UMA ABORDAGEM INTRODUTÓRIA

JOÃO RIBOLDI
LÍDIA DO CARMO SOUZA CABRITA DO NASCIMENTO

262612-5

SÉRIE B, N° 25
PORTO ALEGRE, NOVEMBRO DE 1994

PREFÁCIO

O presente material foi organizado com o objetivo de servir como apoio didático ao Curso de Metodologia de Superfície de resposta, ministrado no Curso de Pós-Graduação, a nível de mestrado, em Ciência dos Alimentos da UFSC.

Apesar do objetivo específico, seguramente servirá de apoio didático às disciplinas de Graduação e Pós-Graduação que ministramos na UFRGS, bem como poderá servir de suporte a outros cursos a serem ministrados sobre o assunto.

O material tem por meta servir como um guia introdutório à Metodologia de Superfície de Resposta, e evidentemente, não esgota o assunto, nem é suficiente para os interessados em se especializar sobre o mesmo.

A literatura sobre Metodologia de Superfície de Resposta é vasta, com um grande número de trabalhos científicos publicados na área, existindo, sobre o assunto, três livros específicos (BOX e DRAPER, 1987; KHURI e CORNELL, 1987 e MYERS, 1971) e capítulos em bons livros de Planejamento de Experimentos, como MONTGOMERY, 1971; BOX, HUNTER e HUNTER, 1978; COCHRAN e COX, 1957 e DAVIES, 1956.

O presente material constitui, basicamente, numa revisão e ampliação da monografia da Lídia do Nascimento, elaborada para obtenção do título de Bacharel em Estatística na UFRGS, sob nossa orientação, em 1991.

Agradecemos aos bolsistas Stela, Flávio e André pelo trabalho de digitação e pelo trabalho de computação na solução dos exemplos utilizando o SAS e a Professora Dinara Fernandez pelas sugestões.

Porto Alegre, março de 1994

Prof. João Riboldi

Í N D I C E

1. INTRODUÇÃO	01
1.1. Características Gerais	01
1.2. A Função e a Superfície de Resposta. Representação Gráfica da Superfície	01
1.3. Porque usar Superfície de Resposta	04
1.4. Aproximações para a Função de Resposta	06
2. MÉTODO DA MÁXIMA INCLINAÇÃO ASCENDENTE	09
2.1. Caracterização do Método	09
2.2. Exemplo	11
3. ANÁLISE DE MODELÔS QUADRÁTICOS	23
3.1. Introdução	23
3.2. Redução à Forma Canônica e Caracterização da Superfície de Resposta	25
3.3. Exemplo	28
3.4. Sistema de "Cumeeira" ou "Aresta"	35
4. DELINEAMENTOS DE SUPERFÍCIE DE RESPOSTA	42
4.1. Delineamentos para Ajustar Modelos de Primeira Ordem	42
4.2. Delineamentos para Ajustar Modelos de Segunda Ordem	44
5. UTILIZAÇÃO DO PROGRAMA RSREG DO PACOTE SAS PARA A METODOLOGIA DE SUPERFÍCIE DE RESPOSTA	49
5.1. Programa	49
5.2. Dados	50
5.3. Resultados	50

6. EXEMPLOS DE APLICAÇÃO DA METODOLOGIA DE SUPERFÍCIE DE RESPOSTA	59
6.1. Exemplo 1	59
6.2. Exemplo 2	64
6.3. Exemplo 3	69
6.4. Exemplo 4	75
7. BIBLIOGRAFIA	81
ANEXO I	82

METODOLOGIA DE SUPERFÍCIE DE RESPOSTA: UMA ABORDAGEM INTRODUTÓRIA

1. INTRODUÇÃO:

1.1. Características Gerais:

Metodologia de Superfície de Resposta (MSR), ou RSM (Response Surface Methodology) é um conjunto de técnicas matemáticas e estatísticas usadas para analisar problemas em que várias variáveis independentes (x_1, x_2, \dots, x_k) influenciam uma variável dependente ou resposta (y) e o objetivo é otimizar esta resposta. Assume-se que as variáveis independentes são contínuas e controladas pelo experimentador com erro desprezível e que a resposta é uma variável aleatória.

Por exemplo: em Engenharia, a resposta pode ser resistência de um material e os fatores quantitativos que a afetam podem ser temperatura, pressão, etc; em Química, a resposta pode ser produção, pureza ou custo de um composto químico e os fatores, pressão, temperatura, etc; em Agronomia a resposta pode ser a produção de uma cultura e os fatores as quantidades de nitrogênio, fósforo e potássio que compõem o adubo adicionado; em Psicologia, a resposta pode ser o resultado de um teste e os fatores, tempo de estudo, Q.I., etc; na Ciência dos Alimentos a resposta pode ser o sabor de um produto e os fatores as concentrações de açúcar e sal, etc.

1.2. A Função e a Superfície de Resposta. Representação Gráfica da Superfície.

Suponhamos que um engenheiro químico deseja obter a temperatura (x_1) e pressão (x_2) que maximiza a produção de um processo. Pode-se escrever a resposta observada y como a função dos níveis de temperatura e pressão como

$$y = f(x_1, x_2) + \epsilon$$

(1.1)

onde ϵ é um erro aleatório. Se assumimos a resposta esperada por $E(y) = \eta$, então a superfície representada por $\eta = f(x_1, x_2)$ é chamada de superfície de resposta. Na prática, por causa de erros experimentais diversos, não observamos o verdadeiro valor da resposta η e sim o valor de uma variável aleatória y , tal que

$$y = \eta + \epsilon$$

(1.2)

Para o exemplo a superfície de resposta tem a representação gráfica apresentada na figura 1.1, onde x_1 e x_2 aparecem no plano e visualiza-se o eixo η perpendicular ao plano.

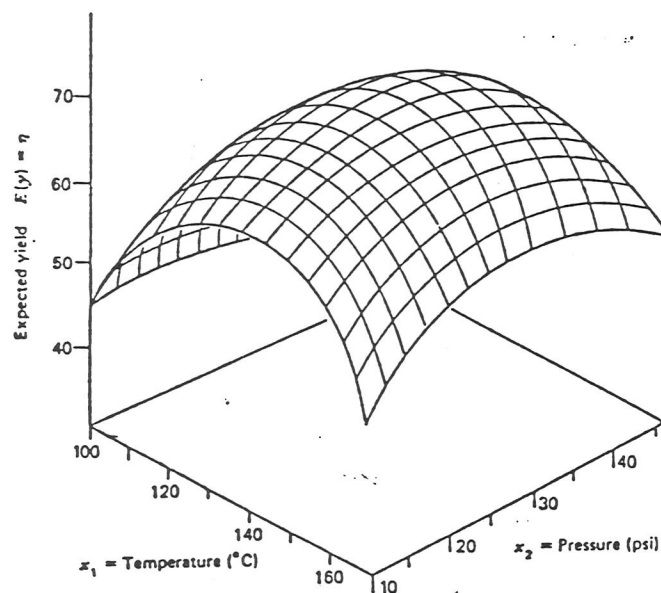


FIGURA 1.1 - SUPERFÍCIE DE RESPOSTA PARA DOIS FATORES
(Extraída de MONTGOMERY, 1991)

Uma alternativa muito útil para a representação geométrica da Superfície de Resposta é o gráfico de contornos de resposta esperada constante. Ele é obtido desenhando-se linhas de igual resposta de um gráfico cujas coordenadas representam os níveis

dos fatores. Na figura 1.2 o gráfico de contornos para a superfície representada na figura 1.1 .

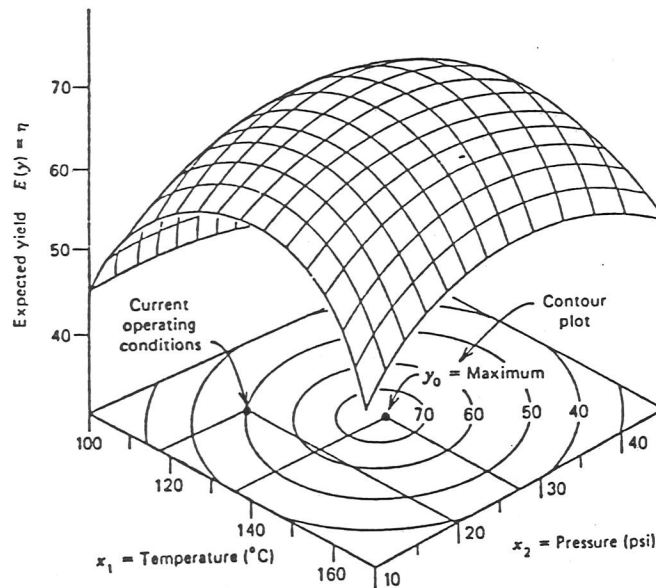


FIGURA 1.2 - GRÁFICO DE CONTORNOS DE RESPOSTA PARA A SUPERFÍCIE DA FIGURA 1.1 (Extraída de MONTGOMERY, 1991).

No caso de três fatores o gráfico de contornos da superfície, no espaço tridimensional, aparece na figura 1.3 .

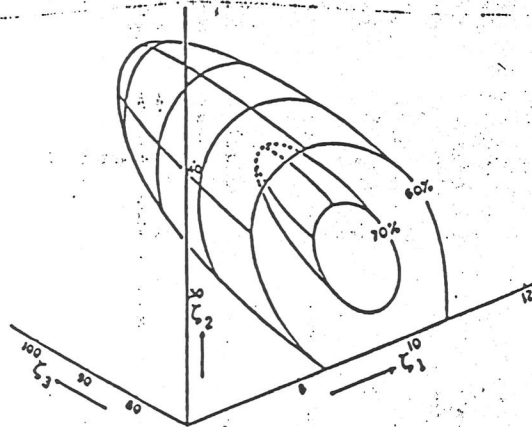


FIGURA 1.3 - GRÁFICO DE CONTORNOS NO ESPAÇO TRIDIMENSIONAL (Extraída de DAVIES, 1956).

Para três ou mais fatores é mais comum a representação gráfica da superfície através do gráfico de contornos no espaço bidimensional, fixando um ou mais fatores no nível ótimo.

1.3. Porque usar Superfície de Resposta.

Suponhamos que um engenheiro químico deseja maximizar a produção de uma reação química, variando o tempo de reação (t) e a temperatura da reação (T).

Se o engenheiro empregar o procedimento clássico de uma variável independente (fator) de cada vez ele deve seguir o que está ilustrado na figura 1.4. Na figura 1.4.a, a temperatura foi fixada em $T=225^{\circ}\text{C}$ e o tempo de reação variou de 60 a 180 minutos. A simples inspeção do gráfico mostra que o melhor tempo de reação, para esta temperatura, é 130 minutos, na qual a produção é 75 gramas.

Seguindo o procedimento clássico, o pesquisador fixará o tempo de reação em 130 minutos e variará a temperatura, como na figura 1.4(b), chegando a conclusão de que a melhor temperatura está ao redor de 225°C , usada no 1º experimento, e que a produção máxima está ao redor de 75 gramas.

Assim, a conclusão que parece válida é de que o máximo de produção está ao redor de 75 g é alcançado com o tempo de reação em 130 minutos e a temperatura em 225°C . O que os experimentos não mostram é o que acontece quando os tempos de reação e as temperaturas mudam, não isoladamente, mas conjuntamente.

Para entender a possível natureza do efeito conjunto do tempo de reação e da temperatura sobre a produção é necessário conduzir um experimento variando ambos os fatores simultaneamente. Os resultados expressando essa dependência podem ser apresentados em um gráfico de contornos (figura 1.5), perfeitamente concordantes com os resultados da figura 1.4.

Pela figura 1.5 nota-se que a máxima produção é de 91 g, e não 75 g como anteriormente, e que esse máximo é alcançado com um tempo de reação de 65 minutos e uma temperatura de 225°C .

Estudos das figuras 1.4 e 1.5 mostram que o procedimento de

uma variável independente (fator) de cada vez falha, pois ele assume que o máximo valor produzido por um fator é independente do máximo produzido pelo outro fator, o que nem sempre acontece.

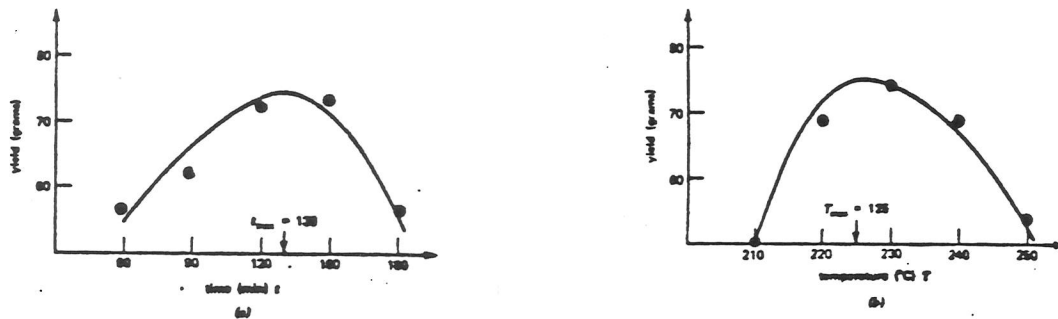


Figura 1.4 - Resultados hipotéticos do procedimento de uma variável independente (fator) de cada vez.

(a) Produção x tempo de reação, à temperatura de 225°C .

(b) Produção x temperatura, ao tempo de reação de 120 minutos (extraída de Box et al., 1978).

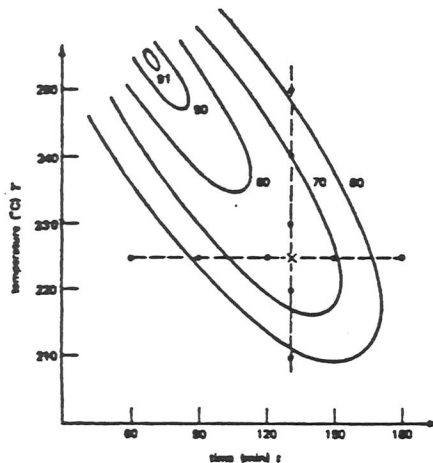


Figura 1.5 - Possível superfície de resposta, representando a produção em função do tempo de reação e da temperatura, com os pontos mostrados no procedimento de uma variável independente (fator) de cada vez (extraída de Box et al., 1978).

1.4. Aproximações para a Função de Resposta.

Em geral, em muitos problemas de MSR, a forma da verdadeira relação entre a resposta e as variáveis independentes não é conhecida, e pode ser complicada. Assim o primeiro passo em MSR é encontrar uma aproximação adequada para a relação funcional entre y e o conjunto de variáveis independentes, dentro de uma faixa limitada do "espaço dos fatores". Usualmente, um polinômio de baixa ordem é empregado. Se a resposta é bem modelada por uma função linear das variáveis independentes, então a função de aproximação é o modelo de 1ª ordem:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (1.3)$$

Se existe curvatura no sistema, então um polinômio de maior grau, tal como o modelo de 2ª ordem

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{\substack{i,j \\ i < j}} \beta_{ij} x_i x_j + \varepsilon \quad (1.4)$$

deve ser usado. Praticamente todos os problemas de MSR utilizam uma ou ambas aproximações polinomiais. De fato, é improvável que um modelo polinomial será uma aproximação razoável da relação funcional verdadeira no espaço completo das variáveis independentes, mas em uma região pequena eles usualmente apresentam bons resultados.

1.5. A Estimação dos Coeficientes, o Planejamento e o

Procedimento da Metodologia de Superfície de Resposta.

O método dos mínimos quadrados é usado para estimar os parâmetros nas aproximações polinomiais. Expressando os modelos polinomiais matricialmente, tem-se

$$y = X\beta + \varepsilon \quad (1.5)$$

e o sistema de equações normais é dado por

$$\boxed{x'x\hat{\beta} = x'y} \quad (1.6)$$

cuja solução é

$$\boxed{\hat{\beta} = (x'x)^{-1} x'y} \quad (1.7)$$

e a matriz de variâncias e covariâncias é

$$\boxed{V = (x'x)^{-1} \sigma^2} \quad (1.8)$$

A análise da superfície de resposta é então feita em termos da superfície ajustada. Se a superfície ajustada é uma adequada aproximação da verdadeira função de resposta, então a análise da superfície ajustada será aproximadamente equivalente à análise do sistema real. Os parâmetros dos modelos podem ser estimados mais efetivamente se delineamentos experimentais apropriados são usados para coletar os dados. Delineamentos bastante utilizados são os Fatoriais das séries 2^k e 3^k e Fatoriais Fracionários destas séries. Delineamentos foram desenvolvidos por BOX e WILSON (1951) especificamente para gerar dados para ajustar superfícies de respostas, conhecidos como Delineamentos Compostos.

A MSR é um procedimento seqüencial. Muitas vezes, quando estamos afastados do ponto onde a superfície de resposta conduz a um ótimo, existe pequena curvatura no sistema, um modelo de 1ª ordem será apropriado. O objetivo nesse ponto é levar o experimento rápida e eficientemente para a vizinhança do ótimo. Uma vez que a região do ótimo tenha sido encontrada, um modelo mais elaborado, como uma superfície de segunda ordem, pode ser empregado, e uma análise realizada para localizar o ótimo. A análise de uma superfície de resposta pode ser pensada como

"subindo uma montanha", onde o topo da montanha representa o ponto de máxima resposta. Se o ótimo verdadeiro é um mínimo, então pode-se pensar como descendo a um vale.

O objetivo principal da MSR é determinar as condições de operação ótimas do sistema ou determinar a região do espaço dos fatores no qual as especificações de operação são satisfeitas. MSR não é usada prioritariamente para obter conhecimento do mecanismo físico do sistema, embora possa auxiliar na busca de tal conhecimento. Além disso, "ótimo" em MSR é usado em um sentido especial. O procedimento de "subida da montanha" da MSR garante convergência somente para um ótimo local.

2. MÉTODO DA MÁXIMA INCLINAÇÃO ASCENDENTE

2.1. Caracterização do Método.

Freqüentemente, as condições iniciais estão afastadas daquelas que otimizam a resposta. Em tais situações, o objetivo é mover o experimento rapidamente para a vizinhança geral do ótimo utilizando um procedimento experimental simples, rápido econômico e eficiente. Quando se está distante do ótimo, usualmente assume-se que um modelo de primeira ordem é uma aproximação adequada para a verdadeira superfície em uma pequena região das variáveis independentes.

O Método da Máxima Inclinação Ascendente (MMIA) é um procedimento usado para mover o experimento seqüencialmente ao longo do caminho de máxima inclinação ascendente, isto é, à direção de máximo incremento na resposta. Se a minimização é desejada então diz-se Método de Máxima Inclinação Descendente.

O modelo de primeira ordem ajustado é:

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^k \hat{\beta}_i X_i \quad (2.1)$$

e gráfico de contornos da superfície de resposta de primeira ordem são uma série de linhas paralelas como mostrado na figura 2.1 . A direção de máxima inclinação ascendente é a direção na qual \hat{y} cresce mais rapidamente e é paralela à normal para a superfície de resposta ajustada. Usualmente, toma-se como caminho de máxima inclinação ascendente a linha a partir do centro da região de interesse e normal à superfície ajustada. Portanto, os passos ao longo do caminho são proporcionais aos coeficientes de regressão $\hat{\beta}_i$. O tamanho do passo é determinado pela experiência do pesquisador.

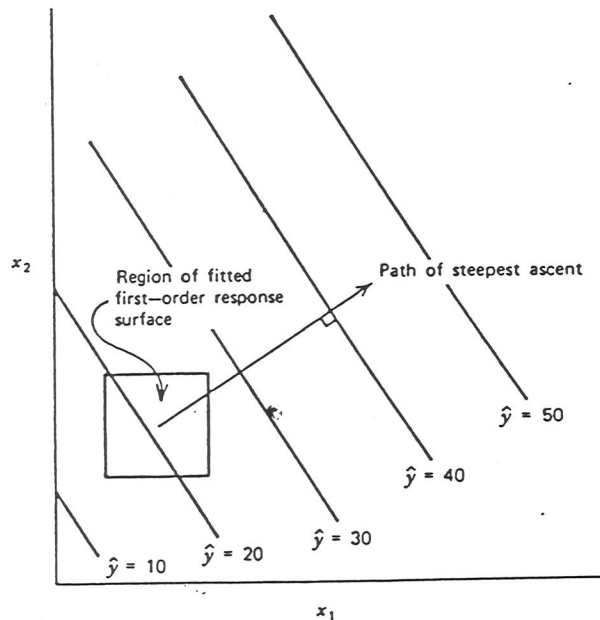


Figura 2.1 - Superfície de resposta de primeira ordem e caminho da máxima inclinação ascendente (extrída de MONTGOMERY, 1991).

Utiliza-se um conjunto de ensaios (tratamentos) em torno do ponto inicial e estima-se por mínimos quadrados as inclinações β_i . A partir das magnitudes e sinais destas inclinações, calcula-se a direção de máxima inclinação ascendente.

Experimentos são conduzidos ao longo do caminho de máxima inclinação ascendente até que nenhum incremento na resposta seja observado. Então, um novo modelo de primeira ordem pode ser ajustado, um novo caminho de máxima inclinação ascendente determinado e o procedimento continuado. Eventualmente, chega-se na vizinhança do ótimo. Isto será indicado pela falta de ajuste do modelo de primeira ordem. A aproximação por um plano se torna insatisfatória pelo fato de coeficientes (efeitos) de ordens mais altas, particularmente os de segunda ordem (quadráticos e de interação linear), se tornarem relativamente mais importantes. Então, o Método de Análise de Modelos Quadráticos será conduzido para obter uma estimativa mais precisa do ótimo.

2.2. Exemplo.

Um engenheiro químico está interessado em determinar os níveis de tempo e temperatura de reação que maximizam a produção de um processo. Normalmente, opera-se com um tempo de 35 minutos e uma temperatura de 155°F, que resulta numa produção de 40% aproximadamente. Como a região que contém o ótimo é desconhecida, ajusta-se um modelo de primeira ordem e aplica-se o Método da Máxima Inclinação Ascendente.

O engenheiro decide que a região experimental será (30,40) minutos para tempo de reação e (150,160)°F para temperatura. Para simplificar os cálculos, as variáveis independentes serão codificadas como (-1,1). Assim, se ξ_1 representa a variável natural para tempo e ξ_2 , a variável natural para temperatura, então, os valores codificados são

$$x_1 = \frac{\xi_1 - 35}{5}$$

$$x_2 = \frac{\xi_2 - 155}{5}$$

Os dados são apresentados na Tabela 2.1. O delineamento usado para coletar estes dados é um Fatorial 2^2 aumentado por cinco pontos (tratamentos) centrais. Os tratamentos repetidos ao centro são usados para estimar o erro experimental e possibilitam testar a adequabilidade ("lack of fit") do modelo de primeira ordem. Além disso, no presente caso, o tratamento central representa as condições de operação normalmente empregadas.

Tabela 2.1 - Dados para ajustar o Modelo de Primeira ordem.

Variáveis Naturais		Variáveis Codificadas		Resposta
ξ_1	ξ_2	x_1	x_2	y
30	150	-1	-1	39,3
30	160	-1	1	40,0
40	150	1	-1	40,9
40	160	1	1	41,5
35	155	0	0	40,3
35	155	0	0	40,5
35	155	0	0	40,7
35	155	0	0	40,2
35	155	0	0	40,6

Ajusta-se aos dados um modelo de primeira ordem, ou seja, obtém-se, pelo método de mínimos quadrados a equação de regressão

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 \quad (2.2)$$

Expressando matricialmente o modelo tem-se:

$$y = \text{vetor das observações} = \begin{bmatrix} 39,3 \\ 40,0 \\ 40,9 \\ 41,5 \\ 40,3 \\ 40,5 \\ 40,7 \\ 40,2 \\ 40,6 \end{bmatrix}$$

$$\begin{array}{l}
 x = \text{matriz dos coeficientes} = \\
 \text{dos parâmetros} \\
 \text{do modelo}
 \end{array}
 \begin{array}{c}
 x_0 \quad x_1 \quad x_2 \\
 \begin{bmatrix}
 1 & -1 & -1 \\
 1 & -1 & 1 \\
 1 & 1 & -1 \\
 1 & 1 & 1 \\
 1 & 0 & 0 \\
 1 & 0 & 0 \\
 1 & 0 & 0 \\
 1 & 0 & 0 \\
 1 & 0 & 0
 \end{bmatrix}
 \end{array}$$

$$\begin{array}{l}
 \beta = \text{vetor dos parâmetros} = \\
 \text{do modelo}
 \end{array}
 \begin{array}{c}
 \beta_0 \\
 \beta_1 \\
 \beta_2
 \end{array}$$

O sistema de equações normais $x'x\hat{\beta} = x'y$ é dado por:

$$\begin{bmatrix}
 9 & & \\
 & 4 & \\
 & & 4
 \end{bmatrix}
 \begin{bmatrix}
 \hat{\beta}_0 \\
 \hat{\beta}_1 \\
 \hat{\beta}_2
 \end{bmatrix}
 =
 \begin{bmatrix}
 364 \\
 3,1 \\
 1,3
 \end{bmatrix}$$

Cuja solução dada por $\hat{\beta} = (x'x)^{-1} x'y$ é

$$\hat{\beta} =
 \begin{bmatrix}
 \hat{\beta}_0 \\
 \hat{\beta}_1 \\
 \hat{\beta}_2
 \end{bmatrix}
 =
 \begin{bmatrix}
 1/9 & & \\
 & 1/4 & \\
 & & 1/4
 \end{bmatrix}
 \begin{bmatrix}
 364 \\
 3,1 \\
 1,3
 \end{bmatrix}
 =
 \begin{bmatrix}
 40,44 \\
 0,775 \\
 0,325
 \end{bmatrix}$$

Assim a equação de Regressão encontrada é

$$\hat{y} = 40,44 + 0,775x_1 + 0,325x_2 \quad (2.3)$$

A variabilidade total dada por

$$SQ_{Total} = (39,3)^2 + \dots + (40,6)^2 - \left[\frac{364}{10} \right]^2 = 3,0022 \quad e$$

particionada na SQRegressão e SQResíduo, ou seja:

$$SQ_{Total} = SQ_{Regressão} + SQ_{Resíduo} \quad , \text{ onde}$$

$$SQ_{Regressão} = (0,775)(3,1) + (0,325)(1,3) = 2,8250 \quad e$$

$$SQ_{Resíduo} = SQ_{Total} - SQ_{Regressão} = 3,0022 - 2,8250 = 0,1772$$

A estimativa do erro experimental com 4 GL é dada por

$$SQ_{Erro Experimental} = SQ_{Erro Puro}$$

$$\begin{aligned} &= (40,3)^2 + (40,5)^2 + (40,7)^2 + (40,2)^2 + (40,66)^2 - (202,3)^2/5 \\ &= 0,1720 \end{aligned}$$

$$\text{Quadrado Médio do Erro Experimental} = QME = \frac{0,1720}{4}$$

O modelo de primeira ordem assume que as variáveis x_1 e x_2 possuem um efeito aditivo sobre a resposta. A interação entre as variáveis é medida pelo coeficiente β_{12} do termo x_1x_2 adicionado ao modelo. A estimativa de mínimos quadrados deste coeficiente é:

$$\begin{aligned} \hat{\beta}_{12} &= 1/4 [(1 \times 39,3) + (-1 \times 40,0) + (-1 \times 40,9) + (1 \times 41,55)] \\ &= 1/4 (-0,1) = -0,025 \end{aligned}$$

A Soma de Quadrados da interação com 1 grau de liberdade é:

$$SQ_{\text{Interação}} = \frac{(-0,1)^2}{4} = 0,0025$$

A estatística de Falta de Ajuste (lack-of-fit) é:

$$F = \frac{SQ_{\text{Interação}}}{QME} = \frac{0,0025}{0,0430} = 0,058$$

a qual é comparada com $F_{.05(1,4)} = 7,71$. Portanto, a interação não é significativa.

Outra verificação da adequabilidade do modelo é obtida pela comparação da resposta média dos quatro pontos do fatorial 2^2 , $\bar{y}_1 = 40,425$, com a resposta média do centro do delineamento, $\bar{y}_2 = 40,40$. Se o delineamento está situado sobre uma superfície curva, então, $\bar{y}_1 - \bar{y}_2$ é uma medida de curvatura geral da superfície. Se β_{11} e β_{22} são os coeficientes dos termos quadráticos puros x_1^2 e x_2^2 , então, $\bar{y}_1 - \bar{y}_2$ é uma estimativa de $\beta_{11} + \beta_{22}$:

$$\hat{\beta}_{11} + \hat{\beta}_{22} = \bar{y}_1 - \bar{y}_2 = 40,425 - 40,46 = -0,035$$

A Soma de Quadrados associada ao contraste $\bar{y}_1 - \bar{y}_2$ com 1 GL é

$$\begin{aligned} SQ_{\text{Quadrático Puro}} &= \frac{n_1 n_2 (\bar{y}_1 - \bar{y}_2)^2}{n_1 + n_2} && (2.4) \\ &= \frac{(4)(5)(-0,035)^2}{4 + 5} \\ &= 0,0027 \end{aligned}$$

onde n_1 e n_2 são o número de pontos (tratamentos) na parte fatorial e o número de pontos (tratamentos) centrais, respectivamente.

A estatística

$$F = \frac{SQ_{\text{Quadrático Puro}}}{QME} = \frac{0,0027}{0,0430} = 0,0663$$

comparada com $F_{05(1,4)}$, indica que o efeito quadrático puro, não é significativo.

A análise de variância completa é apresentada na Tabela 2.2:

Tabela 2.2. Análise de Variância

Causas de Variação	Soma de Quadrados	Graus de Liberdade	Quadrado Médio	F _o
Regressão(β_1, β_2)	2,8250	2	1,4125	47,83 ^a
Resíduo	0,1772	6	0,0295	
(Interação)	(0,0025)	1	0,0025	0,058
(Quadr. Puro)	(0,0027)	1	0,0027	0,063
(Erro Puro)	(0,1720)	4	0,0430	
Total	3,0022	8		

^aSignificante a 1%

Nota-se que ambas verificações de curvatura não são significantes, enquanto que o teste F para a regressão é significativo ao nível de 1% .

Não existe, portanto, nenhuma razão para questionar a adequabilidade do modelo de primeira ordem.

Os passos do caminho de máxima inclinação ascendente aparecem calculados na tabela 2.3 .

Tabela 2.3. Cálculo do caminho de máxima inclinação ascendente e subsequentes ensaios.

	Variáveis Codificadas		Variáveis Naturais		Resposta
	x_1	x_2	ξ_1	ξ_2	y
Nível Base	0	0	35	155	
Unidade			5	5	
Unidade $\times \hat{\beta}_1$			3,875	1,625	
Variação do Nível	1,00	0,42	5	2	
Ensaio 1	1,00	0,42	40	157	41,0
Ensaio 2	2,00	0,84	45	159	42,9
Ensaio 3	3,00	1,26	50	161	47,1
Ensaio 4	4,00	1,68	55	163	49,7
Ensaio 5	5,00	2,10	60	165	53,8
Ensaio 6	6,00	2,52	65	167	59,9
Ensaio 7	7,00	2,94	70	169	65,0
Ensaio 8	8,00	3,36	75	171	70,4
Ensaio 9	9,00	3,78	80	173	77,6
Ensaio 10	10,00	4,20	85	175	80,3
Ensaio 11	11,00	4,62	90	179	76,2
Ensaio 12	12,00	5,04	95	181	75,1

Os fatores devem variar na proporção das estimativas dos $\hat{\beta}_i$, ou seja, para cada 0,775 unidades acrescentadas em x_1 , x_2 deve aumentar 0,325 unidades. A linha 4 define a direção de máxima inclinação ascendente; para cada $5 \times 0,775 = 3,875$ minutos que ξ_1 é mudado, ξ_2 deve ser mudado em $5 \times 0,325 = 1,625$ F. O caminho principal a partir da origem ($x_1=0$, $x_2=0$) nesta direção pode ser obtido dando um incremento conveniente a um dos fatores (o engenheiro decide usar 5 minutos para tempo, o que equivale a $\Delta x_1=1$ para x_1) e calculando as mudanças proporcionais do outro fator: $5 \times (1,625) / 3,875 = 2$ F (o que equivale a $\Delta x_2 = (0,325 / 0,775) \Delta x_1 = 0,42$). Estas quantidades devem ser sucessivamente acrescentadas aos níveis base dando origem ao caminho de máxima inclinação ascendente. Então, o engenheiro realizou os ensaios de acordo

com este caminho e observou a produção nestes pontos até que um decréscimo na resposta foi notado. A representação gráfica do Caminho de Máxima Inclinação Ascendente aparece na Figura 2.2 .

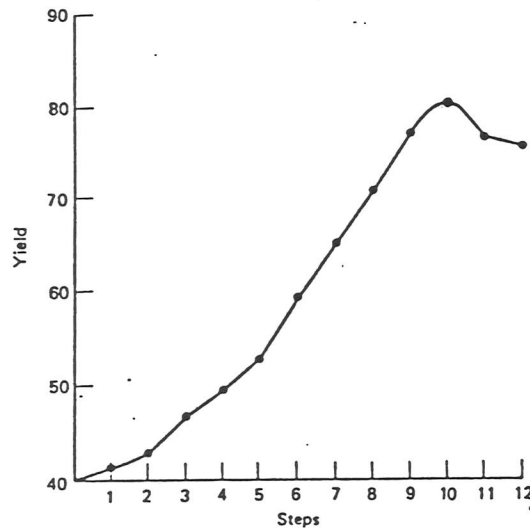


Figura 2.2 - Produção em função dos passos no caminho de Máxima Inclinação Ascendente (Extraída de MONTGOMERY, 1991).

Incrementos na resposta são observados até o décimo passo; entretanto, a partir do décimo primeiro passo há um decréscimo na produção. Portanto, outro modelo de primeira ordem pode ser ajustado em torno do ponto $(x_1 = 85, x_2 = 175)$. A região de exploração para x_1 é $[80, 90]$ e para x_2 é $[170, 180]$. Novamente se codifica os níveis das duas variáveis como $(-1, 1)$, e um delineamento fatorial 2^2 acrescido de cinco pontos centrais é usado. Os dados estão na Tabela 2.4 .

Tabela 2.4 . Dados para ajustar o segundo Modelo de Primeira Ordem.

Variáveis Naturais		Variáveis Codificadas		Resposta
ξ_1	ξ_2	x_1	x_2	y
80	170	-1	-1	76,5
80	180	-1	1	77,0
90	170	1	-1	78,0
90	180	1	1	79,5
85	175	0	0	79,9
85	175	0	0	80,3
85	175	0	0	80,0
85	175	0	0	79,7
85	175	0	0	79,8

O modelo de primeira ordem ajustado aos dados codificados é:

$$\hat{y} = 78,97 + 1,00x_1 + 0,50x_2 \quad (2.5)$$

A Tabela 2.5 apresenta a análise de variância para este modelo.

Tabela 2.5 . Análise de Variância.

Causas de Variação	Soma de Quadrados	Graus de Liberdade	Quadrado Médio	F_o
Regressão(β_1, β_2)	5,0000	2		
Resíduo	11,1200	6		
(Interação)	(0,2500)	1	0,2500	4,72 ^a
(Quadr. Puro)	(10,6580)	1	10,6580	201,09 ^b
(Erro Puro)	(0,2120)	4	0,0530	
Total	16,1200	8		

^aSignificante a 10%

^bSignificante a 1%

Os componentes da interação e do termo quadrático puro são significantes, implicando que o modelo de primeira ordem não é

uma aproximação adequada. Esta curvatura na verdadeira superfície pode indicar que se está próximo do ótimo. Neste ponto, análises adicionais podem ser feitas para localizar o ótimo mais precisamente.

O Caminho de Máxima Inclinação Ascendente é proporcional ao sinal e magnitude dos coeficientes de regressão do modelo de primeira ordem ajustado, expresso em (2.2).

É fácil estabelecer um algoritmo geral para determinar as coordenadas de um ponto no caminho de Máxima Inclinação Ascendente. Assume-se que $x_1=x_2=\dots=x_k=0$ é a base ou ponto de origem. Então

1. Escolhe-se uma amplitude de variação (tamanho do passo) para uma das variáveis independentes (fatores) Δx_j . Usualmente seleciona-se a variável mais conhecida, ou seleciona-se a variável que tem o maior valor absoluto para o coeficiente de regressão, isto é, o maior $|\hat{\beta}_j|$.

2. O tamanho do passo para as outras variáveis é:

$$\Delta x_i = \frac{\hat{\beta}_i}{\hat{\beta}_j / \Delta x_j} \quad i = 1, 2, \dots, k ; i \neq j$$

3. Converte-se os Δx_i das variáveis codificadas para as variáveis originais.

Considerando o exemplo em questão:

1. Desde que x_1 é a variável que tem o maior valor absoluto para o coeficiente de regressão, escolhe-se o tempo de reação como a variável do passo 1 do algoritmo. Cinco minutos de tempo de reação é o tamanho do passo, baseado em conhecimento do processo. Em termos de variável codificada $\Delta x_1 = 1$.

2. O tamanho do passo para temperatura é

$$\Delta x_2 = \frac{\hat{\beta}_2}{\hat{\beta}_1 / \Delta x_1} = \frac{0,325}{(0,775/1)} = 0,42$$

3. Para converter para as variáveis originais usa-se as

relações

$$\Delta x_1 = \frac{\Delta \xi_1}{5} \quad \text{e} \quad \Delta x_2 = \frac{\Delta \xi_2}{5}$$

que resulta em

$$\Delta \xi_1 = \Delta x_1(5) = 1(5) = 5 \text{ minutos e}$$

$$\Delta \xi_2 = \Delta x_2(5) = 0,42(5) = 2^\circ \text{ F} .$$

3. ANÁLISE DE MODELOS QUADRÁTICOS.

3.1. Introdução.

Quando se está numa sub-região do espaço dos fatores onde os efeitos de primeira ordem são pequenos comparados com os de ordem superior, assume-se que o experimento está próximo do ótimo. O objetivo é estudar a natureza da Superfície de Resposta, aproximada por um polinômio de grau dois ou mais, devido a curvatura na superfície, e encontrar a combinação dos níveis dos fatores que otimizam a resposta.

Na maioria dos casos, o modelo de segunda ordem

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^k \hat{\beta}_i x_i + \sum_{i=1}^k \hat{\beta}_{ii} x_i^2 + \sum_{i < j} \sum_j \hat{\beta}_{ij} x_i x_j \quad (3.1)$$

é uma aproximação adequada.

3.2. Localização do Ponto Estacionário

Supomos que se deseja determinar os níveis de x_1, x_2, \dots, x_k que otimizam a resposta estimada. Este ponto, se existir, será o conjunto de x_1, x_2, \dots, x_k para os quais as derivadas parciais

$$\frac{\partial \hat{y}}{\partial x_1} = \frac{\partial \hat{y}}{\partial x_2} = \dots = \frac{\partial \hat{y}}{\partial x_k} = 0$$

Este ponto designado $x_{1,0}, x_{2,0}, \dots, x_{k,0}$ é chamado de ponto estacionário.

O ponto estacionário de resposta representa um ponto de máxima resposta, um ponto de mínima resposta ou um ponto de sela ou minimax.

Pode-se obter uma solução geral para o ponto estacionário. Escrevendo-se o modelo quadrático (3.1) na forma matricial

tem-se:

$$\hat{y} = \hat{\beta}_0 + x'b + x'Bx \quad (3.2)$$

onde

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} \quad b = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} \quad e \quad B = \begin{bmatrix} \hat{\beta}_{11} & \hat{\beta}_{12}/2 & \dots & \hat{\beta}_{1k}/2 \\ & \hat{\beta}_{22} & \dots & \hat{\beta}_{2k}/2 \\ & & \ddots & \vdots \\ \text{sim.} & & & \hat{\beta}_{kk} \end{bmatrix}$$

b ($k \times 1$) é um vetor dos coeficientes de regressão de primeira ordem e B ($k \times k$) é uma matriz simétrica cujos elementos da diagonal principal são os coeficientes quadráticos puros $\hat{\beta}_{ii}$ e fora da diagonal são a metade dos coeficientes quadráticos mistos $\hat{\beta}_{ij}$, $i \neq j$.

A combinação dos níveis dos fatores x_1, x_2, \dots, x_k que maximizam a resposta estimada \hat{y} é obtida fazendo-se

$$\frac{\partial \hat{y}}{\partial x} = b + 2Bx = 0 \quad (3.3)$$

O ponto estacionário é a solução da Equação 3.3:

$$x_0 = -\frac{1}{2} B^{-1} b \quad (3.4)$$

Substituindo-se a Equação 3.4 na Equação 3.2, acha-se a resposta estimada no ponto estacionário

$$\hat{y}_0 = \hat{\beta}_0 + \frac{1}{2} x_0' b \quad (3.5)$$

ou

$$\hat{y}_0 = \hat{\beta}_0 + \frac{1}{2} (\hat{\beta}_1 x_{1,0} + \hat{\beta}_2 x_{2,0} + \dots + \hat{\beta}_k x_{k,0}) \quad (3.6)$$

3.2. Redução à Forma Canônica e Caracterização da Superfície de Resposta.

Além de determinar o ponto estacionário há necessidade de caracterizar a superfície de resposta na vizinhança deste ponto. Para caracterizar a superfície necessita-se determinar se o ponto estacionário é um ponto de máximo, ou mínimo ou de sela para as variáveis x_1, x_2, \dots, x_k .

O Caminho mais imediato para determinar a natureza do ponto estacionário é examinar os contornos de resposta para o modelo ajustado. Quando se tiver um pequeno número de variáveis independentes ou fatores (duas ou três) este procedimento é relativamente fácil. Como nem sempre se tem um pequeno número de variáveis independentes, necessita-se de um procedimento mais formal. O procedimento mais formal consiste em reduzir o modelo (3.1) a sua forma canônica.

O modelo na forma canônica torna possível apreciar a natureza da superfície ajustada, mesmo quando muitos fatores estão envolvidos e construir as linhas de contorno em duas ou três dimensões.

É útil transformar o modelo ajustado em um novo sistema de coordenadas, com a origem no ponto estacionário $x_0 = (x_{1,0}, x_{2,0}, \dots, x_{k,0})$ e, então, fazer uma rotação dos eixos deste sistema até que sejam paralelos ao eixo principal da superfície de resposta ajustada. Esta transformação é mostrada na figura 3.1. O modelo na Forma Canônica é:

$$\hat{y}_0 = \hat{y}_0 + \lambda_1 w_1^2 + \lambda_2 w_2^2 + \dots + \lambda_k w_k^2 \quad (3.7)$$

onde \hat{y}_0 representa a resposta estimada no ponto estacionário x_0 .

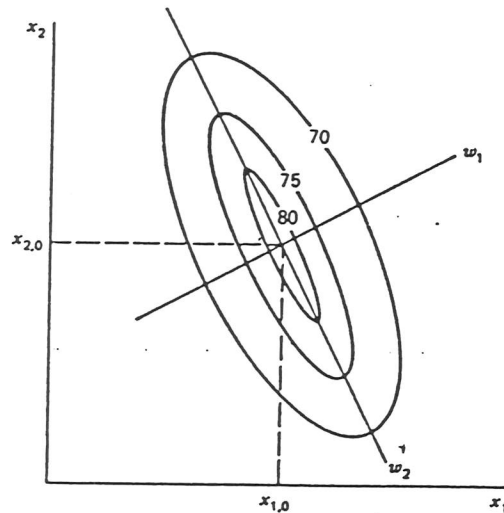


Figura 3.1 - Forma Canônica do Modelo de Segunda Ordem.

(Extraída de MONTGOMERY, 1994).

O procedimento para obtenção da Forma Canônica é calcular:

i) os valores dos coeficientes $\lambda_1, \lambda_2, \dots, \lambda_k$:

Determinar os autovalores ou raízes características λ_i da matriz dos coeficientes de segunda ordem B.

ii) as direções dos eixos w_1, w_2, \dots, w_k :

As variáveis independentes x são relacionadas com as variáveis canônicas w por

$$w = M'(x - x_0) \quad (3.8)$$

onde M ($k \times k$) é uma matriz ortogonal cujas colunas são vetores característicos (autovetores) normalizados associados a cada λ_i .

A natureza da Superfície de Resposta pode ser determinada pelo ponto estacionário x_0 e o sinal e magnitude dos λ_i , obtidos da Forma Canônica.

Supõe-se, inicialmente, que o ponto estacionário (x_0) está dentro da região de exploração para o modelo de segunda ordem ajustado. Se os λ_i , $i=1, 2, \dots, k$ são todos negativos, então x_0

é um ponto de máxima resposta; se os λ_i são todos positivos, então x_0 é um ponto de mínima resposta; se os λ_i tem sinais diferentes, então x_0 é um ponto de sela ou minimax. Além disso a superfície tem uma variação mais acentuada na direção w_i para o qual $|\lambda_i|$ é maior.

A figura 3.1 descreve um sistema no qual os coeficientes λ_1 e λ_2 são negativos. Então, x_0 é um ponto de máxima resposta, ou seja, é a combinação ótima de x_1 e x_2 . Existirá sempre uma perda na resposta, qualquer que seja a direção tomada a partir de x_0 ; λ_1 mede a diminuição na resposta na direção w_1 , e λ_2 na direção w_2 . Os contornos ajustados são elipses e, como $|\lambda_2|$ é menor do que $|\lambda_1|$, estes contornos serão prolongados ao longo do eixo w_2 , como na ilustração.

A figura 3.2 descreve um sistema no qual os coeficientes λ_1 e λ_2 são positivos. Então, x_0 é um ponto de mínima resposta, ou seja existirá sempre um acréscimo na resposta, qualquer que seja a direção tomada a partir de x_0 .

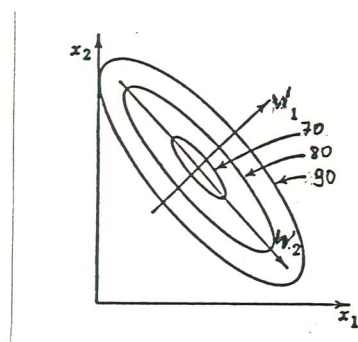


Figura 3.2 - Superfície de Resposta com Ponto de Mínimo (Extraída de DAVIES, 1956).

A figura 3.3 ilustra o sistema obtido quando λ_1 é negativo e λ_2 é positivo; os contornos ajustados são hipérbolas e x_0 é um

ponto de sela ou minimax (a resposta é um mínimo para uma das variáveis e um máximo para a outra). Quando λ_2 é menor em valor absoluto do que λ_1 , os contornos serão prolongados ao longo do eixo w_2 . A resposta aumenta mais rapidamente à medida que se incrementa o w_i que tem o mais alto $|\lambda_i|$. Na ilustração, a superfície tem máxima inclinação na direção w_1 , pois $|\lambda_1|$ é maior.

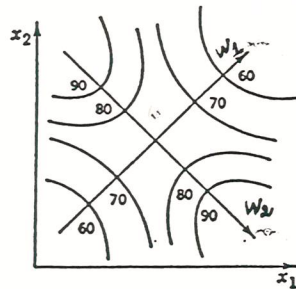


Figura 3.3 - Minimax ou ponto de sela
(extraída de DAVIES, 1950)..

3.3. Exemplo.

Considera-se a análise do processo químico utilizado no exemplo da seção 2.3. Um modelo de segunda ordem nas variáveis x_1 e x_2 não pode ser ajustado usando os dados da tabela 2.4. O experimentador decide aumentar esses dados com pontos suficientes para ajustar um modelo de segunda ordem. Ele obtém quatro observações para $(x_1=0, x_2=\pm 1,414)$ e $(x_1=\pm 1,414, x_2=0)$. O conjunto completo dos dados é mostrado na tabela 3.1, e o delineamento é mostrado na figura 3.4. Este é chamado Delineamento Central Composto Rotacional.

Tabela 3.1. Dados para ajustar o Modelo de Segunda Ordem.

Variáveis Naturais		Variáveis Codificadas		Resposta
ξ_1	ξ_2	x_1	x_2	y
80	170	-1	-1	76,5
80	180	-1	1	77,0
90	170	1	-1	78,0
90	180	1	1	79,5
85	175	0	0	79,9
85	175	0	0	80,3
85	175	0	0	80,0
85	175	0	0	79,7
85	175	0	0	79,8
92,07	175	1,414	0	78,4
77,93	175	-1,414	0	75,6
85	182,07	0	1,414	78,5
85	167,93	0	-1,414	77,0

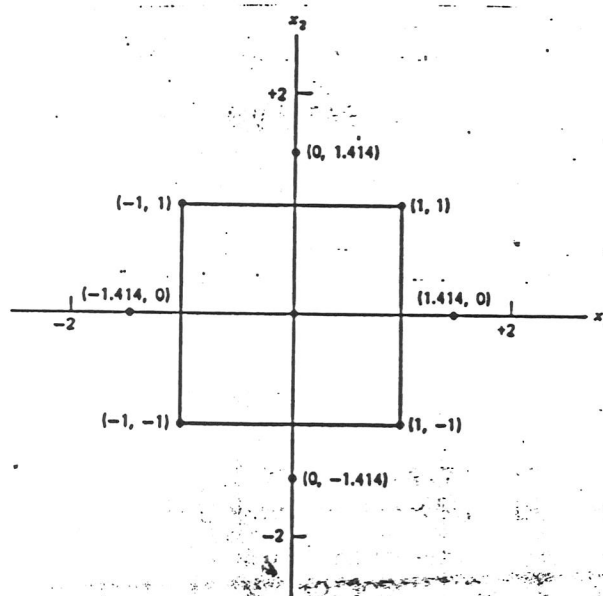


Figura 3.4 - Delineamento Central Composto Rotacional.

(Extraída de MONTGOMERY, 1991)

Um modelo de segunda ordem é ajustado aos dados codificados por mínimos quadrados, obtendo-se:

$$\hat{y} = 79,9408 + 0,9949x_1 + 0,5151x_2 - 1,377x_1^2 - 1,0018x_2^2 + 0,25x_1x_2$$

A análise de variância para este modelo é mostrada na tabela 3.2.

Tabela 3.2. Análise de Variância

Causas de Variação	Soma de Quadrados	Graus de Liberdade	Quadrado Médio	F _o
Regressão	28,2560	5	5,6512	81,20 ^a
Erro	0,4871	7	0,0696	
(Falta de Ajuste)	(0,2751)	3	0,0917	1,73
(Erro Puro)	(0,2120)	4	0,0530	
Total	28,7431	12		

^aSignificante a 1%

A falta de ajuste não é significativa, mas a regressão é; então conclue-se que o modelo de segunda ordem é uma aproximação adequada da verdadeira superfície.

Após esta verificação, determina-se o ponto estacionário. Como

$$b = \begin{bmatrix} 0,9949 \\ 0,5151 \end{bmatrix}$$

$$B = \begin{bmatrix} -1,3770 & 0,1250 \\ 0,1250 & -1,0018 \end{bmatrix}$$

o ponto estacionário x_o é:

$$x_o = -\frac{1}{2}B^{-1}b = -\frac{1}{2} \begin{bmatrix} -0,7345 & -0,0917 \\ -0,0917 & -1,0096 \end{bmatrix} \begin{bmatrix} 0,9949 \\ 0,5151 \end{bmatrix} = \begin{bmatrix} 0,3890 \\ 0,3056 \end{bmatrix}$$

isto é, $x_{1,0} = 0,3890$ e $x_{2,0} = 0,3056$. Em termos de variáveis naturais, o ponto estacionário é:

$$0,3890 = \frac{\xi_1 - 85}{5} \quad \text{e} \quad 0,3056 = \frac{\xi_2 - 175}{5}$$

donde obtém-se

$$\xi_1 = 86,9450 \approx 87 \quad \xi_2 = 176,5280 \approx 176,5$$

A resposta estimada no ponto estacionário é:

$$\hat{y} = \hat{\beta}_0 + \frac{1}{2} x'_0 b = 79,9408 + \frac{1}{2} [0,3890 \quad 0,3056] \begin{bmatrix} 0,9949 \\ 0,5151 \end{bmatrix} = 80,21$$

A representação gráfica da superfície aparece nas figuras 3.5 e 3.6.

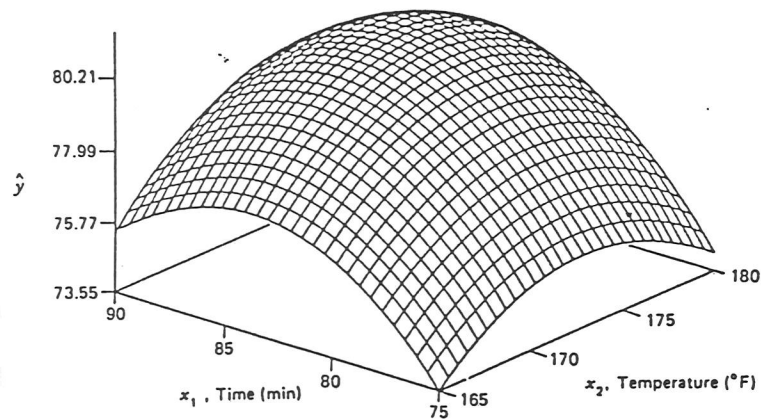


Figura 3.5 - Superfície de Resposta da Produção como função de tempo e temperatura (extraída de MONTGOMERY, 1991).

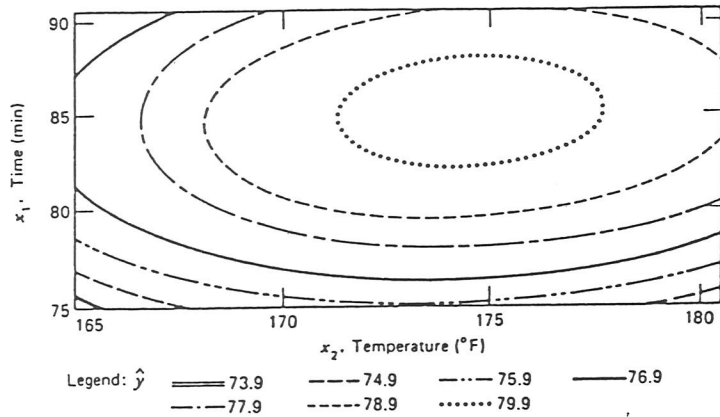


Figura 3.6 - Contornos de Resposta da Produção como função de tempo e temperatura (extraída de MONTGOMERY, 1991).

Examinando-se as figuras 3.5 e 3.6 percebe-se que a resposta ótima está próxima de 135^o F e de 85 minutos, como comprovamos pelos valores determinados para o ponto estacionário. Examinando o gráfico de contornos da figura 3.6, percebe-se que o processo é muito mais sensível a uma variação no tempo de reação do que na temperatura.

Para determinar a natureza da superfície pode-se proceder a redução a forma canônica obtendo-se os autovalores e os autovetores correspondentes.

Os autovalores λ_1 e λ_2 são raízes da equação característica:

$$|B - \lambda I| = 0$$

$$\begin{bmatrix} -1,3770-\lambda & 0,1250 \\ 0,1250 & -1,0018-\lambda \end{bmatrix} = 0$$

a qual reduz a

$$\lambda^2 + 2,3788\lambda + 1,3639 = 0$$

As raízes desta equação quadrática são $\lambda_1 = -0,9641$ e $\lambda_2 = -1,4147$. Assim, a Forma Canônica do modelo ajustado é

$$\hat{y} = 80,21 - 0,9641w_1^2 - 1,4147w_2^2$$

Como λ_1 e λ_2 são negativos e o ponto estacionário está dentro da região experimental, conclue-se que o ponto estacionário x_0 é um máximo. Os contornos são elipses. Como $|\lambda_1|$ é menor do que $|\lambda_2|$, estes contornos são prolongados ao longo do eixo w_1 .

Em alguns problemas de Metodologia de Superfície de Resposta pode ser necessário achar a relação entre as variáveis canônicas w_i e as variáveis independentes x_i . Isto ocorre, principalmente, se for impossível operar o processo no ponto estacionário. Suponha que não se pode operar o processo no ponto $\xi_1 = 87$ e $\xi_2 = 176,5$ porque esta combinação resulta num custo excessivo. Então, deseja-se sair do ponto estacionário para um ponto de baixo custo, sem grande perda na produção. A Forma Canônica do modelo indica que a superfície é menos sensível à perda de produção na direção w_1 . A exploração da Forma Canônica requer a conversão de pontos no espaço (w_1, w_2) para pontos no espaço (x_1, x_2) .

Em geral, as variáveis x são relacionadas com as variáveis canônicas w por

$$w = M'(x - x_0) \quad (3.9)$$

onde M é uma matriz ortogonal $(k \times k)$. As colunas de M são os autovetores normalizados associados com os λ_i . Isto é, se m_i é a

i -ésima coluna de M , então, m_i é a solução de

$$(B - \lambda_i I)m_i = 0 \quad (3.10)$$

para o qual $\sum_{j=1}^k m_{ji}^2 = 1$.

No exemplo para $\lambda_1 = -0,9641$, a Equação 3.10 fica

$$\begin{bmatrix} (-1,3770 + 0,9641) & 0,1250 \\ 0,1250 & (-1,0018 + 0,9641) \end{bmatrix} \begin{bmatrix} m_{11} \\ m_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

ou

$$-0,4129m_{11} + 0,1250m_{21} = 0$$

$$0,1250m_{11} - 0,0377m_{21} = 0$$

Deseja-se obter a solução normalizada para estas equações, isto é, a solução para a qual $m_{11}^2 + m_{21}^2 = 1$. Não existe uma única solução para estas equações e é mais conveniente dar um valor arbitrário para um dos m , resolver o sistema e normalizar a solução. Fazendo $m_{21}^* = 1$, acha-se $m_{11}^* = 0,3027$. Para normalizar esta solução, divide-se m_{11}^* e m_{21}^* por

$$\sqrt{(m_{11}^*)^2 + (m_{21}^*)^2} = \sqrt{(0,3027)^2 + (1)^2} = 1,0448$$

Então, tem-se

$$m_{11} = \frac{m_{11}^*}{1,0448} = \frac{0,3027}{1,0448} = 0,2897$$

$$m_{21} = \frac{m_{21}^*}{1,0448} = \frac{1}{1,0448} = 0,9571$$

que é a primeira coluna da matriz M .

Usando $\lambda_2 = -1,4147$, repete-se o procedimento e obtém-se $m_{12} = -0,9574$ e $m_{22} = 0,2888$ como a segunda coluna de M. Assim, tem-se

$$M = \begin{bmatrix} 0,2897 & -0,9574 \\ 0,9571 & 0,2888 \end{bmatrix}$$

A relação entre as variáveis w e x é

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0,2897 & 0,9571 \\ -0,9574 & 0,2888 \end{bmatrix} = \begin{bmatrix} x_1 - 0,3890 \\ x_2 - 0,3056 \end{bmatrix}$$

ou

$$w_1 = 0,2897(x_1 - 0,3890) + 0,9571(x_2 - 0,3056)$$

$$w_2 = -0,9574(x_1 - 0,3890) + 0,2888(x_2 - 0,3056)$$

Se for desejado explorar a Superfície de Resposta na vizinhança do ponto estacionário, deve-se determinar pontos apropriados aos quais toma-se observações no espaço (w_1, w_2) e então usa-se a relação acima para converter estes pontos para o espaço (x_1, x_2) .

3.4 - Sistema de "Cumeeira" ou "Aresta".

Não é pouco comum obter superfície de resposta com caracterização diferente do que os simples ponto de máximo, mínimo ou ponto de sela. Particularmente os sistemas de "Cumeeira" ou "Aresta" (Ridge Systems) são frequentes.

Supomos que o ponto Estacionário x_0 está dentro da região de experimentação e que um ou mais dos λ_i são muito pequenos, isto é $\lambda_i \cong 0$. A resposta neste caso é insensível a variação das variáveis w_i que são multiplicadas por valores aproximadamente nulos de λ_i .

Quando λ_2 é negativo e λ_1 tende a zero ($\lambda_1 \cong 0$) está

mostrada na figura 3.7 . Nesta superfície o prolongamento na direção do eixo w_1 é infinito. A resposta é sempre a mesma quando se movimenta ao longo de qualquer linha $w_2 = \text{constante}$. Em particular, a resposta máxima pode ser obtida por qualquer combinação de valores de x_1 e x_2 onde $w_2 = 0$.

A equação da Forma Canônica fica:

$$\hat{y} = \hat{y}_0 + \lambda_2 w_2^2 \quad (3.11)$$

onde \hat{y}_0 é a resposta em qualquer ponto do eixo w_1 .

Este tipo de superfície de resposta é chamado de "Cumeeira" ou "Aresta" Estacionária.

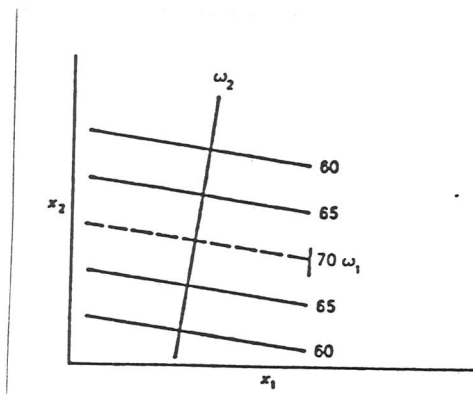


Figura 3.7 - "Cumeeira" ou "Aresta" Estacionária.

(Extraída de MONTGOMERY, 1991)

Caso o ponto estacionário esteja fora da região experimental e um ou mais λ_i , são próximos de zero, então a superfície pode ser uma "Cumeeira" ou "Aresta" crescente. Na figura 3.8 encontra-se uma "Cumeeira" ou "Aresta" crescente no caso de $k=2$. Esta figura mostra um sistema no qual λ_2 é negativo e λ_1 é próximo de zero, com centro sobre o eixo w_1 no infinito; e os contornos ajustados são parábolas. Tomando como origem algum

ponto conveniente sobre o eixo w_1 , a Forma Canônica fica:

$$\hat{y} = \hat{y}_0 + \lambda_1^* w_1 + \lambda_2 w_2^2 \quad (3.12)$$

onde λ_1^* mede a razão de incremento na resposta, isto é, a inclinação ao longo do eixo w_1 . Neste tipo de sistema não se pode inferir sobre a verdadeira superfície ou sobre o ponto estacionário, já que x_0 está fora da região onde se ajustou o modelo. Entretanto, pode-se efetuar outro experimento ao qual se ajusta uma nova superfície quadrática ao longo do eixo w_1 na direção que a resposta se incrementa.

Nas mesmas condições e quando λ_2 é positivo a superfície é chamada de "Cumeeira" ou "Aresta" decrescente.

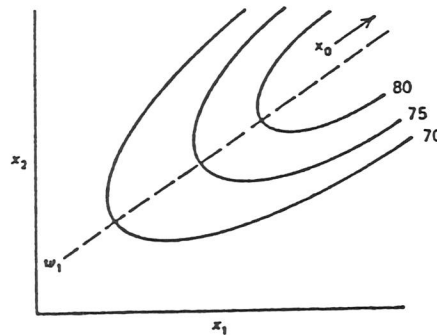


Figura 3.8- "Cumeeira" ou "Aresta" Crescente.

(Extraída de MONTGOMERY, 1991)

É importante lembrar que a equação ajustada pode ser considerada como uma aproximação adequada da superfície de resposta somente na vizinhança em torno do centro do planejamento. A equação ajustada não é de interesse em outras regiões. Por exemplo, é comum encontrar na análise do modelo

quadrático ajustado, que um máximo está indicado longe da região do planejamento, mas que a superfície está prolongada ao longo de um eixo que passa perto do planejamento. Isto indica que a experimentação prévia levou o experimentador não a um máximo, mas perto de uma "Cumeeira" ou "Aresta" crescente. Nenhuma conclusão a respeito da posição ou mesmo da existência do máximo afastado pode ser tirada, mas a atenção deve ser focalizada na elevação local e ensaios adicionais devem ser realizados ao longo deste eixo.

A distância do ponto Estacionário ao centro do delineamento é:

$$d = \left[\sum_{i=1}^k x_{i0}^2 \right]^{1/2}$$

onde x_{i0} , $i = 1, 2, \dots, k$, são as coordenadas do ponto estacionário. Quando interpretamos um sistema de "Cumeeira" ou "Aresta" crescente ou decrescente, d será, normalmente, maior do que a unidade e, conseqüentemente, conclusões sobre a natureza da superfície de resposta em relação a x_0 são arriscadas. Como foi salientado, o melhor procedimento é prosseguir a investigação ao longo da "Cumeeira" ou "Aresta" na direção do ótimo. Nestes casos, outra forma canônica, dada em 3.13, poderá ser útil:

$$\begin{aligned} \hat{y} &= \hat{\beta}_0 + \theta_1 w_1 + \theta_2 w_2 + \dots + \theta_k w_k + \lambda_1 w_1^2 + \lambda_2 w_2^2 + \dots + \lambda_k w_k^2 \\ &= \hat{\beta}_0 + w' \theta + w' \Lambda w \end{aligned} \quad (3.13)$$

onde $\theta = M'b$ e $\Lambda = \text{diagonal}(\lambda_1, \lambda_2, \dots, \lambda_k)$. Nesta forma canônica os λ 's (como antes) determinam o tipo da superfície ajustada e os θ 's medem as inclinações da superfície à origem $x=0$ nas direções dos eixos rotados w_1, w_2, \dots, w_k .

Por exemplo, seja o sistema de "Cumeeira" ou "Aresta" crescente dado na figura 3.9. A equação para esta superfície é

$$\hat{y} = 82,71 + 8,80x_1 + 8,19x_2 - 6,59x_1^2 - 2,07x_2^2 + 7,59x_1x_2$$

e pode ser mostrado que a forma canônica correspondente a equação (3.13) é $\hat{y} = 82,71 - 9,02w_1^2 + 2,97w_2^2$.

Desde que $\lambda_1=0$ e $\lambda_2=-9,02$, o sistema é uma "Cumeeira" ou "Aresta" crescente ($\lambda_2 < 0$ implica em um máximo para o ponto estacionário) e o termo linear $2,97w_2$ é a contribuição para o topo da "Cumeeira" ou "Aresta". Posterior exploração ao longo de w_2 na direção do ótimo é procedida.

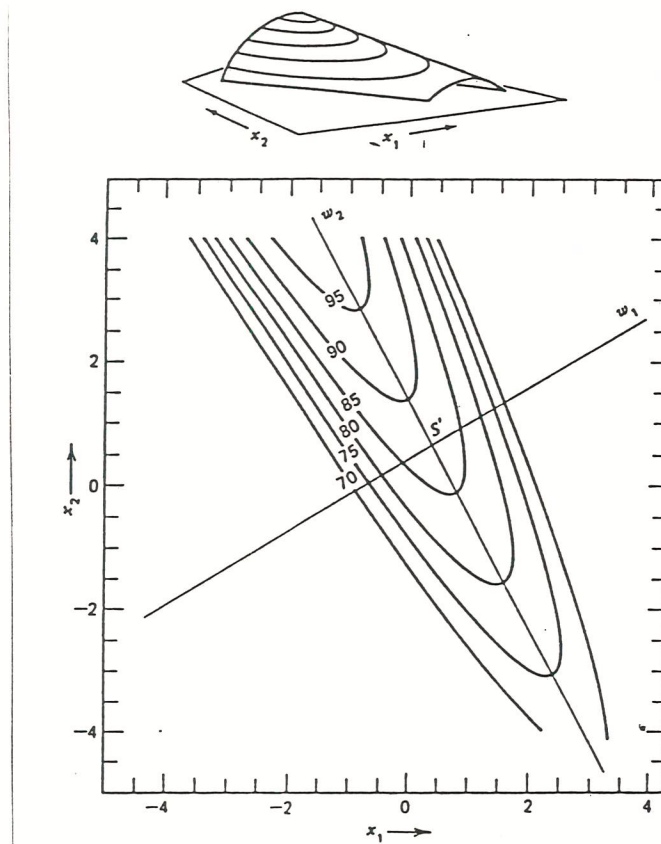


Figura 3.9 - "Cumeeira" ou "Aresta" Crescente.

(Extraída de BOX et al., 1978)

Quando tivermos três fatores ($k=3$) envolvidos no estudo de uma determinada resposta, teremos então sistemas de contornos tri-dimensionais, os quais podem ser considerados como sendo produzidos a partir daqueles em duas dimensões. Reciprocamente, sistemas em duas dimensões são "secções cruzadas" daqueles em três dimensões. Na figura 3.10 temos alguns exemplos de sistemas tri-dimensionais produzidos por equações de segunda ordem.

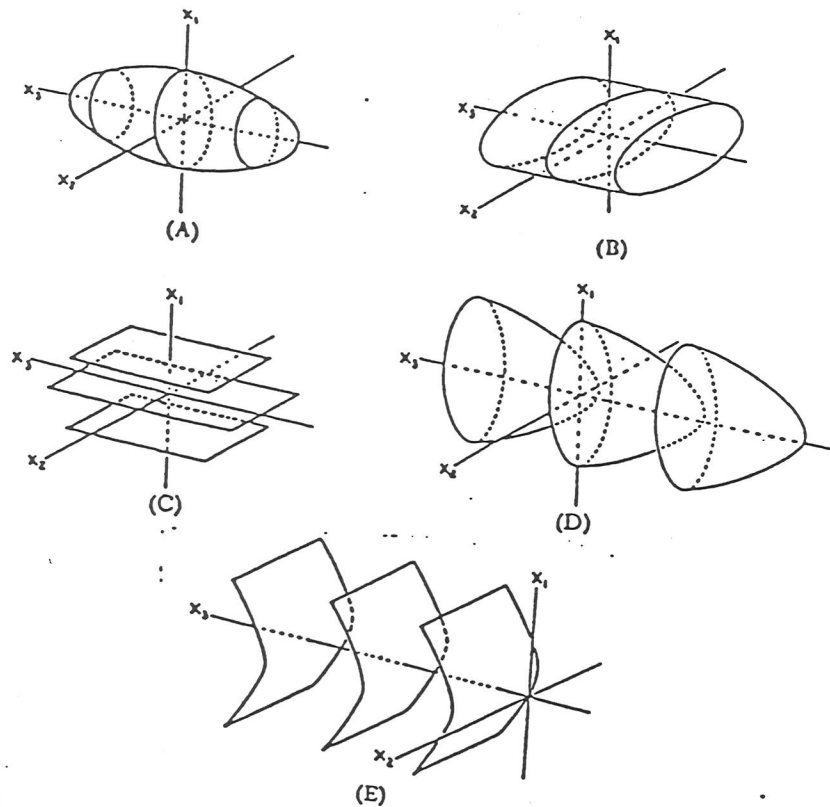


Figura 3.10 - Alguns sistemas de contorno tri-dimensionais em uma região Quase Estacionária.

(Extraída de DAVIES, 1956)

A figura 3.10 (A) mostra uma superfície na qual λ_1 , λ_2 e λ_3 são todos negativos, os contornos são elipsóides e existe um ponto de máximo.

Em (B), λ_1 e λ_2 são ambos negativos e λ_3 é zero, os contornos são cilindros elípticos, e o eixo w_3 é uma "linha máxima" (linha estacionária). Sobre esta linha uma resposta máxima seria obtida, e partindo dela em qualquer direção resultaria em decréscimo na resposta.

Em (C), λ_2 e λ_3 são ambos zero, a superfície contém um plano máximo e os contornos são planos paralelos ao plano que contém os eixos w_2 e w_3 .

Em (D), os contornos são parabolóides elípticos. O centro

está no infinito e λ_3 é zero. Pegando a origem no eixo w_3 , a equação do segundo grau pode ser reduzida à forma

$$\hat{y} = \hat{y}_0 + \lambda_1 w_1^2 + \lambda_2 w_2^2 + \lambda_3^* w_3 \quad (3.14)$$

onde λ_1 e λ_2 são negativos, λ_3 é zero e λ_3^* mede a inclinação do eixo w_3 . O sistema é uma linha estacionária como (B), mas a elevação é crescente em vez de estacionária.

O sistema correspondente a (C), mas representando um aclave em vez de uma elevação estacionária, está mostrado em (E).

Em geral quando um ou mais dos coeficientes, $\lambda_1, \lambda_2, \dots, \lambda_k$ na Forma Canônica são pequenos, comparados com os outros, alguma forma de sistema de "Cumeeira" ou "Aresta" é indicada. Quando o centro do sistema está longe da vizinhança do centro do planejamento, o sistema de "Cumeeira" ou "Aresta" será aproximadamente estacionário. Mesmo quando o centro do sistema está longe da vizinhança do planejamento, quase sempre será descoberto que uma "Cumeeira" ou "Aresta" crescente passa perto da vizinhança do planejamento e para determinar sua natureza, uma nova origem é escolhida na elevação. Por exemplo, se o coeficiente λ_t é pequeno comparado com os outros, então o eixo w_t define uma "Cumeeira" ou "Aresta". Uma nova origem é tomada neste eixo perto do planejamento, e então a correspondente equação conterà um coeficiente λ_t^* o qual, se o centro está longe do planejamento, mede a inclinação da "Cumeeira" ou "Aresta". Quando dois coeficientes λ_t e λ_s são pequenos comparados com os outros, o plano contendo os eixos w_t e w_s define uma "Cumeeira" ou "Aresta" o qual passará perto do planejamento. Uma nova origem é tomada neste plano, e então a equação correspondente conterà coeficientes λ_t^* e λ_s^* . Se estes coeficientes não são pequenos, uma "Cumeeira" ou "Aresta" crescente é indicada e, a direção de maior incremento sobre ela será encontrada pelo Método da Máxima Inclinação Ascendente.

4 - DELINEAMENTOS DE SUPERFÍCIE DE RESPOSTA

Uma superfície de resposta pode ser ajustada mais eficientemente se atenção apropriada é dada à escolha do delineamento experimental.

4.1. Delineamentos para Ajustar Modelos de Primeira Ordem

Suponha que se quer ajustar o modelo de primeira ordem (relação linear entre a resposta y e os k fatores)

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \varepsilon \quad (4.1)$$

Existe uma única classe de delineamentos que minimizam a variância dos coeficientes de regressão $\hat{\beta}_i$, são os Delineamentos de Primeira Ordem *Ortogonais*. Nestes delineamentos os elementos fora da diagonal principal da matriz $(X'X)$ são todos nulos. Isto implica que os produtos cruzados das colunas da matriz X somam zero.

A classe desses delineamentos inclui o Fatorial 2^k e as frações das séries 2^k , nas quais um efeito principal não tem como alias outro efeito principal. No uso destes delineamentos, assume-se que os fatores são codificados aos níveis ± 1 . Como um exemplo, tem-se o delineamento 2^3 para ajustar o modelo de primeira ordem

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon \quad (4.2)$$

A matriz X para ajustar este modelo é

$$X = \begin{matrix} & \beta_0 & \beta_1 & \beta_2 & \beta_3 \\ \begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{matrix}$$

É fácil verificar que os elementos fora da diagonal de $(X'X)$ são nulos para este delineamento.

O delineamento 2^k não fornece uma estimação da variância do erro experimental, com a qual pode-se verificar se a equação linear se ajusta adequadamente. Esta variância pode ser obtida:

- i) pela repetição de todo o experimento;
- ii) pelo uso de uma estimação de uma experimentação anterior, caso exista evidência de que a variância do erro permanece estável através do tempo;
- iii) agregando ao fatorial 2^k várias observações no centro (o ponto $x_i=0, i=1, 2, \dots, k$). Esta adição de pontos centrais não influencia os $\hat{\beta}_i$ para $i \geq 1$, mas a estimativa de β_0 torna-se a média geral de todas as observações do experimento. Além disso, esta adição de pontos centrais não altera a propriedade de ortogonalidade do delineamento.

Suponha que o ponto central se repete n_2 vezes. A soma de quadrados dos desvios das n_2 respostas no centro, de sua média, dão $(n_2 - 1)$ graus de liberdade para estimar o erro experimental. Se \bar{y}_2 é a resposta média dos pontos centrais, e \bar{y}_1 é a resposta média dos pontos do fatorial 2^k , o contraste $(\bar{y}_1 - \bar{y}_2)$ fornece um grau de liberdade adicional para medir a falta de ajuste. Se a superfície de resposta verdadeira é quadrática, em vez de linear, este contraste estima a soma dos coeficientes $\beta_{11}, \beta_{22}, \dots, \beta_{kk}$ dos termos $x_1^2, x_2^2, \dots, x_k^2$. A soma de quadrados para este único grau de liberdade se calcula da forma:

$$\frac{n_1 n_2}{n_1 + n_2} (\bar{y}_1 - \bar{y}_2)^2 \quad (4.3)$$

onde n_1 é o número de pontos do fatorial 2^k .

Outro delineamento de primeira ordem ortogonal é o *Simplex*. Ele é uma figura de lados regulares com $k+1$ vértices em k dimensões. Para $k=2$ o delineamento Simplex é um triângulo equilátero, enquanto, para $k=3$ ele é um tetraedro regular. Delineamentos simplex para 2 e 3 dimensões aparecem na figura 4.1.

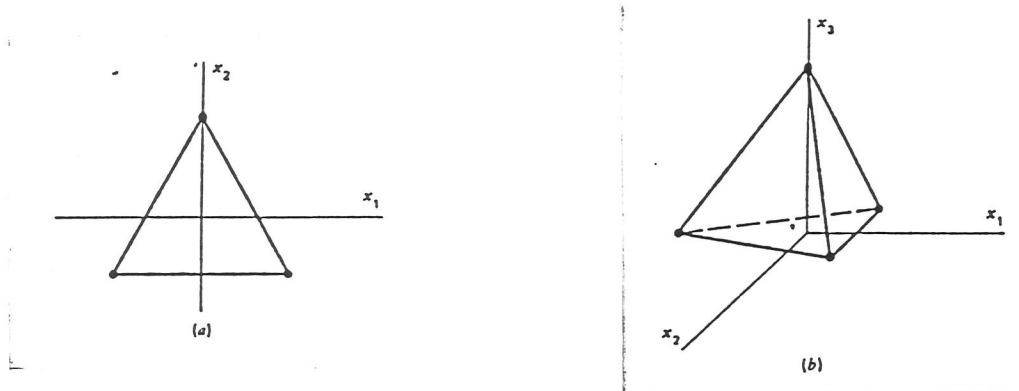


Figura 4.1 Delineamento Simplex para (a) $k=2$ e (b) $k=3$ variáveis
(Extraída de MONTGOMERY, 1991)

4.2 Delineamentos para Ajustar Modelos de Segunda Ordem

São usados para ajustar o modelo de segunda ordem (polinômio quadrático):

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{\substack{i < j \\ i, j}} \beta_{ij} x_i x_j + \varepsilon \quad (4.4)$$

Um delineamento experimental para ajustar este modelo deve ter no mínimo três níveis para cada fator para que os coeficientes possam ser estimados. Se os três níveis de cada variável são codificados como $-1, 0, 1$ a superfície de segundo grau se ajusta facilmente aos resultados de um fatorial 3^k . Mas com mais de três fatores, os experimentos se tornam grandes e desvantajosos. Por este motivo, foram desenvolvidos novos delineamentos para serem usados na Metodologia de Superfície de Resposta.

4.2.1 Delineamento Central Composto

O procedimento consiste em escolher um planejamento no qual todos os efeitos de primeira ordem (efeitos principais) e todos os efeitos de interação (de 2 fatores) de segunda ordem possam ser estimados. Esse planejamento é então suplementado com pontos adicionais de modo a permitir a estimação dos efeitos quadráticos β_{11} , β_{22} , ..., β_{kk} . Ainda, adiciona-se pontos centrais para obter uma estimativa da variância residual (erro experimental).

Então, ao planejamento fatorial 2^k ou fatorial fracionário (codificado com a notação usual ± 1) acrescenta-se n_0 pontos centrais $(0, 0, \dots, 0)$ e $2k$ pontos axiais

$(\pm\alpha, 0, \dots, 0), (0, \pm\alpha, \dots, 0), \dots, (0, 0, \dots, \pm\alpha)$

Delineamentos Centrais Compostos são mostrados na Figura 4.2. Uma vantagem é que podem ser realizados em estágios. Principia-se com um fatorial 2^k ao qual se ajusta uma superfície de resposta linear e aplica-se o Método de Máxima Inclinação Ascendente. Quando o ajuste linear não for mais significativo e os efeitos de segunda ordem forem, basta acrescentar as $2k$ combinações de tratamentos (pontos axiais) para ajustar a superfície quadrática.

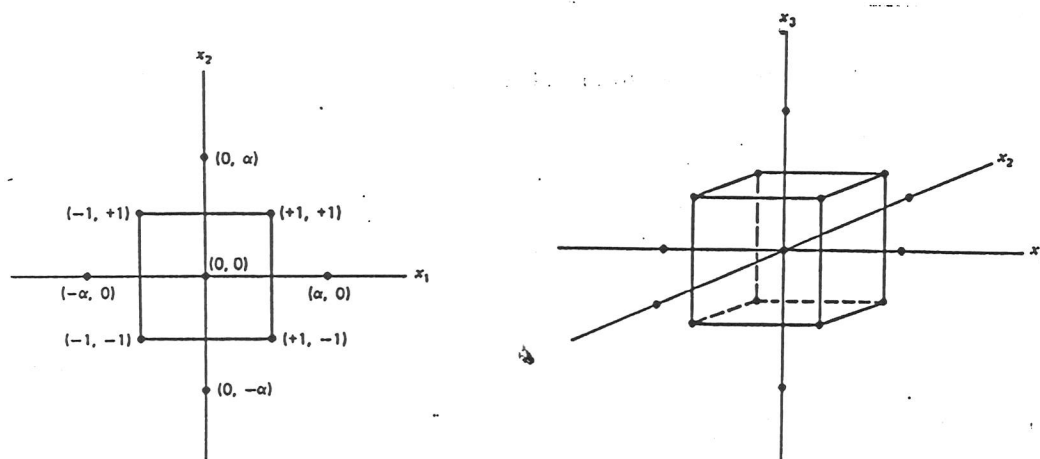


Figura 4.2. Delineamentos Centrais Compostos para $k=2$ e $k=3$
(Extraída de MONTGOMERY, 1991)

4.2.2. Delineamentos Compostos Ortogonais

Pode-se construir Planejamentos Compostos através da escolha adequada do valor de α . Este deve ser escolhido de modo que a distância (α) do ponto central a cada um dos pontos axiais seja tal que todos os efeitos e interações estimados no modelo de segunda ordem sejam ortogonais. Para que isto seja possível, a distância α é dada pela expressão

$$\alpha = \left\{ \frac{[(F + T)^{1/2} - F^{1/2}]^2 F}{4n^2} \right\}^{1/4} \quad (4.5)$$

onde F é o número de combinações de tratamentos no planejamento fatorial tomado como núcleo; e T é o número de pontos adicionais multiplicado pelo número de observações por combinação de tratamento.

Outras propriedades do Delineamento Central Composto podem ser controladas pela escolha do número de pontos centrais, n_0 . Com a escolha apropriada de n_0 , o Delineamento Central Composto pode ser feito *Ortogonal*, ou pode ser feito um Delineamento de *Precisão Uniforme*. Neste último, a variância da resposta estimada \hat{y} na origem é igual à variância de \hat{y} numa distância unitária da origem.

4.2.3. Delineamentos Rotacionais

A classe de Delineamentos de Superfície de Resposta de Segunda Ordem preferida é classe de Delineamentos Rotacionais, pois oferece condição para que a resposta seja estimada com a mesma precisão para pontos que estejam a uma mesma distância do centro. Um delineamento experimental é dito rotacional se a variância da resposta estimada \hat{y} a um ponto x é função somente da distância do ponto ao centro do planejamento, e não uma função da direção. Isto implica que os contornos de variância de \hat{y} sejam círculos concêntricos. Além do mais, um delineamento com esta propriedade deixa a variância de \hat{y} invariante quando o delineamento é rotado em torno do centro $(0, 0, \dots, 0)$.

Um Delineamento Central Composto é feito rotacional pela escolha de α . Neste caso, o valor de α depende do número de pontos na parte fatorial do delineamento, isto é,

$$\alpha = (F)^{1/4} \quad (4.6)$$

onde F é o número de pontos usados na parte fatorial.

A tabela 4.1 mostra os parâmetros para os Delineamentos Centrais Compostos Rotacionais de Precisão Uniforme e Ortogonais para vários valores de k .

Tabela 4.1. Delineamentos Centrais Compostos Rotacionais de Precisão Uniforme e Ortogonais

k	2	3	4	5	5	6	6	7	8
fração					1/2		1/2	1/2	1/2
F	4	8	16	32	16	64	32	64	128
Pontos Axiais	4	6	8	10	10	12	12	14	16
n_o (P.U.)	5	6	7	10	6	15	9	14	20
n_o (Ort.)	8	9	12	17	10	24	15	22	33
N (P.U.)	13	20	31	52	32	91	53	92	164
N (Ort.)	16	23	36	59	36	100	59	100	177
α	1.414	1.682	2.000	2.378	2.000	2.828	2.378	2.828	3.364

4.2.4. Outros Delineamentos

Existem outros Delineamentos Rotacionais que são ocasionalmente usados em problemas envolvendo duas ou três variáveis. Consistem de pontos igualmente espaçados sobre um círculo ($k=2$) ou uma esfera ($k=3$) e são polígonos regulares ou poliedros. Como os pontos estão equidistantes da origem, estes arranjos são chamados de Delineamentos *Equiradiais*.

Para $k=2$, um Delineamento Equiradial Rotacional é obtido combinando $n_1 \geq 5$ pontos igualmente espaçados sobre um círculo com $n_2 \geq 1$ pontos no centro do círculo. Particularmente, delineamentos usados são o Hexagonal e o Pentagonal. Estes são mostrados na Figura 4.3. Para $k=3$ somente os arranjos equiradiais que contém pontos suficientes para permitir que todos os parâmetros do modelo de segunda ordem possam ser estimados são o Icosaedro (20 pontos) e o Dodecaedro (12 pontos).

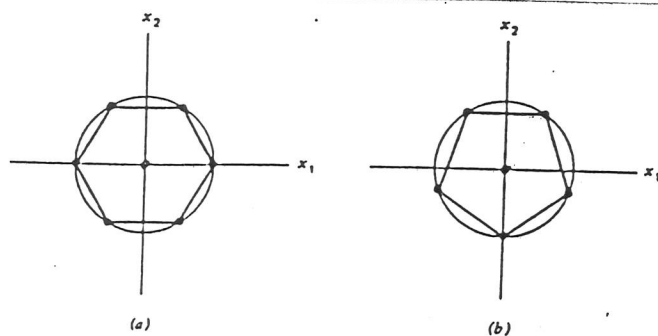


Figura 4.3. Delineamentos Equiradiais para 2 variáveis
(a) Hexagonal, (b) Pentagonal.
(Extraída de MONTGOMERY, 1991)

5. UTILIZAÇÃO DO PROGRAMA RSREG DO PACOTE SAS PARA A METODOLOGIA DE SUPERFICIE DE RESPOSTA

O Proc RSREG do pacote SAS (Statistical Analysis System) realiza a análise de modelos quadráticos e faz o gráfico de contornos.

Para exemplificar a utilização do Proc RSREG do SAS utilizar-se-á o exemplo da seção 3.3, cujos dados aparecem na figura 3.1. Foram usadas as variáveis codificadas x_1 e x_2 e a resposta y para proceder a análise.

5.1 - Programa

O programa para o RSREG, incluindo a realização do gráfico de contornos, para o exemplo é o seguinte:

```
DATA A;
INFILE 'A:EXEMPLO.DAT';
INPUT X1 X2 Y;
LABEL
  X1="TEMPO DE REACAO"
  X2="TEMPERATURA DE REACAO"
  Y="PRODUCAO";
PROC RSREG;
  MODEL Y=X1 X2 / LACKFIT;
DATA B;
  SET A END=EOF;
  OUTPUT;
  IF EOF THEN DO;
    Y=.;
    DO X1=-1.5 TO 1.5 BY .1;
      DO X2=-1.5 TO 1.5 BY .1;
        OUTPUT;
      END;
    END;
  END;
PROC RSREG DATA=B OUT=C NOPRINT;
  MODEL Y=X1 X2 / PREDICT;
DATA D;
  SET C;
PROC PLOT DATA=D;
PLOT X1*X2=Y / CONTOUR=6 HPOS=100 VPOS=36 HSPACE=10
      HAXIS=-1.5 TO 1.5 BY .5
      VAXIS=-1.5 TO 1.5 BY .5;
RUN;
```

5.2 - Dados

Os dados utilizados na análise são listados a seguir :

-1	-1	76.5
-1	1	77.0
1	-1	78.0
1	1	79.5
0	0	79.9
0	0	80.3
0	0	80.0
0	0	79.7
0	0	79.8
1.414	0	78.4
-1.414	0	75.6
0	1.414	78.5
0	-1.414	77.0

5.3 - Resultados

A listagem da saída do Proc RSREG é apresentada a seguir :

Coding Coefficients for the Independent Variables

	Factor	Subtracted off	Divided by
1	X1	0	1.414000
	X2	0	1.414000

Response Surface for Variable Y: PRODUCAO

2	Response Mean	78.476923
3	Root MSE	0.266290
4	R-Square	0.9827
5	Coef. of Variation	0.3393

6	7	8	9	10	11
Regression	Degrees of Freedom	Type I Sum of Squares	R-Square	F-Ratio	Prob > F
Linear	2	10.042955	0.3494	70.814	0.0000
Quadratic	2	17.953749	0.6246	126.6	0.0000
Crossproduct	1	0.250000	0.0087	3.526	0.1025
Total Regress	5	28.246703	0.9827	79.669	0.0000

Residual	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Prob > F
12 Lack of Fit	3	0.284373	0.094791	1.789	0.2886
Pure Error	4	0.212000	0.053000		
Total Error	7	0.496373	0.070910		

13

Parameter	Degrees of Freedom	14 Parameter Estimate	15 Standard Error	16 T for H0: Parameter=0	17 Prob > T
INTERCEPT	1	79.939955	0.119089	671.3	0.0000
X1	1	0.995050	0.094155	10.568	0.0000
X2	1	0.515203	0.094155	5.472	0.0009
X1*X1	1	-1.376449	0.100984	-13.630	0.0000
X2*X1	1	0.250000	0.133145	1.878	0.1025
X2*X2	1	-1.001336	0.100984	-9.916	0.0000

Parameter	Parameter Estimate from Coded Data	18
INTERCEPT	79.939955	
X1	1.407001	
X2	0.728497	
X1*X1	-2.752067	
X2*X1	0.499849	
X2*X2	-2.002067	

19 Factor	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Prob > F
X1	3	21.344008	7.114669	100.3	0.0000
X2	3	9.345251	3.115084	43.930	0.0001

Canonical Analysis of Response Surface
(based on coded data)

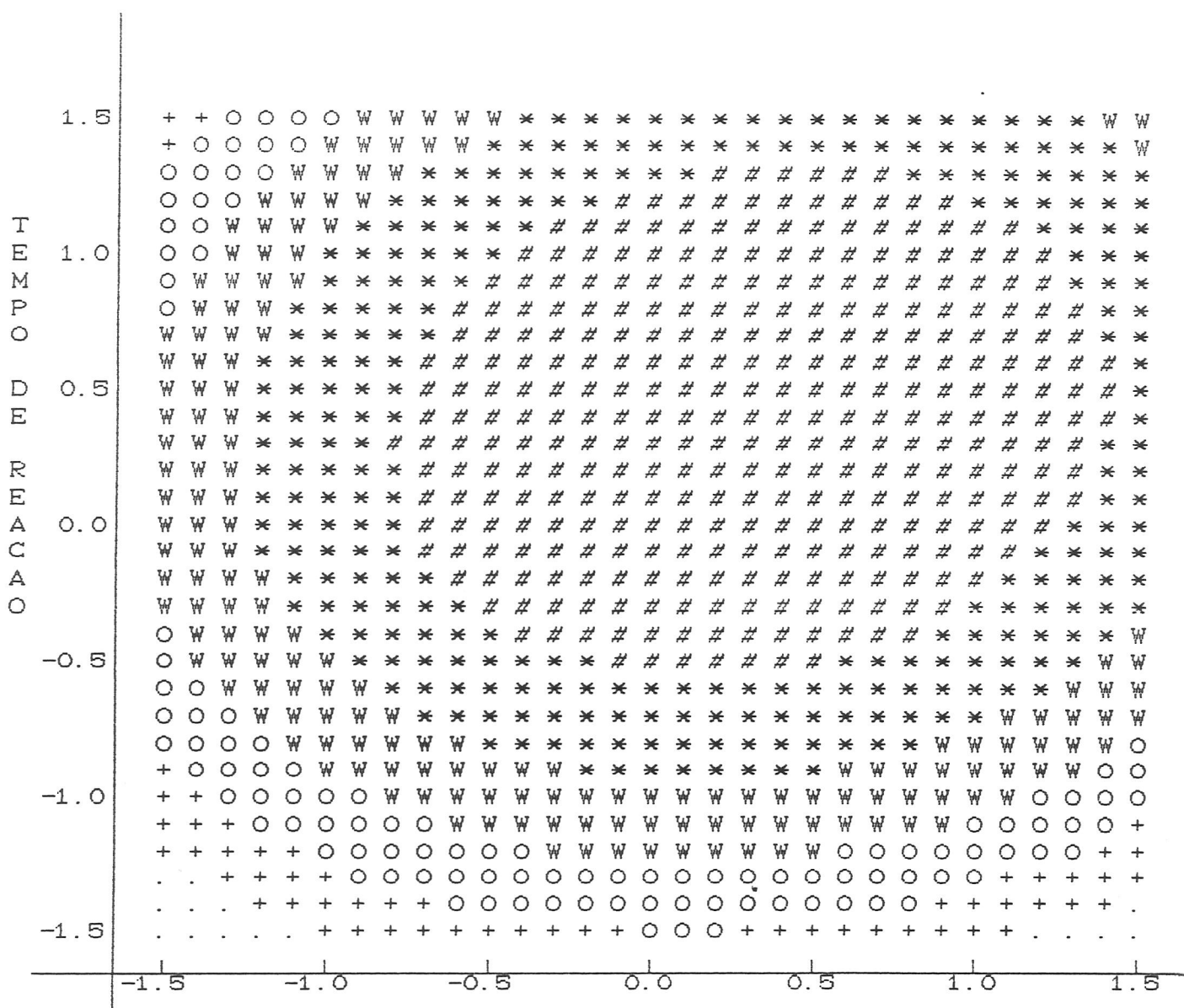
Factor	20 Critical Value		
	Coded	Uncoded	
X1	0.275269	0.389230	TEMPO DE REACAO
X2	0.216299	0.305847	TEMPERATURA DE REACAO

Predicted value at stationary point 80.212393

21 Eigenvalues	Eigenvectors	
	X1	X2
-1.926415	0.289717	0.957112
-2.827719	0.957112	-0.289717

Stationary point is a maximum.

Contour plot of X1*X2.



O significado de cada item dos resultados é o seguinte :

1. Estimação e análise de variância

1 - Operação de codificação onde para cada variável obtém-se valor codificado = $(\text{valor original} - M)/S$ onde M é a média entre o maior e o menor nível do fator (variável independente) e S é a metade da diferença entre estes níveis.

2 - Média da variável resposta.

3 - Raiz quadrada do quadrado médio do erro total. Estima a variabilidade residual.

4 - Coeficiente de determinação R^2 . Representa a proporção da variação na variável resposta que é atribuída ao modelo.

5 - Coeficiente de variação.

6 - Fontes de variação referentes aos componentes de regressão, indicando o conjunto de termos lineares, quadráticos e de interação linear, e todos os termos do modelo considerados conjuntamente.

7 - Os GL de cada conjunto de componentes indicado em 6.

8 - As somas de quadrados do tipo I, conhecida como soma de quadrados sequencial, ou seja mede a redução na soma de quadrados para cada conjunto de componentes (linear, quadrático e de interação linear) adicionado ao modelo.

9 - O coeficiente de determinação R^2 , que mede a proporção do R^2 total devido a contribuição quando cada conjunto de termos (linear, quadrático e de interação linear) é adicionado ao modelo.

10 - Valor da estatística F para testar a hipótese que todos os parâmetros no conjunto de termos são nulos, usando o quadrado médio do erro total como denominador. O teste de hipótese é do tipo I, ou seja o numerador do teste F é obtido na condição de que os efeitos das variáveis subsequentes não estão no modelo.

11 - Prob > F, é o nível mínimo de significância (n.m.s) que indica a probabilidade de se obter no mínimo um valor igual ao F calculado dado que a hipótese de nulidade é verdadeira.

12 - A soma de quadrados para soma total pode ser particionada em falta de ajuste (Lack of Fit) e erro puro. Quando a falta de ajuste é significativa, indica que existem outros componentes (tal como os componentes cúbicos dos fatores) que deveriam ser considerados no modelo e portanto isolados do erro total.

13 - Quadrado médio do erro total, que é estimativa da variância σ^2 .

14 - Estimativas dos parâmetros do modelo com base nos valores não codificados para os fatores (codificação subentende-se a codificação do SAS designada em 1). Se um efeito é combinação linear de efeitos anteriores, parâmetro do efeito não é estimado. Quando isto ocorre o grau de liberdade é nulo e a estimativa do parâmetro é tomada como zero, e as estimativas e teste dos demais parâmetros é condicional ao parâmetro não estimado e tomado como zero, o que não é mostrado.

15 - Estimativa do erro padrão para as estimativas dos parâmetros com base nos dados não codificados.

16 - Teste t para H_0 : parâmetro = 0, apresenta os valores de t para o teste da hipótese de nulidade que o verdadeiro parâmetro é nulo, quando os valores não codificados para os fatores são usados.

17 - Prob $|t|$ apresenta o nível mínimo de significância (n.m.s) que representa a probabilidade de se ter no mínimo um valor , em módulo, no mínimo igual ao t calculado dado que a hipótese de nulidade é verdadeira.

18 - Estima os parâmetros para os dados codificados (codificação SAS, indicado em 1). Estas estimativas são usadas na análise canônica e análise de "cumeeira" (Ridge Analysis).

19 - Teste para o fator consiste no teste conjunto para todos os parâmetros que envolvem o fator. No exemplo, o teste para x_1 consiste da hipótese de nulidade onde os parâmetros de x_1 , x_1^2 e x_1x_2 são todos nulos.

II - Análise Canônica

20 - Valores críticos. Valores para os fatores no ponto estacionário da superfície ajustada , que pode ser um ponto de máximo, mínimo ou de sela.

21 - Os autovalores e autovetores da matriz das estimativas dos parâmetros dos termos quadráticos, baseado nos dados codificados (codificação SAS apresentada em 1). Eles caracterizam a forma da superfície ajustada com base nos dados codificados (codificação SAS indicada em 1), eles serão em geral, diferentes daqueles obtidos no Proc RSREG de versões anteriores do SAS. Pela mesma razão são diferentes daquelas obtidas na seção 3.3.

Quando a superfície estimada não tem um ótimo único, utiliza-se a análise de "cumeeira" (Ridge Analysis) para determinar a região na qual se tem a linha de ótimo.

No exemplo, o título de ilustração, uma vez que não há necessidade, pois existe um ótimo único, o programa básico, executando a parte para obter o gráfico de contornos, é o seguinte:

```
DATA A;  
INFILE 'A:EXEMPLO.DAT';  
INPUT X1 X2 Y;
```



```

LABEL
  X1="TEMPO DE REACAO"
  X2="TEMPERATURA DE REACAO"
  Y="PRODUCAO";
PROC RSREG;
  MODEL Y=X1 X2 / LACKFIT;
RIDGE MAX;
RUN;

```

Os resultados obtidos são os mesmos anteriores, acrescentando as estimativas Ridge de resposta máxima que são os seguintes :

Estimated Ridge of Maximum Response for Variable Y: PRODUCAO

22 Coded Radius	23 Estimated Response	23 Standard Error	24 Uncoded Factor X1	Values X2
0.0	79.939955	0.119089	0	0
0.1	80.074757	0.118658	0.121681	0.072026
0.2	80.163290	0.117567	0.235056	0.157240
0.3	80.206973	0.116426	0.339864	0.253847
0.4	80.207052	0.116292	0.436470	0.359718
0.5	80.164560	0.118620	0.525650	0.472801
0.6	80.080325	0.125042	0.608353	0.591345
0.7	79.954995	0.136954	0.685531	0.713969
0.8	79.789072	0.155143	0.758051	0.839627
0.9	79.582948	0.179744	0.826659	0.967546
1.0	79.336925	0.210484	0.891984	1.097160

O significado de cada item desses resultados é o seguinte :

III - Análise de "cumeeira"(Ridge Analysis)

22 - Raios codificados. É a distância do ponto codificado associado a origem codificada da "cumeeira". A origem é dada pelo ponto de raio zero.

23 - A resposta estimada é o valor estimado no ponto associado. É dado também o erro padrão da estimativa. Esta quantidade é usada para assegurar a relativa credibilidade da predição em um dado raio. Tipicamente este erro padrão aumenta rapidamente quando a "cumeeira" ou "aresta" é conduzida para distante do perímetro do delineamento, refletindo a dificuldade

inerente de produzir predições longe da região de experimentação.

24 - Valores não codificados para os fatores. São os valores não codificados para os fatores que fornecem a resposta ótima ao raio estabelecido da origem da "cumeeira" ou "aresta".

6. EXEMPLOS DE APLICAÇÃO DA METODOLOGIA DE SUPERFÍCIE DE RESPOSTA

6.1 - EXEMPLO 1

6.1.1 - Dados

Schneider e Stockett (1963) citados por JOHN (1971) realizaram um experimento com o objetivo de verificar a influência dos fatores Temperatura (ξ_1), taxa de gás/líquido (ξ_2) e altura da embalagem (ξ_3), na redução do odor desagradável de um produto químico que está sendo estocado para uso residencial.

O delineamento utilizado foi um cubo octahedron com três pontos (tratamentos) centrais.

Os valores das variáveis independentes (fatores) e do odor y são apresentados na tabela 6.1 .

Tabela 6.1 - Dados do exemplo 6.1

Variáveis Naturais			Variáveis Codificadas			Resposta
ξ_1	ξ_2	ξ_3	x_1	x_2	x_3	y
40	0.3	4	-1	-1	0	66
120	0.3	4	1	-1	0	39
40	0.7	4	-1	1	0	43
120	0.7	4	1	1	0	49
40	0.5	2	-1	0	-1	58
120	0.5	2	1	0	-1	17
40	0.5	6	-1	0	1	-5
120	0.5	6	1	0	1	-40
80	0.3	2	0	-1	-1	65
80	0.7	2	0	1	-1	7
80	0.3	6	0	-1	1	43
80	0.7	6	0	1	1	-22
80	0.5	4	0	0	0	-31
80	0.5	4	0	0	0	-35
80	0.5	4	0	0	0	-26

As codificações utilizadas foram

$$x_1 = \frac{\xi_1 - 80}{40}, \quad x_2 = \frac{\xi_2 - 0.5}{0.5}, \quad x_3 = \frac{\xi_3 - 4}{2}$$

6.1.2 - Resultados

Os comandos do PROC RSREG do SAS para o exemplo 6.1 e os respectivos resultados são dados a seguir:

```
DATA A;
INFILE 'A:EXEMPLO1';
INPUT Y X1 X2 X3;
  LABEL Y="ODOR"
        X1="TEMPERATURE"
        X2="GAS-LIQUID RATIO"
        X3="PACKING HEIGHTH";
PROC RSREG;
  MODEL Y=X1-X3 / LACKFIT;
RUN;
DATA B;
  SET A END=EOF;
  OUTPUT;
  IF EOF THEN DO;
    Y=.;
    X3=1.77;
    DO X1=-1.5 TO 1.5 BY .1;
      DO X2=-2 TO 2 BY .1;
        OUTPUT;
      END;
    END;
  END;
PROC RSREG DATA=B OUT=C NOPRINT;
  MODEL Y=X1-X3 / PREDICT;
DATA D;
  SET C;
  IF X3=1.77;
PROC PLOT DATA=D;
  PLOT X1*X2=Y / CONTOUR=6 HPOS=100 VPOS=36 HSPACE=10
        HAXIS=-1.5 TO 1.5 BY .5
        VAXIS=-1.5 TO 1.5 BY .5;

RUN;
```

66	-1	-1	0
39	1	-1	0
43	-1	1	0
49	1	1	0
58	-1	0	-1
17	1	0	-1
-5	-1	0	1
-40	1	0	1
65	0	-1	-1
7	0	1	-1
43	0	-1	1
-22	0	1	1
-31	0	0	0
-35	0	0	0
-26	0	0	0

Coding Coefficients for the Independent Variables

Factor	Subtracted off	Divided by
X1	0	1.000000
X2	0	1.000000
X3	0	1.000000

Response Surface for Variable Y: ODOR

Response Mean	15.200000
Root MSE	22.478508
R-Square	0.8820
Coef. of Variation	147.8849

Regression	Degrees of Freedom	Type I Sum of Squares	R-Square	F-Ratio	Prob > F
Linear	3	7143.250000	0.3337	4.712	0.0641
Quadratic	3	11445	0.5346	7.550	0.0264
Crossproduct	3	293.500000	0.0137	0.194	0.8965
Total Regress	9	18882	0.8820	4.152	0.0657

Residual	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Prob > F
Lack of Fit	3	2485.750000	828.583333	40.750	0.0240
Pure Error	2	40.666667	20.333333		
Total Error	5	2526.416667	505.283333		

Parameter	Degrees of Freedom	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEPT	1	-30.666667	12.977973	-2.363	0.0645
X1	1	-12.125000	7.947353	-1.526	0.1876
X2	1	-17.000000	7.947353	-2.139	0.0854
X3	1	-21.375000	7.947353	-2.690	0.0433
X1*X1	1	32.083333	11.698187	2.743	0.0407
X2*X1	1	8.250000	11.239254	0.734	0.4959
X2*X2	1	47.833333	11.698187	4.089	0.0095
X3*X1	1	1.500000	11.239254	0.133	0.8990
X3*X2	1	-1.750000	11.239254	-0.156	0.8824
X3*X3	1	6.083333	11.698187	0.520	0.6252

Parameter	Parameter Estimate from Coded Data
INTERCEPT	-30.666667
X1	-12.125000
X2	-17.000000
X3	-21.375000
X1*X1	32.083333
X2*X1	8.250000
X2*X2	47.833333
X3*X1	1.500000
X3*X2	-1.750000
X3*X3	6.083333

Factor	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Prob > F	
X1	4	5258.016026	1314.504006	2.602	0.1613	TEMPERATURE
X2	4	11045	2761.150641	5.465	0.0454	GAS-LIQUID RATIO
X3	4	3813.016026	953.254006	1.887	0.2510	PACKING HEIGHT

Canonical Analysis of Response Surface
(based on coded data)

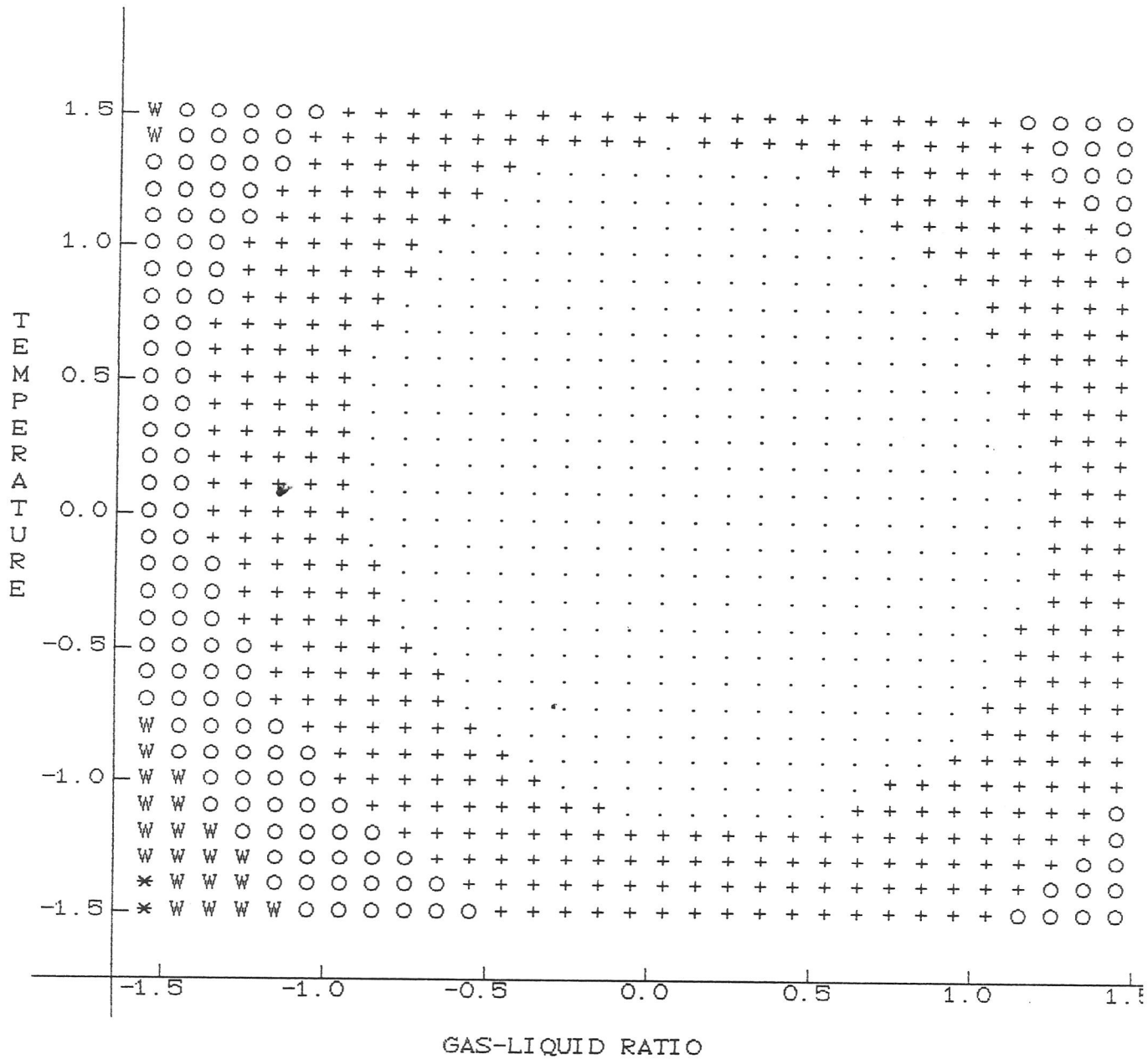
Factor	Critical Value		
	Coded	Uncoded	
X1	0.121913	0.121913	TEMPERATURE
X2	0.199575	0.199575	GAS-LIQUID RATIO
X3	1.770525	1.770525	PACKING HEIGHT

Predicted value at stationary point -52.024631

Eigenvalues	Eigenvectors		
	X1	X2	X3
48.858807	0.238091	0.971116	-0.015690
31.103461	0.970696	-0.237384	0.037399
6.037732	-0.032594	0.024135	0.999177

Stationary point is a minimum.

Contour plot of X1*X2.



Symbol	Y	Symbol	Y	Symbol	Y
.....	-60 - 0	OOOOO	60 - 120	*****	180 - 240
+++++	0 - 60	WWWWW	120 - 180	#####	240 - 300

6.2 - EXEMPLO 2

6.2.1 - Dados

Realizou-se um experimento para verificar a combinação ótima de açúcar e sal na paçoca de soja. Nesse experimento organoléptico foram utilizados 11 degustadores, onde os degustadores davam nota de 1 a 10 numa escala hedônica. Registrou-se como valor y a resposta média das notas. Os teores de açúcar e sal são dados em %. O experimento em si aparece na tabela 6.2.

Tabela 6.2 - Dados do exemplo 6.2

Variáveis Naturais		Variáveis Codificadas		Resposta
ξ_1	ξ_2	x_1	x_2	y
20	0.4	-1	-1	4.00
100	0.4	1	-1	6.45
20	0.6	-1	1	5.00
100	0.6	1	1	6.09
3.44	0.5	-1.414	0	3.73
116.56	0.5	1.414	0	6.64
60	0.3586	0	-1.414	6.00
60	0.6434	0	1.414	5.64
60	0.5	0	0	7.00
60	0.5	0	0	6.64
60	0.5	0	0	6.64
60	0.5	0	0	6.64
60	0.5	0	0	6.91

As codificações utilizadas foram

$$x_1 = \frac{\xi_1 - 60}{40}, \quad x_2 = \frac{\xi_2 - 0.5}{0.1}$$

6.2.2 - Resultados

Os comandos do PROC RSREG do SAS para o exemplo 6.2 e os resultados obtidos são os seguintes :

```
DATA A;
INFILE 'A:EXEMPLO2';
INPUT X1 X2 Y;
LABEL
  X1="ACUCAR"
  X2="SAL"
  Y="SABOR";
PROC RSREG;
  MODEL Y=X1 X2 / LACKFIT;
DATA B;
  SET A END=EOF;
  OUTPUT;
  IF EOF THEN DO;
    Y=.;
    DO X1=-1.5 TO 1.5 BY .1;
      DO X2=-1.5 TO 1.5 BY .1;
        OUTPUT;
        END;
      END;
    END;
  END;
PROC RSREG DATA=B OUT=C NOPRINT;
  MODEL Y=X1 X2 / PREDICT;
DATA D;
  SET C;
PROC PLOT DATA=D;
PLOT X1*X2=Y / CONTOUR=6 HPOS=100 VPOS=36 HSPACE=10
      HAXIS=-1.5 TO 1.5 BY .5
      VAXIS=-1.5 TO 1.5 BY .5;
RUN;
```

```

-1      -1  4.00
  1      -1  6.45
-1       1  5.00
  1       1  6.09
-1.414   0  3.73
 1.414   0  6.64
  0 -1.414  6.00
  0  1.414  5.64
  0       0  7.00
  0       0  6.64
  0       0  6.64
  0       0  6.64
  0       0  6.64
  0       0  6.91
```

Coding Coefficients for the Independent Variables

Factor	Subtracted off	Divided by
X1	0	1.414000
X2	0	1.414000

Response Surface for Variable Y: SABOR

Response Mean	5.952308
Root MSE	0.225852
R-Square	0.9744
Coef. of Variation	3.7944

Regression	Degrees of Freedom	Type I Sum of Squares	R-Square	F-Ratio	Prob > F
Linear	2	7.327631	0.5260	71.827	0.0000
Quadratic	2	5.782936	0.4151	56.685	0.0000
Crossproduct	1	0.462400	0.0332	9.065	0.0196
Total Regress	5	13.572967	0.9744	53.218	0.0000

Residual	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Prob > F
Lack of Fit	3	0.233944	0.077981	2.534	0.1954
Pure Error	4	0.123120	0.030780		
Total Error	7	0.357064	0.051009		

Parameter	Degrees of Freedom	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEPT	1	6.766014	0.101004	66.987	0.0000
X1	1	0.956987	0.079857	11.984	0.0000
X2	1	0.016372	0.079857	0.205	0.8434
X1*X1	1	-0.820034	0.085649	-9.574	0.0000
X2*X1	1	-0.340000	0.112926	-3.011	0.0196
X2*X2	1	-0.502438	0.085649	-5.866	0.0006

Parameter	Parameter Estimate from Coded Data
INTERCEPT	6.766014
X1	1.353180
X2	0.023151
X1*X1	-1.639573
X2*X1	-0.679795
X2*X2	-1.004573

Factor	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Prob > F	
X1	3	12.463817	4.154606	81.448	0.0000	ACUCAR
X2	3	2.219918	0.739973	14.507	0.0022	SAL

Canonical Analysis of Response Surface
(based on coded data)

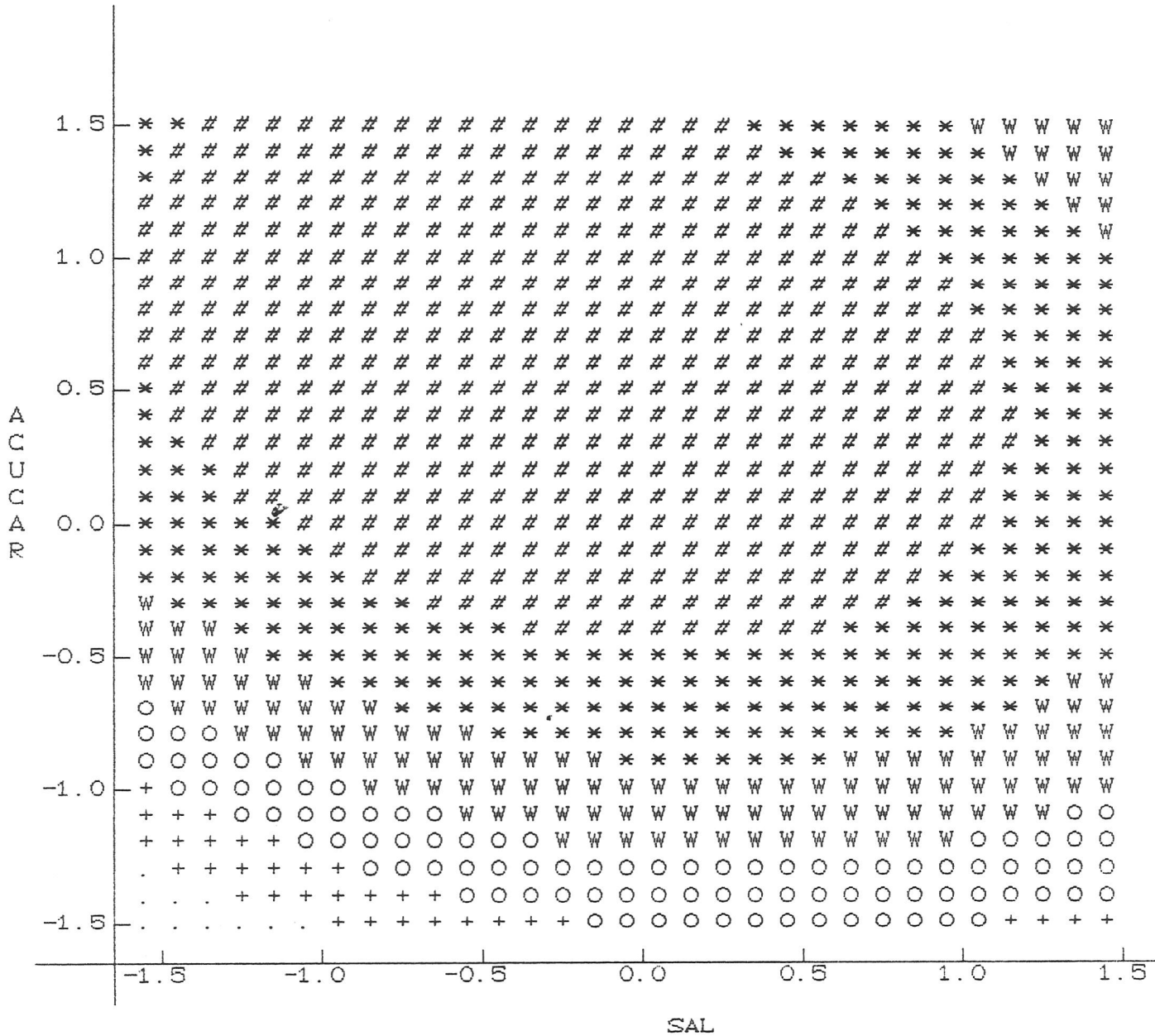
Factor	Critical Value		
	Coded	Uncoded	
X1	0.441222	0.623888	ACUCAR
X2	-0.137765	-0.194799	SAL

Predicted value at stationary point 7.062946

Eigenvalues	Eigenvectors	
	X1	X2
-0.856953	-0.398359	0.917229
-1.787193	0.917229	0.398359

Stationary point is a maximum.

Contour plot of X1*X2.



Symbol	Y	Symbol	Y
.....	1.5654 - 2.4816	WWWWW	4.3139 - 5.2301
+++++	2.4816 - 3.3977	*****	5.2301 - 6.1463
OOOOO	3.3977 - 4.3139	#####	6.1463 - 7.0624

6.3 - EXEMPLO 3

6.3.1 - Dados

Em um experimento, avaliou-se os efeitos de fósforo (ξ_1) em Kg/ha, e calcário (ξ_2) em t/ha, sobre a produção de uma mistura forrageira (Rhodes, Siratro e Desmodium). Utilizou-se o delineamento central composto rotacional com precisão uniforme e mediu-se as produções de Materia seca da mistura (y), em t/ha. Os dados para o exemplo aparecem na tabela 6.3.

Tabela 6.3 - Dados do exemplo 6.3

Variáveis Naturais		Variáveis Codificadas		Resposta
ξ_1	ξ_2	x_1	x_2	y
100	2	-1	-1	3.3
100	6	-1	1	4.5
300	2	1	-1	4.9
300	6	1	1	7.1
58.6	4	-1.414	0	2.5
341.4	4	1.414	0	5.8
200	1.172	0	-1.414	4.5
200	6.828	0	1.414	7.2
200	4	0	0	5.1
200	4	0	0	4.6
200	4	0	0	5.4
200	4	0	0	5.0
200	4	0	0	6.1

6.3.2 - Resultados

Os comandos do PROC RSREG do SAS para o exemplo 6.3 e os resultados obtidos são os seguintes :

```
DATA A;
INFILE 'A:EXEMPLO3';
INPUT X1 X2 Y;
LABEL
  X1="FOSFORO"
  X2="CALCARIO"
  Y="PRODUCAO DE MS";
PROC RSREG;
  MODEL Y=X1 X2 / LACKFIT;
  RIDGE MAX;
RUN;
DATA B;
  SET A END=EOF;
  OUTPUT;
  IF EOF THEN DO;
    Y=.;
    DO X1=-1.5 TO 1.5 BY .1;
      DO X2=-1.5 TO 1.5 BY .1;
        OUTPUT;
      END;
    END;
  END;
PROC RSREG DATA=B OUT=C NOPRINT;
  MODEL Y=X1 X2 / PREDICT;
DATA D;
  SET C;
PROC PLOT DATA=D;
PLOT X1*X2=Y / CONTOUR=6 HPOS=100 VPOS=36 HSPACE=10
      HAXIS=-1.5 TO 1.5 BY .5
      VAXIS=-1.5 TO 1.5 BY .5;
RUN;
```

```

-1      -1 3.3
-1      1 4.5
 1      -1 4.9
 1      1 7.1
-1.414  0 2.5
 1.414  0 5.8
 0 -1.414 4.5
 0  1.414 7.2
 0      0 5.1
 0      0 4.6
 0      0 5.4
 0      0 5.0
 0      0 6.1
```

Coding Coefficients for the Independent Variables

Factor	Subtracted off	Divided by
X1	0	1.414000
X2	0	1.414000

Response Surface for Variable Y: PRODUCAO DE MS

Response Mean	5.076923
Root MSE	0.431980
R-Square	0.9378
Coef. of Variation	8.5087

Regression	Degrees of Freedom	Type I Sum of Squares	R-Square	F-Ratio	Prob > F
Linear	2	16.340735	0.7780	43.784	0.0001
Quadratic	2	3.106093	0.1479	8.323	0.0141
Crossproduct	1	0.250000	0.0119	1.340	0.2850
Total Regress	5	19.696828	0.9378	21.110	0.0004

Residual	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Prob > F
Lack of Fit	3	0.054249	0.018083	0.0578	0.9794
Pure Error	4	1.252000	0.313000		
Total Error	7	1.306249	0.186607		

Parameter	Degrees of Freedom	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEPT	1	5.240006	0.193187	27.124	0.0000
X1	1	1.108442	0.152740	7.257	0.0002
X2	1	0.902361	0.152740	5.908	0.0006
X1*X1	1	-0.557653	0.163818	-3.404	0.0114
X2*X1	1	0.250000	0.215990	1.157	0.2850
X2*X2	1	0.292603	0.163818	1.786	0.1172

Parameter	Parameter Estimate from Coded Data
INTERCEPT	5.240006
X1	1.567338
X2	1.275939
X1*X1	-1.114970
X2*X1	0.499849
X2*X2	0.585030

Factor	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Prob > F	
X1	3	12.240054	4.080018	21.864	0.0006	FOSFORO
X2	3	7.358400	2.452800	13.144	0.0029	CALCARIO

Canonical Analysis of Response Surface
(based on coded data)

Factor	Critical Value		
	Coded	Uncoded	
X1	0.418362	0.591564	FOSFORO
X2	-1.269214	-1.794668	CALCARIO

Predicted value at stationary point 4.758144

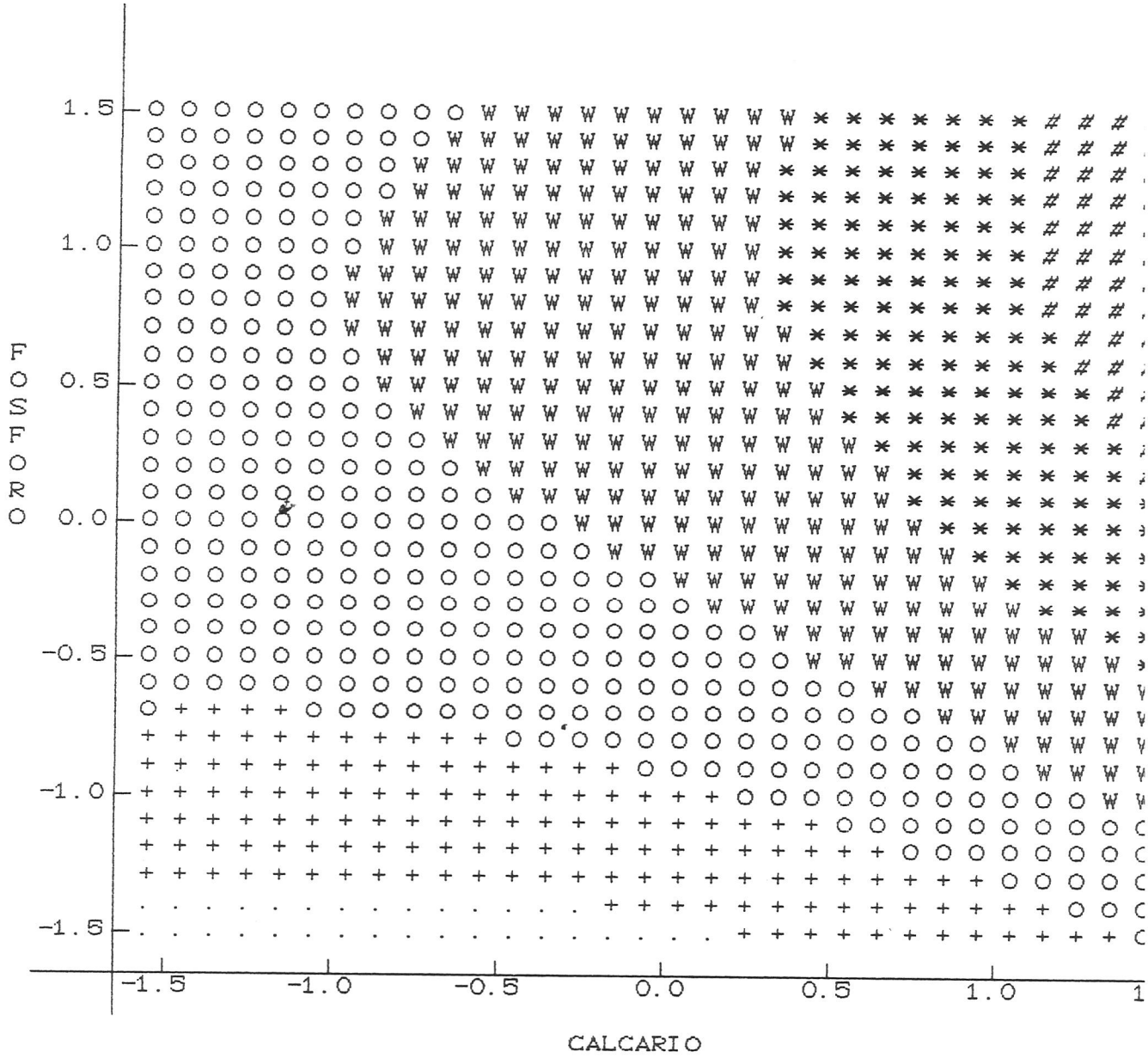
Eigenvalues	Eigenvectors	
	X1	X2
0.621011	0.142498	0.989795
-1.150951	0.989795	-0.142498

Stationary point is a saddle point.

Estimated Ridge of Maximum Response for Variable Y: PRODUCAO DE MS

Coded Radius	Estimated Response	Standard Error	Uncoded Factor Values	
			X1	X2
0.0	5.240006	0.193187	0	0
0.1	5.440936	0.192490	0.101681	0.098259
0.2	5.642308	0.190718	0.187679	0.211547
0.3	5.847617	0.188869	0.259958	0.335212
0.4	6.059397	0.188651	0.321325	0.465461
0.5	6.279385	0.192427	0.374386	0.599737
0.6	6.508751	0.202847	0.421198	0.736461
0.7	6.748290	0.222174	0.463285	0.874683
0.8	6.998549	0.251684	0.501752	1.013833
0.9	7.259918	0.291596	0.537405	1.153563
1.0	7.532676	0.341467	0.570834	1.293655

Contour plot of X1*X2.



Symbol	Y	Symbol	Y	Symbol	Y
.....	1.25 - 2.50	OOOOO	3.75 - 5.00	*****	6.25 - 7.50
+++++	2.50 - 3.75	WWWWW	5.00 - 6.25	#####	7.50 - 8.75

6.4 - EXEMPLO 4

6.4.1 - Dados

Frankel (1961) citado por MYERS (1971) discute um experimento realizado com o objetivo de maximizar a produção de mercaptobenzothiazole (MTB) variando o tempo (ξ_1) em horas, e a temperatura (ξ_2) em °C, de processamento. Os dados para o exemplo aparecem na tabela 6.4.

Tabela 6.4 - Dados do exemplo 6.4

Variáveis Naturais		Variáveis Codificadas		Resposta
ξ_1	ξ_2	x_1	x_2	y
4	250	-1	0	83.8
20	250	1	0	81.7
12	220	0	-1	84.7
12	280	0	1	57.9
6.3	229	-0.7	-0.7	81.3
6.3	271	-0.7	0.7	83.1
17.7	229	0.7	-0.7	85.3
17.7	271	0.7	0.7	72.7
12	250	0	0	82.4
12	250	0	0	82.9
12	250	0	0	81.2
4	250	0	0	82.0

As codificações utilizadas foram

$$x_1 = \frac{\xi_1 - 12}{8}, \quad x_2 = \frac{\xi_2 - 250}{30}$$

6.4.2 - Resultados

Os comandos do PROC RSREG do SAS para o exemplo 6.4 e os resultados obtidos são os seguintes :

```
DATA A;
INFILE 'A:EXEMPLO4';
  INPUT TIME TEMP MBT;
  LABEL
    TIME="REACTION TIME (HOURS)"
    TEMP="TEMPERATURE (DEGRESS CENTIGRADE)"
    MBT="PERCENT YIELD MERCAPTOTHIAZOLE";
PROC RSREG;
  MODEL MBT=TIME TEMP / LACKFIT;
  RIDGE MAX;
RUN;
DATA B;
  SET A END=EOF;
  OUTPUT;
  IF EOF THEN DO;
    MBT=.;
    DO TIME=4.0 TO 20.0 BY .2;
      DO TEMP=220 TO 280 BY 2;
        OUTPUT;
      END;
    END;
  END;
PROC RSREG DATA=B OUT=C NOPRINT;
  MODEL MBT=TIME TEMP / PREDICT;
DATA D;
  SET C;
PROC PLOT DATA=D;
  PLOT TIME*TEMP=MBT / CONTOUR=6 HPOS=100 VPOS=36 HSPACE=10
    HAXIS=220 TO 280 BY 10
    VAXIS=4.0 TO 20.0 BY 4;
RUN;
```

```
20.0 250 81.7
12.0 220 84.7
12.0 280 57.9
 6.3 229 81.3
 6.3 271 83.1
17.7 229 85.3
17.7 271 72.7
12.0 250 82.4
12.0 250 82.9
12.0 250 81.2
 4.0 250 82.0
 4.0 250 83.8
```

Coding Coefficients for the Independent Variables

Factor	Subtracted off	Divided by
TIME	12.000000	8.000000
TEMP	250.000000	30.000000

Response Surface for Variable MBT: PERCENT YIELD MERCAPTOBENZOTHAZOLE

Response Mean	79.916667
Root MSE	4.615964
R-Square	0.8003
Coef. of Variation	5.7760

Regression	Degrees of Freedom	Type I Sum of Squares	R-Square	F-Ratio	Prob >
Linear	2	313.585803	0.4899	7.359	0.024
Quadratic	2	146.768144	0.2293	3.444	0.100
Crossproduct	1	51.840000	0.0810	2.433	0.168
Total Regress	5	512.193947	0.8003	4.808	0.041

Residual	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Prob >
Lack of Fit	3	124.696053	41.565351	39.628	0.006
Pure Error	3	3.146667	1.048889		
Total Error	6	127.842720	21.307120		

Parameter	Degrees of Freedom	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob >
INTERCEPT	1	-545.867976	277.145373	-1.970	0.0964
TIME	1	6.872863	5.004928	1.373	0.2188
TEMP	1	4.989743	2.165839	2.304	0.0608
TIME*TIME	1	0.021631	0.056784	0.381	0.7164
TEMP*TIME	1	-0.030075	0.019281	-1.560	0.1698
TEMP*TEMP	1	-0.009836	0.004304	-2.285	0.0623

Parameter	Parameter Estimate from Coded Data
INTERCEPT	82.173110
TIME	-1.014287
TEMP	-8.676768
TIME*TIME	1.384394
TEMP*TIME	-7.218045
TEMP*TEMP	-8.852519

Factor	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Prob >
TIME	3	61.290957	20.430319	0.959	0.4704
TEMP	3	461.250925	153.750308	7.216	0.0205

Canonical Analysis of Response Surface
(based on coded data)

Factor	Critical Value		
	Coded	Uncoded	
TIME	-0.441758	8.465935	REACTION TIME (HOURS)
TEMP	-0.309976	240.700718	TEMPERATURE (DEGRESS CENTIGRAD)

Predicted value at stationary point 83.741940

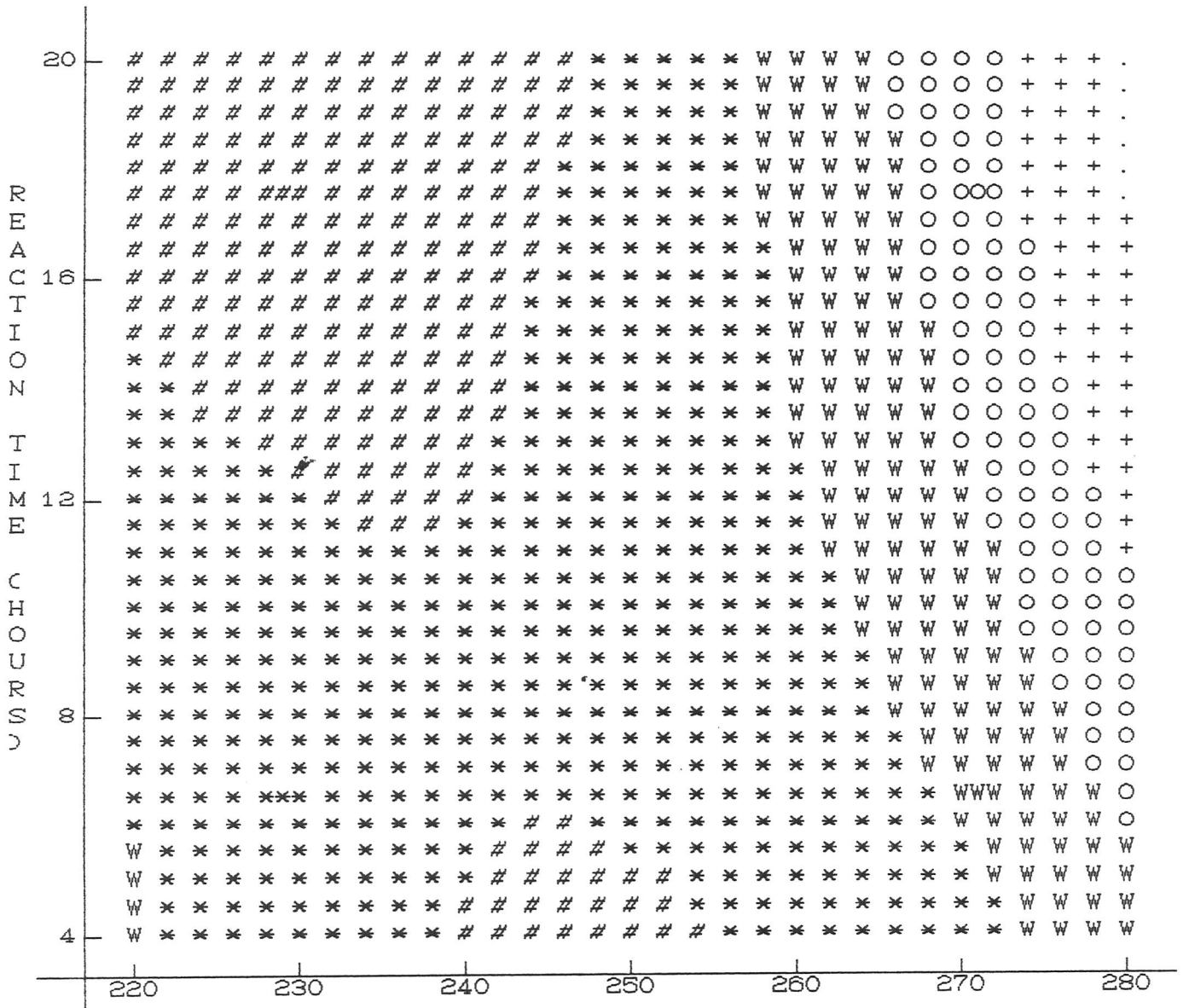
Eigenvalues	Eigenvectors	
	TIME	TEMP
2.528816	0.953223	-0.302267
-9.996940	0.302267	0.953223

Stationary point is a saddle point.

Estimated Ridge of Maximum Response for Variable MBT

Coded Radius	Estimated Response	Standard Error	Uncoded Factor Values	
			TIME	TEMP
0.0	82.173110	2.665023	12.000000	250.000000
0.1	82.952909	2.648671	11.964493	247.002956
0.2	83.558260	2.602270	12.142790	244.023941
0.3	84.037098	2.533296	12.704153	241.396084
0.4	84.470454	2.457836	13.517555	239.435227
0.5	84.914099	2.404616	14.370977	237.919138
0.6	85.390012	2.410981	15.212247	236.624811
0.7	85.906767	2.516619	16.037822	235.449230
0.8	86.468277	2.752355	16.850813	234.344204
0.9	87.076587	3.130961	17.654321	233.284652
1.0	87.732874	3.648568	18.450682	232.256238

Contour plot of TIME*TEMP.



Symbol	MBT	Symbol	MBT	Symbol	MBT
.....	54 - 60	OOOOO	66 - 72	*****	78 - 84
+++++	60 - 66	WWWWW	72 - 78	####	84 - 90

7. BIBLIOGRAFIA

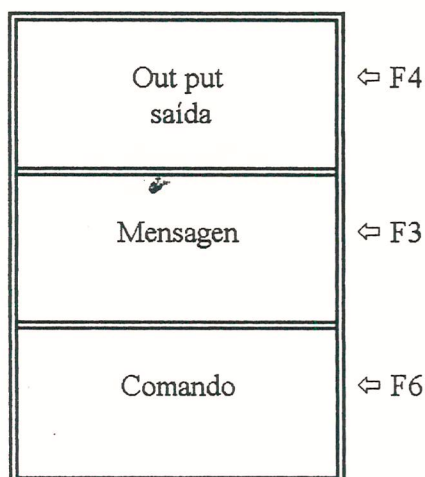
- BOX, G.E.P. & WILSON, K.B. On the Experimental Attainment of Conditions. *Journal of the Royal Statistical Society, série B*, 13 : 1 - 45. 1951.
- BOX, G.E.P. ; HUNTER, W.G. & HUNTER, J.S. *Statistical for Experimenters*. New York, John Willey, 1978. 653 p.
- BOX, G.E.P. & DRAPER, N.R. *Empirical Model - Building and Response Surfaces*. New York, John Willey, 1987. 669 p.
- COCHRAN, W.G. & COX, G.M. *Experimental Designs*. 2.ed. New York, John Willey, 1957. 611 p.
- DAVIES, O.L. *The Design and Analysis of Industrial Experiments*. London, Longman, 1956. 636 p.
- GOMES, F.P. *Curso de Estatística Experimental*. 11. ed., Piracicaba, Nobel, 1985. 466 p.
- JOHN, P.W.M. *Statistical Design and Analysis of Experiments*. New York, Macmillan, 1971. 386 p.
- KHURI, A.I. & CORNELL, J.A. *Response Surface Designs and Analysis*. New York, Marcel Dekker, 1987. 405 p.
- MONTGOMERY, D.C. *Design and Analisis of Experiments*. 3. ed., New York, John Willey, 1991. 649 p.
- MYERS, R.H. *Response Surface Methodology*. Boston, Allyn and Bacon, 1971. 431 p.

Anexo I - Instruções gerais sobre o uso do SAS

SAS

1. Composição Geral do SAS

O SAS é dividido em três janelas



Alguns Comandos do SAS :

F7 - Amplia a tela (zoom)

F9 - Chama o último programa que foi rodado (restaura na tela)

F10 - Roda o programa que está na tela (executa o programa)

F1 - Help (tecla de ajuda)

- Chamar um programa : **comand** : Include 'a: _____ . pro¹'
- Gravar : **comand** : File 'a: _____ . lis '
- Imprimir : salvar a saída e imprimir posteriormente no ws
- Para limpar a tela de saída Alt E ou **comand** : clear
- Sair do SAS **comand** : bye

¹ Esta extensão dos arquivos é opcional (usamos estas para facilitar)

2. Modo de utilização do SAS

- Cria-se um arquivo não-documento em um editor de textos qualquer, este arquivo conterá os dados a serem analisados.

Exemplo : EXEMPLO.DAT

1	-1	76.5
-1	1	77.0
1	-1	78.0
1	1	79.5
0	0	79.9
0	0	80.3
0	0	80.0
0	0	79.7
0	0	79.8
-1.414	0	78.4
-1.414	0	75.6
0	1.414	78.5
0	-1.414	77.0

- Cria-se um arquivo não-documento, que conterá o programa a ser rodado (este arquivo poderá ser feito em um editor de textos qualquer ou mesmo no próprio SAS).

Exemplo: EXEMPLO.PRO

```
DATA A;
INFILE 'A:EXEMPLO.DAT';
INPUT X1 X2 Y;
LABEL
  X1="TEMPO DE REAÇÃO"
  X2="TEMPERATURA DE REAÇÃO"
  Y="PRODUÇÃO";
PROC RSREG;
  MODEL Y=X1 X2 / LACKFIT;
DATA B;
  SET A END=EOF;
  OUTPUT;
  IF EOF THEN DO;
    Y=.;
    DO X1=-1.5 TO 1.5 BY .1;
      DO X2=-1.5 TO 1.5 BY .1;
        OUTPUT;
      END;
    END;
  END;
PROC RSREG DATA=B OUT=C NOPRINT;
  MODEL Y=X1 X2 / PREDICT;
DATA D;
  SET C;
PROC PLOT DATA=D;
PLOT X1*X2=Y / CONTOUR=6 HPOS=100 VPOS=36 HSPACE=10
  HAXIS=-1.5 TO 1.5 BY .5
  VAXIS=-1.5 TO 1.5 BY .5;
RUN;
```

- Criados estes dois arquivos vamos para o SAS

- Comandos no DOS

Exemplo:

C:\ CD SAS

C:\SAS\ SAS

- Uma vez dentro do SAS

O cursor do SAS começará piscando no terceiro módulo, o de comando, onde será chamado o programa.

Procedimentos

Comandos

Tecla: Home (para o cursor se descolar para a linha do **comand:** da janela de comando)

Home

Chame o programa (veja primeira página)

include 'a:² EXEMPLO.PRO'

Mande rodar o programa

F10

Tecla F4 para se deslocar para a janela das saídas

F4

Tecla F7 para obter um ampliamto da tela

F7

Para visualizar as saídas tecla

Page Up ou Page Down

Page Up

Page Down

Tecla Home caso o cursor não se encontre na linha do **comand:** da janela das saídas

Home

Grave as saídas

file 'a: EXEMPLO.LIS'

Saia do SAS (com o cursor no **comand:** de qualquer janela)

bye

² Caso os arquivos estejam gravados no Drive A

Publicações do Instituto de Matemática da UFRGS
Cadernos de Matemática e Estatística

Série B: Trabalho de Apoio Didático

1. Elsa Mundstock - Curso Básico Sobre Wordstar 3.45 - MAR/89
2. Jaime B. Ripoll - Introdução ao Cálculo Diferencial Via Funções de Uma Variável Real - OUT/89
3. Edmund R. Puczyłowski - Dimension of Modular Lattices - JUN/90
4. Marcos Sebastiani - Geometrias Não Euclidianas - JUL/90
5. Sandra R. C. Pizzatto - Cálculo Numérico - AGO/91
6. Vera Clotilde G. Carneiro - Elementos de Cálculo para Biologia - AGO/91
7. Elsa Mundstock - Iniciação ao SPSS/PC - SET/91
8. Elisa Hagg, Loiva C. de Zeni, Maria Alice Gravina e Vera Carneiro - Notas da 1ª Oficina de Matemática da UFRGS - JAN/92
9. Paulo Werlang de Oliveira, Elisabete Rambô, Suzana Lima dos Santos, Coordenação: Profª Maria Alice Gravina - A Tartaruga no Espaço Tridimensional - FEV/92
10. Silvio Possoli - Análise Multivariada - JUL/92
11. Dinara Westphalen Fernandez - Números Índices - OUT/92
12. Maria Teresinha Albanese - Coeficiente de Fidedignidade de um Instrumento de Medida - OUT/92
13. Vera Clotilde Carneiro e Sérgio Cláudio Ramos - Gráficos na Escola - DEZ/92

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
NÚCLEO DE ATIVIDADES EXTRACURRICULARES

Os Cadernos de Matemática e Estatística publicam as seguintes séries:

Série A: Trabalho de Pesquisa

Série B: Trabalho de Apoio Didático

Série C: Colóquio de Matemática SBM/UFRGS

Série D: Trabalho de Graduação

Série F: Trabalho de Divulgação

Série G: Textos para Discussão

Toda correspondência com solicitação de números publicados e demais informações deverá ser enviada para:

NAEC - NÚCLEO DE ATIVIDADES EXTRACURRICULARES
INSTITUTO DE MATEMÁTICA - UFRGS
AV. BENTO GONÇALVES, 9500 - PRÉDIO 43111
CEP 91509 - 900 AGRONOMIA - POA/RS
FONE: 336 92 22 OU 339 13 55 OU 228 16 33
RAMAL 6197
FAX: 336 15 12